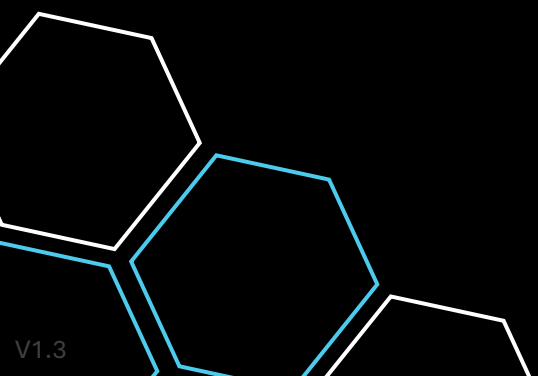




Copilot Studio

Implementation Guide



Success by Design



Success by Design is Microsoft's framework for Power Platform implementations

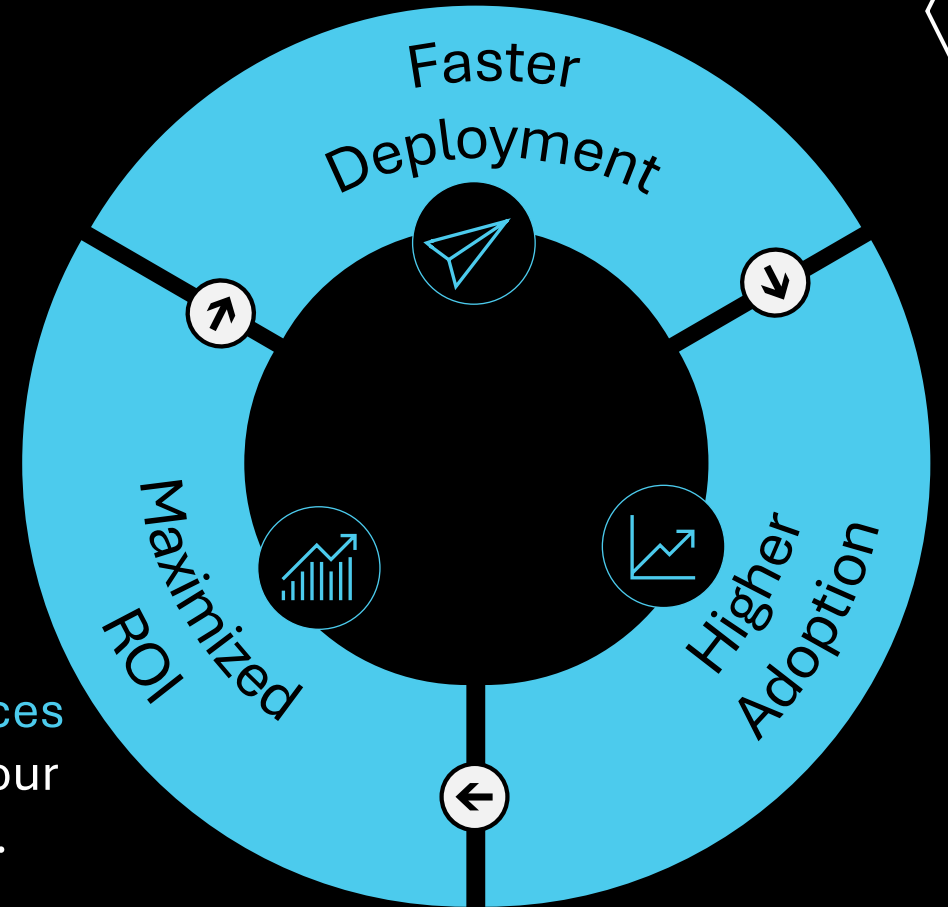


The framework is designed around 3 principles:

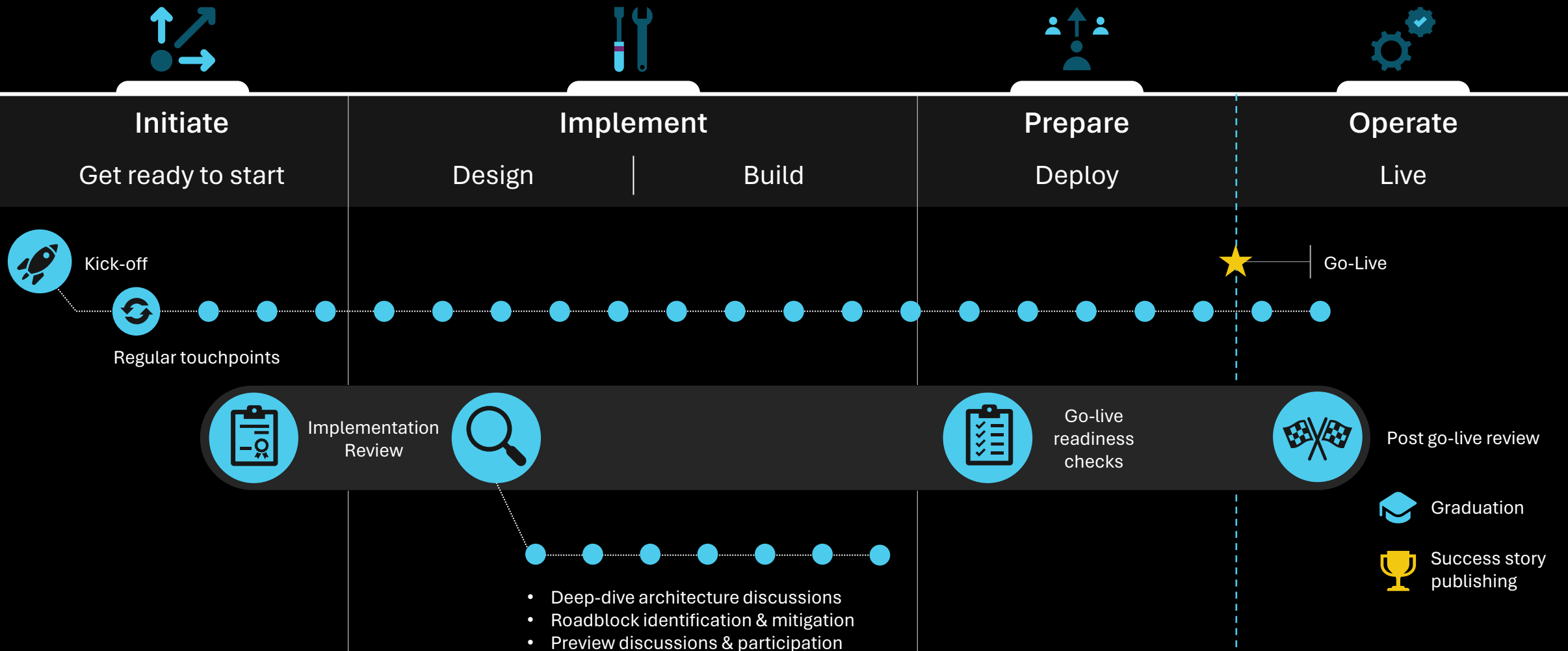
- ✓ Early Discovery
- ✓ Proactive Guidance
- ✓ Predictable Success



Success by Design brings the learnings and experiences from thousands of customer deployments to make your journey to the cloud smoother, faster and successful.



Typical Success by Design engagement lifecycle



Implementation Review



Agenda

- ✓ Project overview
- ✓ Architecture overview
- ✓ Language
- ✓ AI capabilities
- ✓ Integrations & channels
- ✓ Security, monitoring & governance
- ✓ Application lifecycle management
- ✓ Analytics & KPIs
- ✓ Fit-gap & top requests
- ✓ Dynamics 365 Omnichannel (optional)

The objective is to understand the various aspects of the project and provide recommendations, guidance, best practices, and highlight any risks or gaps in the plan. It is also an opportunity to discuss product roadmap and release alignment.



Attendees: key stakeholders from the customer and partner: project leads, solution architects, functional and technical leads.

Requirements: filling the implementation review document, at least partially. If needed, the content can be reviewed and discussed in multiple sessions throughout the implementation.

Out of scope: defining a solution architecture based on the review, providing product training, or performing code reviews.

Instructions for customers and partners

How to fill this document:

- ✔ Complete each slide providing the required information
- ✔ Post-it notes should be removed: their purpose is only to provide example answers.
- ✔ Skip and hide slides that are not applicable to your project.
- ✔ Add more slides to capture the required information if needed.
- ✔ Do not hesitate to paste your own slides and existing content.
- ✔ Reach out if you have questions on the format or content.
- ✔ Send the completed deck before the review workshops.
- ✔ Store the document in a shared SharePoint location if possible.
- ✔ Feedback on this guide? Share it [here](#)

Project Overview

-
- Business overview
 - Project overview & copilot purpose
 - Stakeholders
 - Project planning
 - Target KPIs
 - Assumption & concerns

Why *copilot*?

Copilots are the new chatbots

- ✔ Throughout this document, we refer to copilots instead of bots or chatbots. This reflects the new branding of Power Virtual Agents into Copilot Studio.
- ✔ Microsoft Copilot Studio is an end-to-end conversational AI platform that empowers you to create and customize copilots using natural language or a graphical interface. With Copilot Studio, you can easily design, test, and publish copilots that suit your specific needs for internal or external scenarios across your industry, department, or role.
- ✔ With more immersive UX and next-generation AI capabilities, copilots are the new chatbots.
- ▶ Watch: [Understanding Copilots at Microsoft \(video\)](#)

Business overview

- ✔ Provide an overview of your organization with a focus on your business unit

Example answer (you can delete this)



Contoso Corporation is a fictional but representative global manufacturing conglomerate with its headquarters in Paris.

2022 key figures:

- 40k employees
- 16 countries
- \$20B revenue in 2022

The **Global Modern Workplace** department is a team of 100 reporting to the CDO and is responsible for the modernization of workforce tools and processes across the company. One of the key principles of the modernization strategy is to use low code solutions and AI to empower employees and help them achieve more and be more productive.

Why do we ask these questions?

- Knowing more about your organization and business unit helps us better understand your use cases.
- Depending on your industry, we may have specific guidance or share similar implementation patterns.

Information we're looking for:

- Overview of your organization.
- Feel free to include logos and key metrics (number of employees, revenue, etc.).
- If the project applies to a specific subsidiary, business unit, department or team, please also describe it here.

Project overview & copilot purpose

- ✔ Provide an overview of the project and the business problems the copilot is trying to solve

Example answer (you can delete this)

Enterprise information is disseminated across the organization and maintained in silos that different teams are responsible for. An internal employee survey revealed a high dissatisfaction about the time spent to find answers internally.

The **Employee Copilot** is planned to be available to all 40k Contoso Corporation employees in Microsoft Teams as an app and in the global internal portal as a chatbot.

The copilot is expected to leverage AI and existing knowledge sources to answer employee questions on HR, IT, Sales and Finance topics. Answers must be grounded and sourced in enterprise data locations (SharePoint, ServiceNow, Dynamics 365, SAP).

Employees can also use the copilot to perform common actions such as unlocking an account or booking meeting rooms.

Why do we ask these questions?

- Understanding the project and what problem your copilot is trying to solve is key to make relevant recommendations or share similar implementation patterns.
- Knowing who your end-users are will also be useful to understand the copilot deployment channels requirements.

Information we're looking for:

- What are the current business challenges that this project will help address?
- What is the purpose and main features of the copilot?
- Who are the end-users of the copilot?
- Is this a migration project?

Key stakeholders



Customer, Partner and Microsoft teams

Customer team			
Role	Name	Email	Title
Business Sponsor	[Jane Doe]	[jane.doe@contoso.com]	[CIO]
Project Lead			
Product Owner			
Technical Lead			
Solution Architect			
Conversation Designer			
Tenant Admin			
Partner team (if applicable)			
Role	Name	Email	Title
Project Manager	[John Smith]	[john.smith@fabrikam.com]	[Program Lead]
Architect			
Developer			
Functional Consultant			

Why do we ask these questions?

- This helps us know who the key stakeholders are for the projects.
- The Microsoft team can also be described below.

Microsoft Team

Account Executive
[John Doe jdoe@microsoft.com]

Account Technology Strategist
[John Doe jdoe@microsoft.com]

Customer Success Account Manager
[John Doe jdoe@microsoft.com]

Sales Specialist
[John Doe jdoe@microsoft.com]

Technical Specialist
[John Doe jdoe@microsoft.com]

Cloud Solution Architect
[John Doe jdoe@microsoft.com]

Engineering Sponsor
[John Doe jdoe@microsoft.com]

Power CAT Program Manager
[John Doe jdoe@microsoft.com]

FastTrack Solution Architect
[John Doe jdoe@microsoft.com]

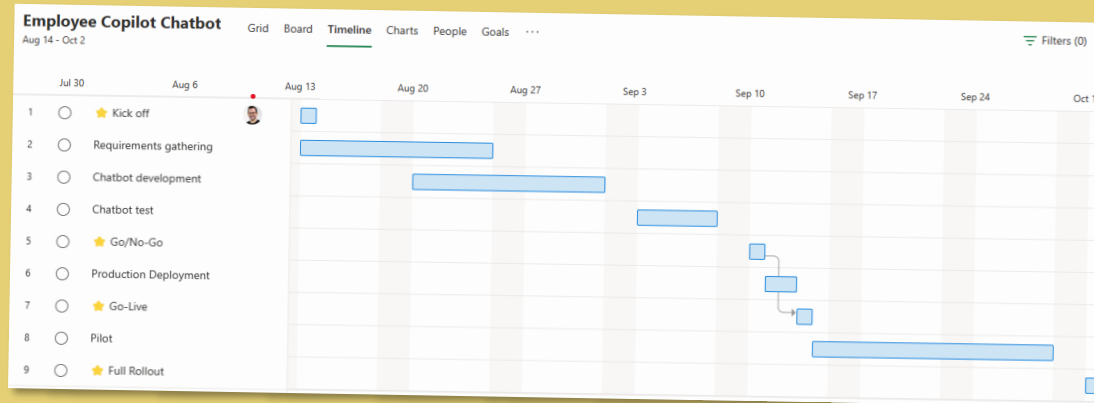
Planning

- ✓ What is the project timeline, from start to go-live, and the major milestones?

Example answer (you can delete this)

Kick off with business, technical and partner stakeholders on 8/1.

Phased rollout of the copilot to 1,000 users in a 2-week pilot before global deployment to all 40k employees.



Why do we ask these questions?

- Understanding the project key milestones and planned go-live dates will help setup meetings accordingly.

Information we're looking for:

- Please provide your project timeline, with target dates for major milestones.
- Feel free to paste your current project planning.

Target KPIs

✓ What are the success metrics for the project, and what are their targets?

Example answer (you can delete this)



60% deflection rate: end-user conversations leading to a resolution without any human intervention (today 20%).



4.5+ CSAT score: high satisfaction from end-users (today 3.0).



20,000 conversations / month: employees are expected on average to use the copilot once every two months (today 5,000).



\$1m. support savings/year: cost of ticket handling and human support avoided through automation.



Same-day updates and deployments: business subject matter experts can autonomously update, test and deploy the copilot.

Why do we ask these questions?

- To make relevant recommendations, we need to understand what success means for your copilots and project.

Information we're looking for:

- How is the copilot expected to help your end-users and lower the support cost?
- How is end-user satisfaction expected to increase with the copilot?
- Is there a specific return on investment you're expecting from this project?
- Are there other metrics you're looking to track?

Assumptions & concerns

✓ About the capabilities of your copilot:

What are your assumptions?

Example answer (you can delete this)

- Out-of-the-box deployment capabilities across all channels, including WhatsApp, Slack, and SharePoint intranet pages.
- Generative AI is expected to fully remove the need to maintain knowledge within the copilot itself.
- In the era of large language models (LLMs), state of the art intent recognition and entity extraction.

What are your concerns?

Example answer (you can delete this)

- Limited control of the answers generated by a large language model (LLM).
- Internal employee pushback about copilots replacing or reducing customer service teams.
- Some IT teams would prefer to host the conversational AI software and LLMs on internal servers.

Why do we ask these questions?


- Understanding your assumptions and concerns to help you either validate or address them early in the project.

Architecture Overview


-
- Solution architecture
 - Conversation volumes
 - Performance
 - Identified technical challenges

Copilot Studio Overview


✔ Product capabilities



Post to multiple channels



IVR/
Telephony



Extend Copilot for
Microsoft 365

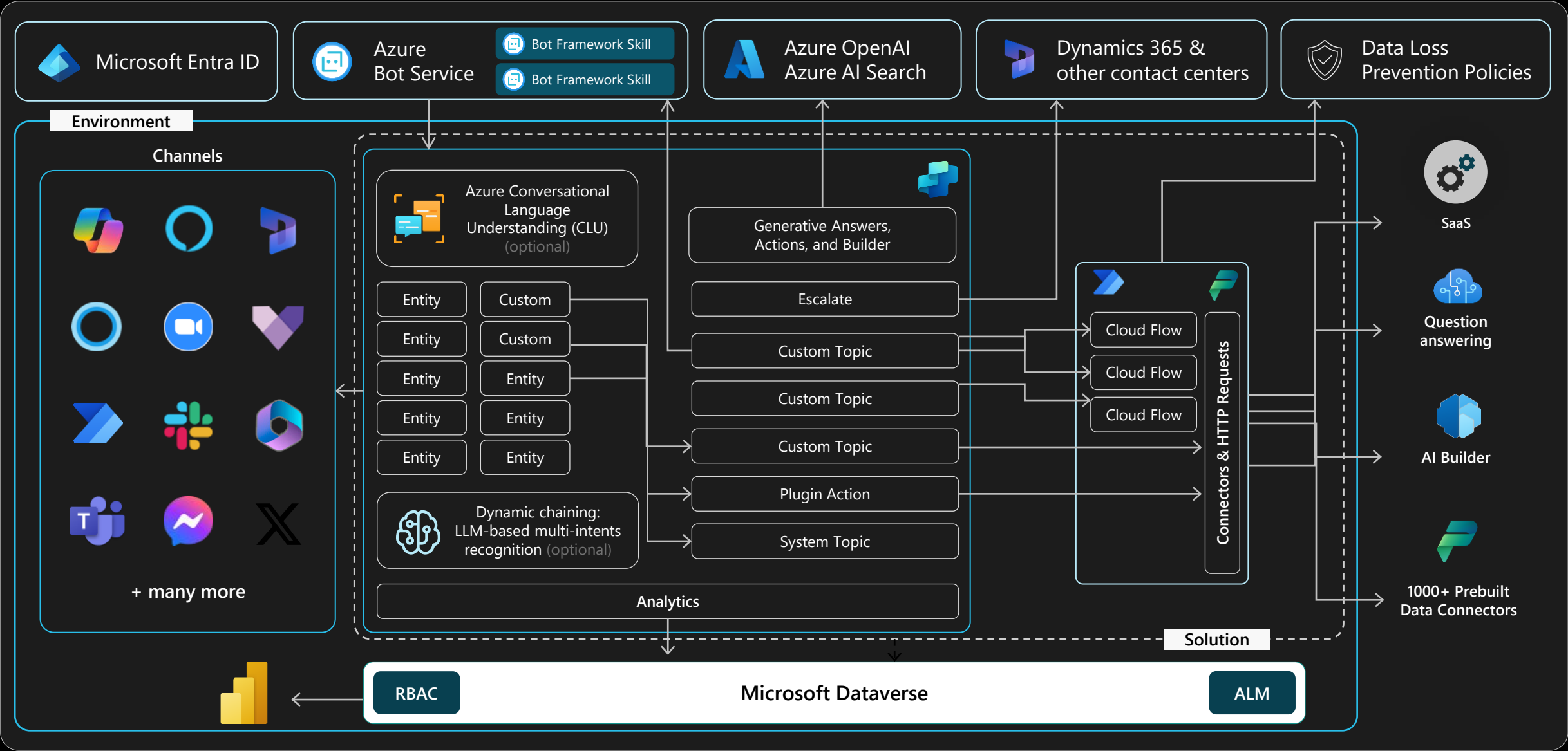
World Class UX	Declarative	Rich Response	Intuitive	Ease of use	Copilot Assistance
Logic + Automation	Power Fx	Generative Answers + Actions		Power Automate	Bot Framework skills
Next-Gen AI	Prebuilt LLM	Generative AI	Proactive suggestions		Bring your own model (Azure CLU)
Connected experience	Microsoft Teams	Dynamics 365	1000+ Connectors		Azure Bot Service
Copilot Lifecycle	Test pane	Collaboration	Solution management		ALM Automation
Security + Governance	Trusted identity	Full visibility	Advanced RBAC		Granular DLP control

Microsoft's end-to-end copilot building platform

Why do we share this example?

- Overview of the native capabilities of Copilot Studio to help you assess the features you're planning to use in the project.

Copilot Studio architecture



Copilots and Conversational AI

Microsoft 365
Conversational experiences
for Microsoft 365

Copilot for
Microsoft 365



Biz Apps & Power Copilots
Conversational experiences for
Dynamics 365 and Power Platform
products

Sales Copilot

Service
Copilot

Copilot for
Power Platform



**Other Microsoft
Copilots**

...



Custom Copilots
Custom copilots and bots
for enterprises and third parties

Custom
Enterprise
Copilots



Microsoft Copilot Studio
Extend and customize 1st party copilots | Build custom Copilots

Bot
Framework
/ SDK

Bot Service
Channels

Azure AI
Studio

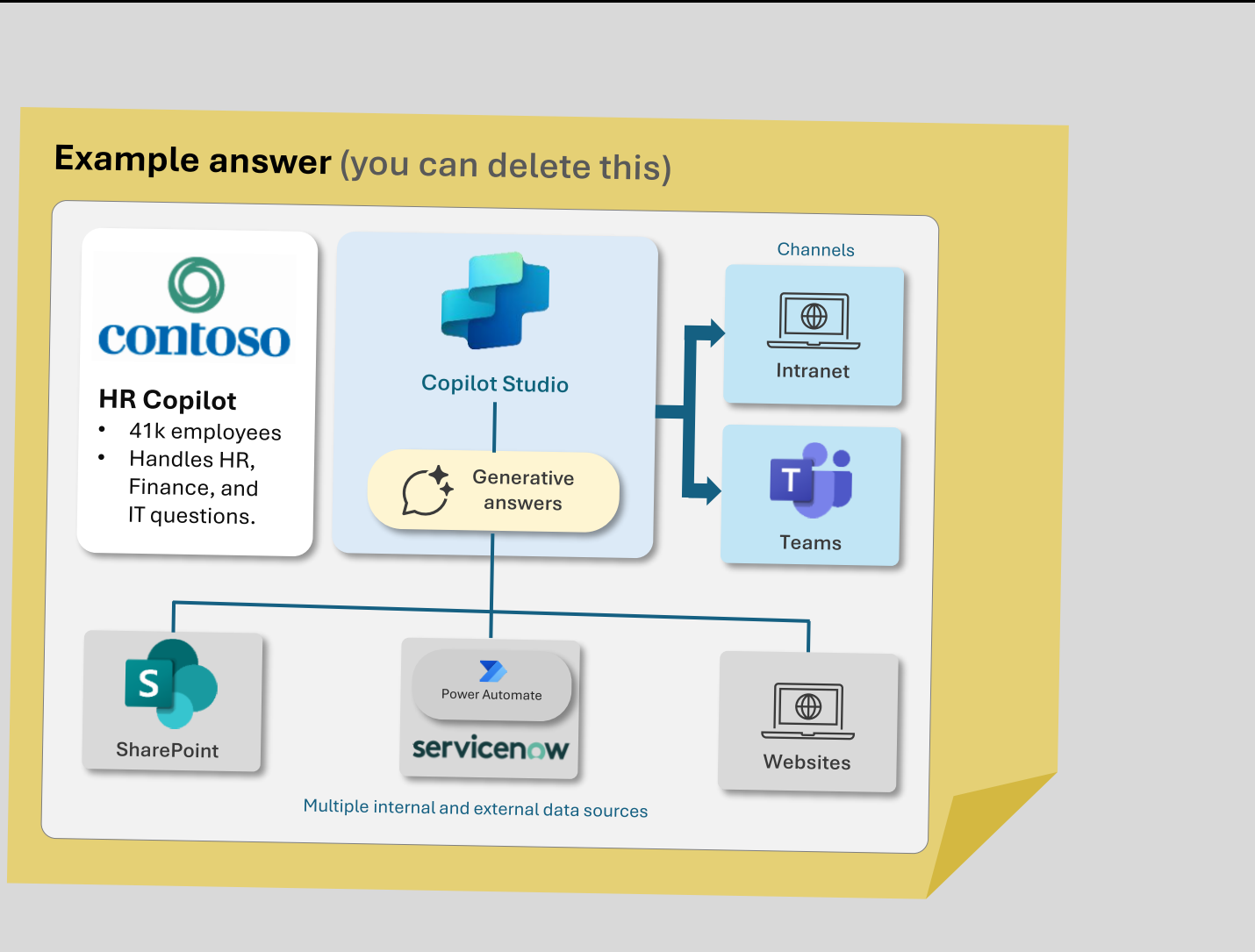
Azure AI
Search

Power
Platform
Connectors

AI Builder

Solution architecture

- ✓ Provide an architecture diagram with the overview of the technical implementation



Why do we ask these questions?

- To have the big picture of your copilot project and its integration in a broader technical and functional landscape.

Information we're looking for:

- Architecture diagrams or blueprints
- Functional and technical overview of the copilot and its integration in the broader technical landscape (other Power Platform or Azure services, internal APIs, etc.)

Conversation volumes

✔ What are the expected volumes of chat messages or conversations?

Monthly target volume of messages or conversations and growth

Example answer (you can delete this)

Initially 200,000 chat sessions per month (with an average of 8 messages per session) when deploying to the first channel and growing to 400,000 when deployed over all channels.

Expected seasonality impact

Example answer (you can delete this)

800,000 monthly chat sessions in December and January.

Expected maximum peak of concurrent conversations

Example answer (you can delete this)

Maximum of 5,000 concurrent chats when sales season begins.

Why do we ask these questions?

- Understanding target volumes helps validate the target architecture and scale.
- When integrating with external systems, for example through Power Automate or HTTP requests, it's important to validate that every part can handle the load.
- Target volumes also help validate the licensing aspects of the copilot and the potential impact on Dataverse storage for the conversation transcripts.

Useful resources:

- [Rate limits for copilots](#)
- [Pricing plans](#)
- [Microsoft Power Platform Licensing Guide](#)

Performance

✔ Detail the expected copilot response times

First chat load time and first message expectation

Example answer (you can delete this)

5 seconds for the web chat control to load and for the initial messages to be displayed to the user.

Expected maximum latency for the copilot to answer user queries

Example answer (you can delete this)

3 seconds after a user provides an input.

Approach for handling long-running actions (e.g., waiting for an external system to return data)

Example answer (you can delete this)

A specific message is displayed to the user asking them to wait while the cloud flow runs and returns data. Typing indicator shows that the copilot is doing something.

Why do we ask these questions?

- While there are no guaranteed service level agreements (SLAs) for Copilot Studio response times – mainly because they depend on complexity, integration patterns, etc. – it's still important to understand performance expectations for the copilot response times to make sure the copilot design optimizes for performance.

Useful resources:

- [Capture telemetry with Application Insights](#)

Identified technical challenges

✓ Please share the list of technical challenges or roadblocks you have already identified

Challenge #1

Example answer (you can delete this)

Connecting to on-premises resources that are not exposed to the public internet.

Challenge #2

Example answer (you can delete this)

Deploying to a WhatsApp channel.

Challenge #3

Example answer (you can delete this)

Allowing copilot end-users to download the conversation transcript at the end of the chat session.

Why do we ask these questions?






- To understand the challenges that have already been identified in the architecture that could get in the way of the copilot success and for which it's important to define mitigation plans.

Language

- Natural Language Understanding
- Intent recognition and slot filling
- Azure CLU integration
- Topic structure
- Handling unrecognized intents
- Localization & languages
- Miscellaneous

Natural Language Understanding

Choosing the right option for intent recognition and entity extraction

	 Standard NLU model	 Custom Azure CLU model	 Dynamic chaining
	<ul style="list-style-type: none"> ✓ Default, out-of-the-box, model that comes pre-trained, with many predefined entity types. ✓ Configuration is done by adding trigger phrases and custom entities (either closed lists with values and synonyms, or regular expressions). 	<ul style="list-style-type: none"> ✓ Supports additional languages, with native models. ✓ Allows to further customize the intent triggering model for better intent recognition or to address specific industry requirements. ✓ Allows for complex entity extraction (e.g., of the same type). ✓ Entity extraction can also leverage Copilot Studio standard NLU. 	<ul style="list-style-type: none"> ✓ Uses a GPT large language model. ✓ Can handle multiple intents and chain topics and/or plugins. ✓ Automatically generate questions for missing inputs and answers complex entities and questions from the conversation context. ✓ Configuration is done by simply describing topics, plugin actions, and inputs/outputs.
	<ul style="list-style-type: none"> ✓ Single intent recognition per query. ✓ Cannot be extended. ✓ Slot-filling multiple entities of the same type in the same query requires disambiguation for each (e.g., from and to cities) 	<ul style="list-style-type: none"> ✓ Single intent recognition per query. ✓ Configuration is done in Azure and involves additional costs. ✓ Has its own service limits that need to be evaluated. ✓ Azure CLU intents and Copilot Studio topics must be carefully kept in sync. 	<ul style="list-style-type: none"> ✓ As it's a generative AI feature, the licensing burn rate of messages is higher than in regular topic triggering. ✓ Preview capability.

Why do we share this?

- NLU (Natural Language Understanding) is a subfield of NLP (Natural Language Processing) that specializes in the machine's ability to comprehend and make sense of human language in a valuable way, focusing on understanding context, sentiment, and intent.
- In Copilot Studio, topic or action triggering can be done in different ways: customers can choose to override the standard NLU model with Azure CLU (Conversational Language Understanding) – the modern equivalent of LUIS – or can be fully replaced with dynamic chaining, a GPT LLM-based one.
- Azure CLU is a feature of Azure AI Language.

Useful resources:

- [Trigger phrases best practices](#)
- [Slot filling best practices](#)
- [Copilot Studio supported languages](#)
- [Azure CLU supported languages](#)
- [Azure CLU limits](#)
- [Azure CLU pricing](#)
- [How to use CLU as your NLU \(video\)](#)

Intent recognition and slot filling

✔ Leveraging native capabilities for better intent recognition and efficient conversations

How are you planning to define the topics trigger phrases?

Example answer (you can delete this)

From an existing FAQ base and from agent chat transcripts.

What is your expected use of entities (prebuilt, closed list, regular expressions).

Example answer (you can delete this)

Regular expressions to identify an order ID, prebuilt entity for email, and closed list for operation type and synonyms for each.

Do you plan to test topic triggering with test utterances?

Example answer (you can delete this)

Yes, bulk testing to make sure the relevant topics are triggered based on a set of test user utterances.

Why do we ask these questions?

Trigger phrases:

- Trigger phrases train your copilot's natural language understanding (NLU) model.
- It's important to appropriately configure them to make sure that the right topic triggers for the user utterances.
- When a copilot is uncertain what topic to trigger, it can suggest up to 3 potential topic candidates (Multiple Topics Matched topic) or go to Fallback if no topic is identified.

Entity extraction and slot filling:

- Slot filling relates to the use of entities that let a copilot acquire and use information more easily by identifying and extracting them from the user query.
- For example, when a user asks: *"I'd like to order 3 large blue t-shirts"*, the Natural Language Understanding (NLU) can immediately understand:
 - ✓ Topic is **Order**.
 - ✓ Quantity is **3**.
 - ✓ Color is **Blue**.
 - ✓ Size is **Large**.
 - ✓ Item Type is **T-Shirt**.

And the related questions be skipped.

Useful resources:

- [Trigger phrases best practices](#)
- [Slot filling best practices](#)

Azure CLU integration

✔ Overriding the standard NLU model with Azure Conversational Language Understanding

Do you plan to use Azure CLU? If yes, why? What benefits are you expecting?

Example answer (you can delete this)

Using Azure CLU to support additional languages, industry specific dictionaries and complex entity extraction.

Are you also doing custom entity extractions with Azure CLU? Why?

Example answer (you can delete this)

To do silent or lucky slot filling without having to ask questions and to handle unsupported scenarios in Copilot Studio such as source and destination city in the same phrase.

How are you making sure the related Azure CLU service quotas and limits match the use of the copilot?

Example answer (you can delete this)

Azure CLU is setup using the S tier. Less than 1,000 calls requiring intent recognition are expected per minute.

Why do we ask these questions?

- You can integrate a Conversational Language Understanding (CLU) model with a Copilot Studio copilot.
- Azure CLU intents must be mapped with Copilot Studio topics.
- Copilot Studio prebuilt entities can be used in conjunctions with Azure CLU entities.
- Using Azure CLU requires separate Azure resources and involves separate costs.

Useful resources:

- [Azure CLU integration overview](#)
- [Azure CLU best practices](#)
- [Azure CLU limits](#)
- [Azure CLU pricing](#)

Topic structure example for disambiguation

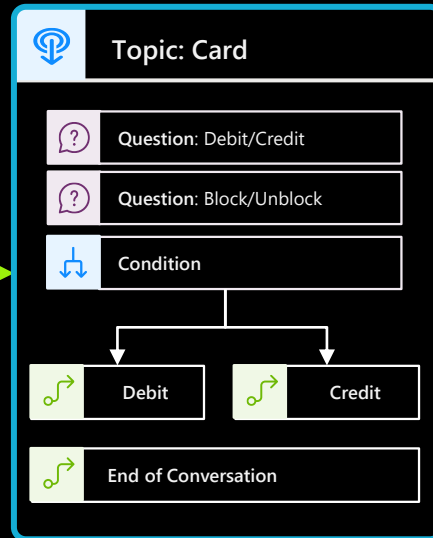
✓ Mixing topics triggered by trigger phrases and redirected topics

Natural Language Understanding
Trigger phrases

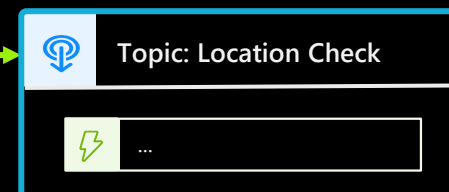
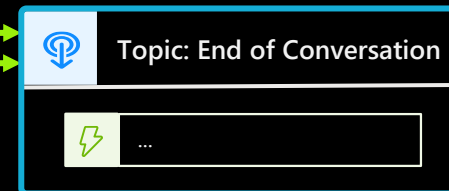
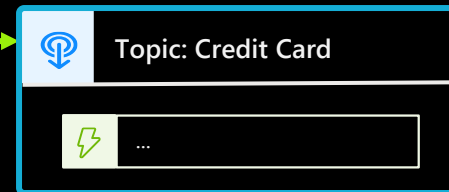
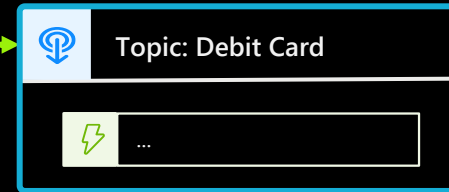
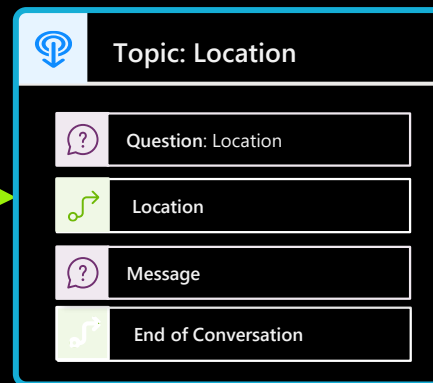
"Trigger Phrases" Topics
Triggerred from trigger phrases or redirects

Reusable, Bite-Size, Topics
Explicitly called by topics

- Issue with my card
- Block card
- I need to block my credit card
- Need help blocking debit card
- Unblock card
- Need to unblock credit card
- How to unblock my debit card



- Find my nearest bank
- Check branch location
- Find a bank
- Find me your nearest location
- Banks near me



Why do we share this example?

This example showcases the use of a disambiguation topic. It's very useful when it's hard or impossible to be certain of triggering the right topic if the user doesn't provide enough details in their query. This also prevents triggering a multiple topics matched topic that sometime confuse users.

- In this example, trigger phrases are associated to two main topics, that then break down their logic into multiple sub-topics that are called with redirect actions.
- Some topics can be called by multiple other topics: that's typically the case with the **End Of Conversation** topic.
- Thanks to slot filling and entity extraction, if a user says "*I need to unblock my credit card*", the **Card** topic will get triggered and both **Debit/Credit** and **Block/Unblock** questions will be skipped, as the **Card Type** and **Operation Type** will be deduced from the trigger phrase. That way, the appropriate **Credit Card** child topic will automatically be called, without the user providing any additional input or even noticing they went through the disambiguation topic.

Useful resources:

- [Topics best practices](#)

Topic structure

✓ Creating and designing efficient topics

What is your approach to designing topics?

Example answer (you can delete this)

A few custom topics for key scenarios with relevant trigger phrases and redirects, with a parent-child topic structure. Unrecognized intents will trigger generative answers and fallback.

Do you plan to use disambiguation topics?

Example answer (you can delete this)

Yes, for user account operations with clarifying questions on the operation (e.g. create, unlock, suspend, etc.) and system (e.g. SAP, ServiceNow, Microsoft, etc.)

How are you avoiding duplicating topic content multiple times?

Example answer (you can delete this)

Whenever a dialog path needs to be repeated, creating reusable topics that can be called by a parent topic before resuming the parent topic conversation logic once the child topic is complete.

Why do we ask these questions?

- Topics are discrete conversation paths that, when used together, allow for users to have a conversation with a copilot that feels natural and flows appropriately.
- While there's no one size fits all, given how topics can be triggered, it's a good practice to distinguish between:
 - ✓ **Topics that will trigger based on user utterances.** These can almost be seen as your entry points topics. If you have trigger phrases that overlap multiple topics, consider having a catch-all topic and then redirect to other topics after clarifying questions. With entity extraction and slot filling, clarifying questions can be skipped if already answered.
 - ✓ **Topics that will trigger when called from a redirect action, activity or event.** These can be called by multiple topics and can have input and output variables. They're ideally reusable, bite-size, topics.
 - ✓ **A topic can also be both,** triggered through intent recognition or by an explicit redirect.
 - ✓ **Conversational boosting and fallback:** topics that trigger when no matching topic is triggered based on a user query.

Useful resources:

- [Topics best practices](#)

Handling unrecognized intents

✔ Answering for unplanned user queries

How are you planning to manage unrecognized intents?

Example answer (you can delete this)

First by going to the Conversational Boosting topic to look for the answer in corporate resources (SharePoint sites and public websites) and if no result to a fallback ChatGPT experience.

Are you integrating with an external system as part of fallback? If yes, how?

Example answer (you can delete this)

The Azure OpenAI GPT-4 model is called through an HTTP request. It uses a custom system prompt to offer a compliant ChatGPT-like experience.

What is the expected % of conversations hitting fallback?

Example answer (you can delete this)

25% initially, and these are regularly reviewed to enrich existing topics or create new ones.

Why do we ask these questions?

- The Fallback topic gets triggered when Copilot Studio doesn't understand a user utterance and doesn't have sufficient confidence to trigger any of the existing topics.
- If Generative Answers are enabled on the copilot, the Conversational Boosting topic also triggers on the unknown intent event and triggers before the Fallback one.
- There are many ways to handle unrecognized intents: using Generative Answers to look for the answer on various data sources and/or using the Fallback topic to integrate with other systems. For example, question answering in Azure Cognitive Service for Language allows you to offload large volumes of question-and-answer pairs. It also has a chitchat model to handle random questions to the copilot.
- While it's important to leverage the Conversational Boosting and Fallback capabilities, it's also important to make sure that the core scenarios and topics of your copilots are properly handled through custom topics and their outcomes defined (resolved, etc.).

Useful resources:

- [Generative Answers](#)
- [Using the Fallback topic](#)

Localization & languages

✓ What languages should your copilot speak?

What languages and markets should your copilot support?

Example answer (you can delete this)

Primarily US English (70% of end-users), Spanish (20%), French, Portuguese (Brazilian) and Dutch (Belgium Flemish).

Do you plan for a single copilot multi-languages copilot, or for one copilot per language?

Example answer (you can delete this)

A single copilot enabled for multiple languages, with condition logic based on the end-user language so that some topics are only available for specific regions.

Should the translations be set during configuration or real-time (i.e., provided at runtime)?

Example answer (you can delete this)

Not yet, but this could be considered to deploy to more markets more rapidly.

Why do we ask these questions?

- There are multiple ways to address copilot localizations. Understanding your requirements will help make better architecture recommendations.
- The main approaches are:
 - ✓ One copilot per language.
 - ✓ One copilot for multiple languages, with translations provided as part of the copilot configuration. Translations need to be updated each time the copilot is updated or when new content is added.
 - ✓ One copilot for multiple languages, with translations provided real-time, at runtime, through a relay copilot sitting between the user and the copilot. This allows deploying more languages rapidly, but it also adds a dependency on a relay copilot and a real-time translation layer (e.g., Azure Service Copilot and Azure Cognitive Services Translator).

Useful resources:

- [Supported languages](#)
- [Azure CLU integration](#)
- [Real-time multilingual copilot sample](#)
- [Configure multilingual chatbots](#)
- [How to make multilingual copilots \(video\)](#)
- [Make generative answers multilingual \(video\)](#)

Miscellaneous

✓ Miscellaneous conversation design questions and requests

Do you have specific requirements on session timeout? (i.e., how long a copilot session can last)

Example answer (you can delete this)

We need chat sessions to last up to 7 days on the Teams channel.

Do you have other requirements or questions on conversation design?

Example answer (you can delete this)

No, the copilot business subject matter experts in the project all trained and certified with Conversation Designer Certification from the Conversation Design Institute.

Why do we ask these questions?

- Sessions can be configured to last for more than the default 30-minute by using the inactivity trigger.

Useful resources:

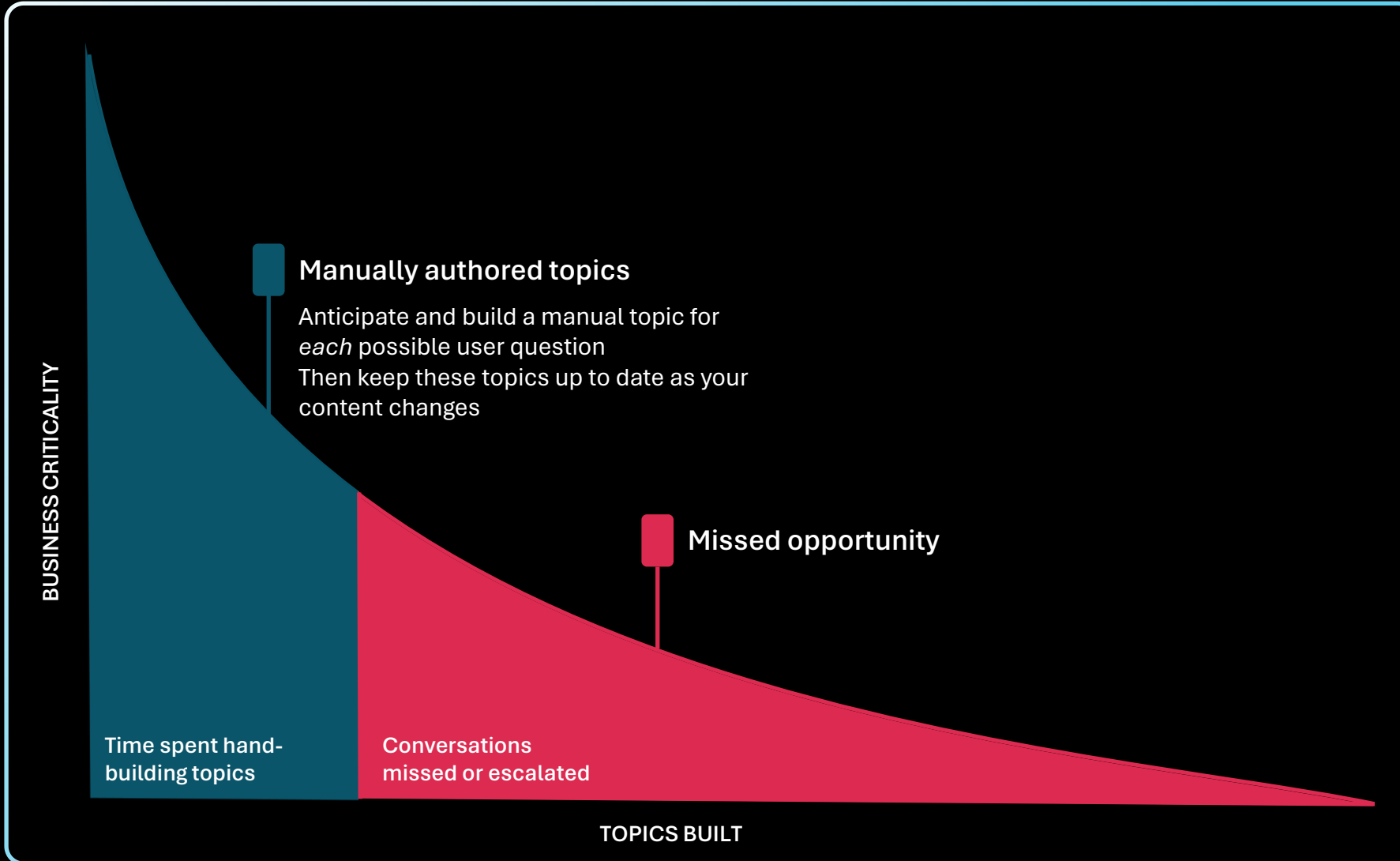
- [Inactivity trigger](#)
- [Conversation Design Institute](#)

AI capabilities

- Generative AI is changing conversational AI
- Generative answers
- Generative actions
- Generative builder
- Other AI requirements

Generative AI is changing conversational AI

✓ Before generative AI

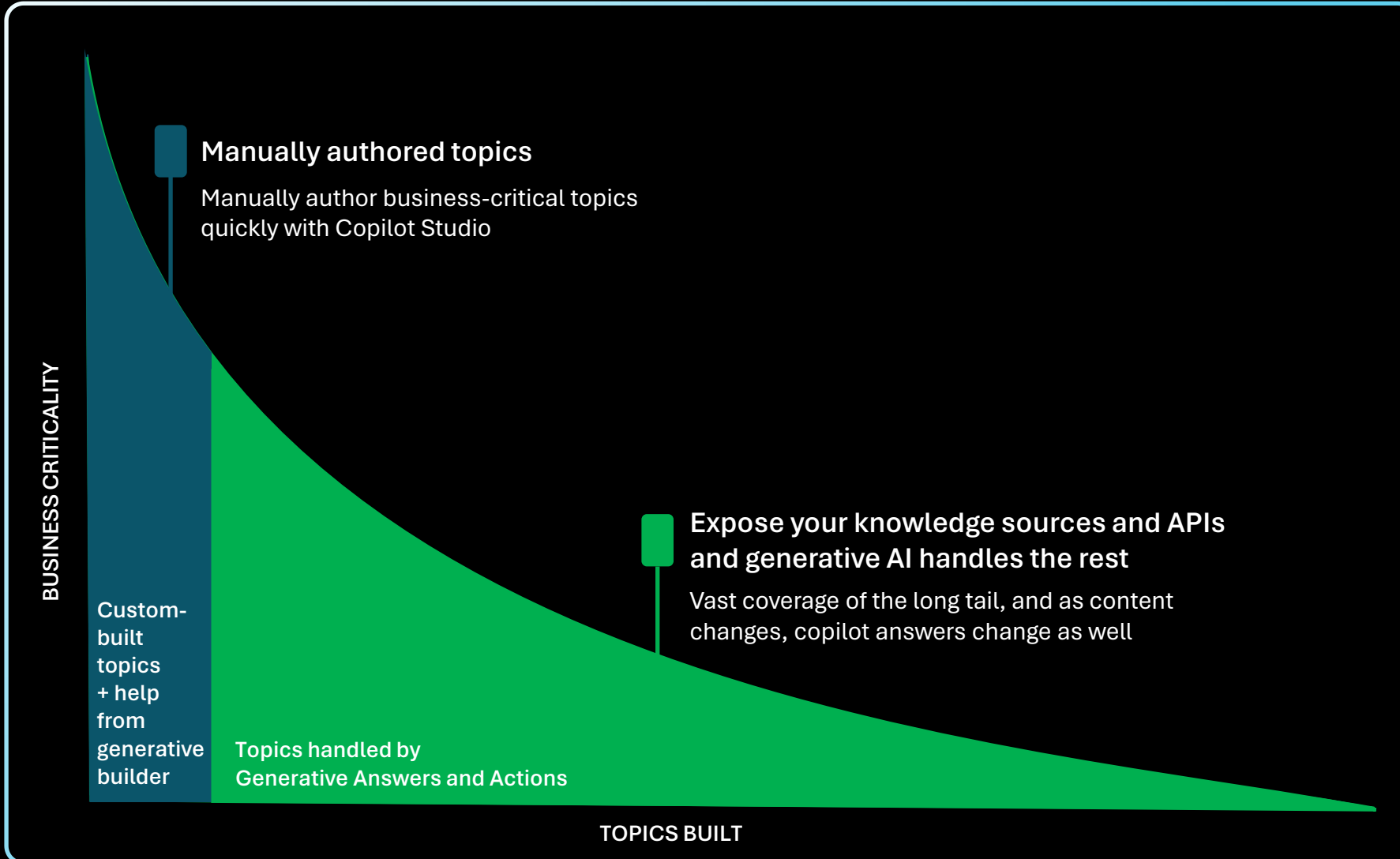


Why do we share this example?

- Before generative AI, companies used to create every topic their bots needed by hand, and they could only create so many, perhaps a few hundreds – and then they had a large cliff of unsupported things a user couldn't talk about.
- These then would create user frustration with the bot saying "sorry I didn't understand you", or large costs by having it just to dump the user to a human agent.

Generative AI is changing conversational AI

✓ After generative AI



Why do we share this example?

- Now, with Generative AI, companies are increasingly only building a few, business critical topics where they want tight control over the dialog.
- And for everything else, they're just pointing to knowledge sources and allowing the bots to talk over those corpuses.
- That means vastly increased support coverage for anything a user could ask, and far less topics that need to be handled by human agents.
- But it also means that maintaining topics is far less onerous - as the information changes, the copilot answers change automatically.
- There are no dialogs to continually update and keep relevant. So, this dramatically reduces costs.

Generative answers considerations

Generative answers processes and data sources

1

Relevant text content retrieval from configured sources

2

Summarization using Azure OpenAI

- Custom instructions
- Past 10 turns for contextualization
- References
- Content summary

3

Response Validation

- Native Azure OpenAI moderation layer
- Responsible AI moderation layer
- Grounding validation / hallucination removal

Public data

- Website content must be indexed by Bing.
- Bing is a global service, no geo boundary.

Public Website	<ul style="list-style-type: none"> Up to 4 URLs, but the list can be dynamic. Max 2 subpages depth (/en/help/), no direct pages.
Bing Custom Search	<ul style="list-style-type: none"> One configuration ID, but can be set using a formula. Azure costs covered by Microsoft Up to 400 URLs, custom ranking options. Max 2 subpages depth (/en/help/), supports direct pages.

SharePoint Internal only

- Requires the user to be authenticated with Entra ID to make delegated calls to the Graph API.
- Up to 4 SharePoint or OneDrive for Business sites (but the list can be dynamic).
- 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 OpenAI on your data

- Returns results from the linked Azure AI 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.

Uploaded documents

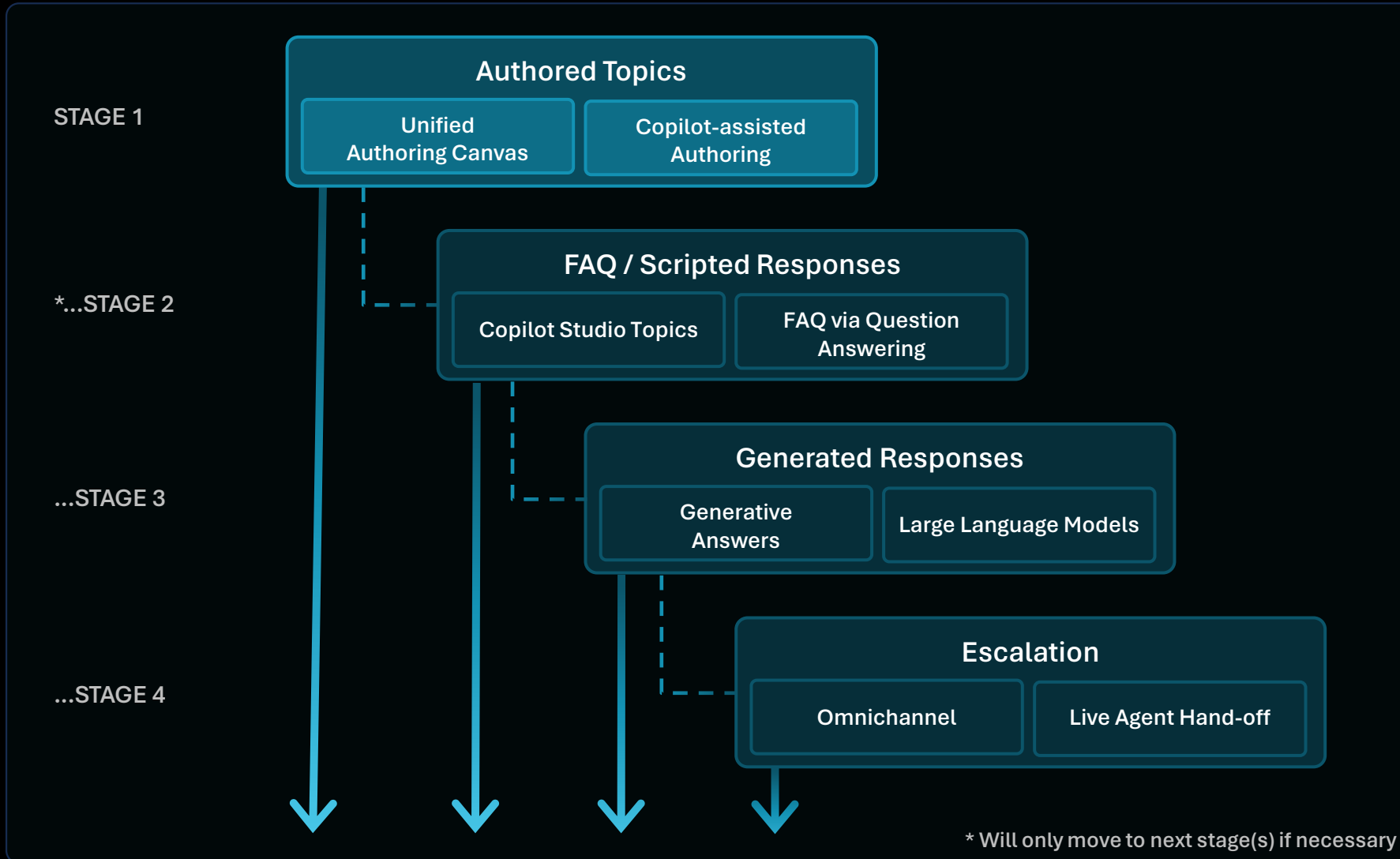
- Files (max 3 MB) are stored in Dataverse file storage and indexed with Dataverse Search.
- Not possible to offer a URL to the source documents in the citations.
- Configured at the bot level: possible to disable file search at the node level but can't filter them.

Custom data

- Allows to first query any data source – through Power Automate or HTTP requests – and then pass the data to generative answers to formulate an answer for the user query.
- Generative answers expects a table variable with 2 properties: Content (typically snippets of relevant content) and ContentLocation (typically a URL).

New design patterns with generative AI

✓ New paradigm in topic responses



Why do we share this example?

- This shows how generated responses are plugged into a dialog manager (Copilot Studio) to ensure that you remain in control of the user experience while allowing for authored experiences that are key functions of your copilot, scripted responses that are key for managing the answers you want scripted (such as marketing-controlled responses) work in conjunction with generated responses allowing the enterprise to remain in full control of the experience.
- This allows for an end-to-end enterprise conversational platform to build effective and managed experiences that delight customers, while reducing the overhead and cost of maintaining fully scripted experiences of the past.

Topic examples:

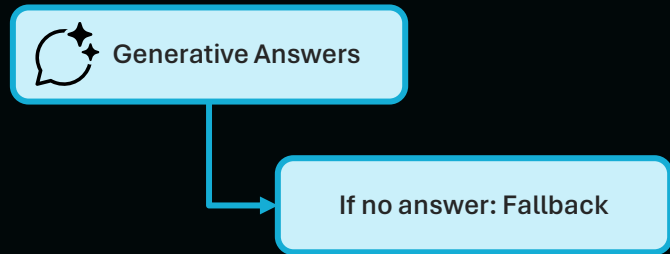
- **Authored Topics**
“How many miles do I have in my account?”
- **FAQ / Scripted Responses**
“Why are all flights currently grounded?”
- **Generated Responses**
“What type of snacks are available on your flights?”
- **Escalation**
“Can I talk to an agent?”

Infusing generative AI into topics example

✓ Finding the right place for generative AI in your new or existing copilots

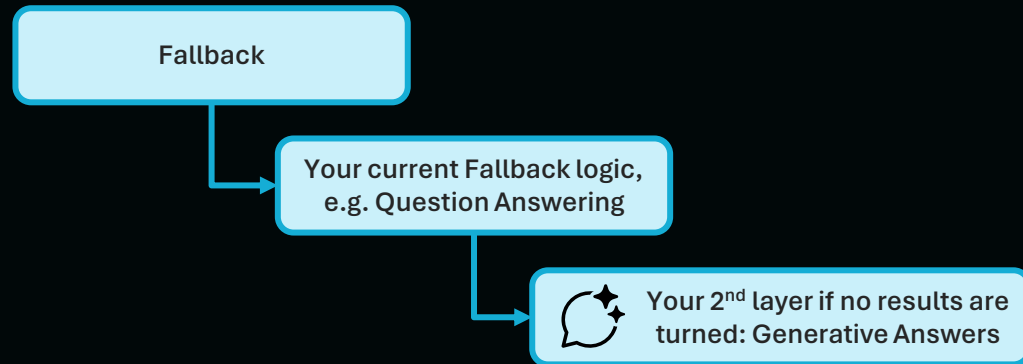
Pattern #1:

Placing Generative Answers as the first level of unknown intent logic before Fallback.



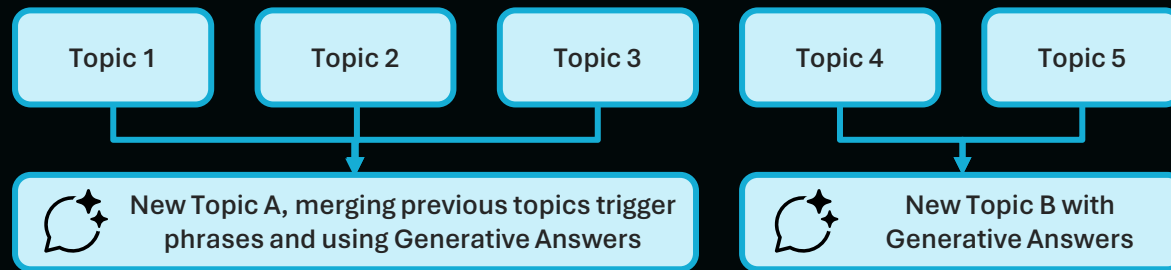
Pattern #2:

Enriching your existing Fallback



Pattern #3:

For simple, single-turn question & answer topics, group them into a single or multiple new topics with their past trigger phrases



Why do we share this example?

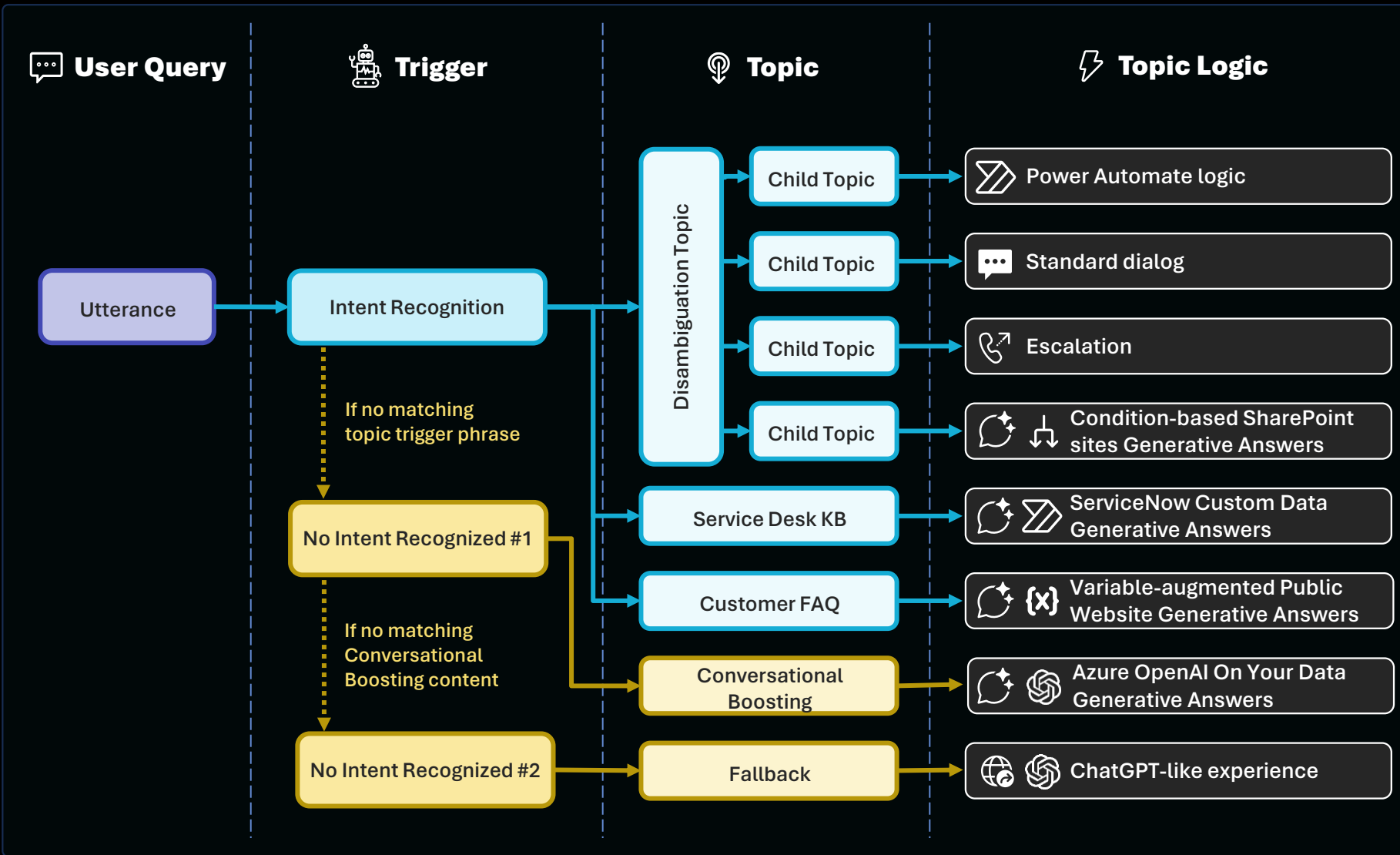
- Generative answers unlock new use cases for copilots where dialog paths no longer need to be fully anticipated.
- By plugging generative answers into your internal and external data sources, the copilot can search and summarize answers for user queries.
- It's not an 'either' choice: you can configure generative answers wherever you choose in your topics, and the data sources can both be dynamic, hard coded, or enriched with context variables.
- It's still a good idea to have generative answers also configured in the Conversational boosting topic that triggers before Fallback, to try to catch and answer user queries with broader data sources before they get to Fallback.
- While planning for generative answers, it's important to plan for follow-up questions (should they remain in context of the previous answer or trigger a new topic?) and how you will measure, track, and validate that the generated answers are accurate and answer the user questions.

Useful resources:

- Data sources: [public websites](#), [Bing Custom Search](#), [SharePoint/OneDrive](#), [Custom Data](#), [uploaded documents](#), [Azure OpenAI on your data](#).

Infusing generative AI into topics example

✓ Mixing scripted dialog paths and generative answers from different data sources



Why do we share this example?

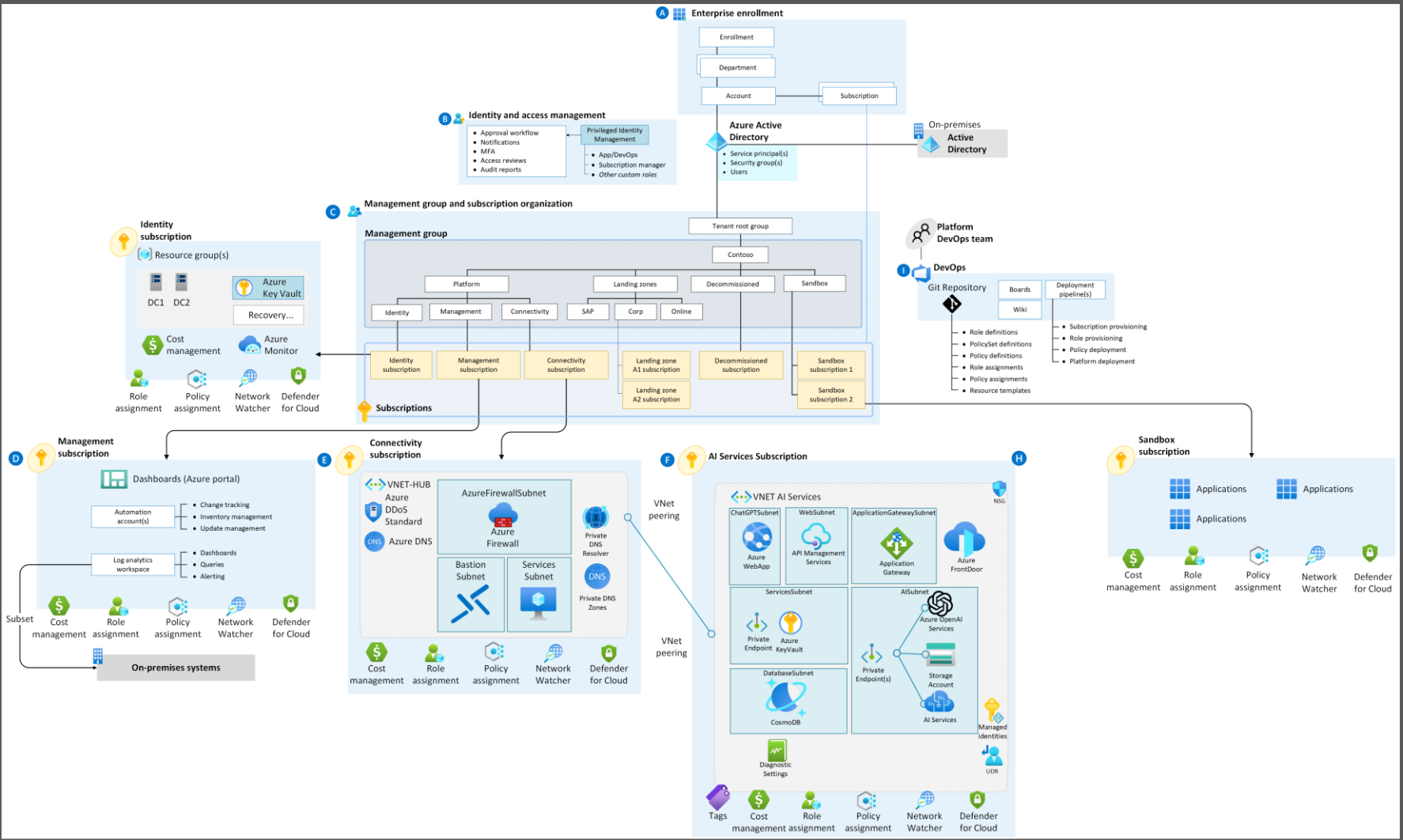
- Generative answers unlock new use cases for copilots where dialog paths no longer need to be fully anticipated.
- By plugging generative answers into your internal and external data sources, the copilot can search and summarize answers for user queries.
- It's not an 'either' choice: you can configure generative answers wherever you choose in your topics, and the data sources can both be dynamic, hard coded, or enriched with context variables.
- It's still a good idea to have generative answers also configured in the Conversational boosting topic that triggers before Fallback, to try to catch and answer user queries with broader data sources before they get to Fallback.
- While planning for generative answers, it's important to plan for follow-up questions (should they remain in context of the previous answer or trigger a new topic?) and how you will measure, track, and validate that the generated answers are accurate and answer the user questions.

Useful resources:

- Data sources: public websites, Bing Custom Search, SharePoint/OneDrive, Custom Data, uploaded documents, Azure OpenAI on your data.

Azure OpenAI reference architecture

✓ Azure OpenAI Landing Zone reference architecture



Why do we share this example?

- Using your own Azure OpenAI service? Consider reviewing the Azure OpenAI Landing Zone.
- Azure Landing Zones provide a solid foundation for your cloud environment. When deploying complex AI services such as Azure OpenAI, using a Landing Zone approach helps you manage your resources in a structured, consistent manner, ensuring governance, compliance, and security are properly maintained.
- The Azure OpenAI Landing Zone is a reference architecture that integrates a variety of services to create a seamless infrastructure for running OpenAI workloads.

Useful resources:

- [Azure OpenAI Landing Zone reference architecture](#)

Generative answers

- ✓ Dynamically generate answers real-time based on content you choose

What data sources are you using for generative answers?
Do you use custom instructions?

Example answer (you can delete this)

Yes, over public websites, internal SharePoint sites, and ServiceNow knowledge bases.

Instructions to return friendly and brief answers (40 words max).

Where, in your topic structure, do you leverage generative answers?

Example answer (you can delete this)

Mostly in custom topics that will contain trigger phrases and entity questions that will direct the user towards the right data source. Conversational boosting will use public FAQ websites.

Do you have more advanced scenarios that require your own large language models (LLMs)?

Example answer (you can delete this)

Yes, as a last resort and to handle chitchat and unrelated queries, the copilot will be able to answer in a ChatGPT manner (e.g. help write me a SQL query, etc.)

Why do we ask these questions?

- By default, generative answers is available in the 'Conversational boosting' system topic that triggers just before 'Fallback', when an intent isn't recognized.
- Generative answers nodes can however be used throughout your custom and system topics, and the data sources be different and even set dynamically.

Useful resources:

- [Generative answers](#)
- [Generative answers node](#)
- [Generative answers with Bing Custom Search \(video\)](#)
- [Generative answers with custom data – ServiceNow \(video\)](#)
- [Generative answers with uploaded documents \(video\)](#)
- [Customize responses with generative answers \(video\)](#)

Generative actions

✔ Intelligently chain together plugins

Do you plan to use plugin actions and if yes, how?

Example answer (you can delete this)

- Yes, to perform actions in various systems and letting AI ask for any missing required property for the configured APIs.

Do you plan to replace the standard NLU with dynamic chaining?

Example answer (you can delete this)

- Not initially, but dynamic chaining will be evaluated for its advanced multi-intent detection (for topics and plugin actions) and multi-entity extraction (even for entities of the same type) compared to traditional natural language understanding capabilities.

Why do we ask these questions?

- Generative actions enable copilots, with the Azure OpenAI-powered conversation engine, to select and chain together plugins, to complete tasks for your users without any manual authoring required.
- Dynamic chaining allows to replace Copilot Studio NLU (including Azure CLU) with one based on a GPT large language model, allowing to do multi-intent recognition and advanced entity extraction.

Useful resources:

- [Copilot Studio copilot building experience with generative AI](#)
- [How to use plugin actions \(video\)](#)

Generative builder

✔ Intelligent topic authoring support

Do you plan to use the generative builder to design new topics or update existing ones?

Please describe how

Example answer (you can delete this)

Yes, this helps business users easily get started without an intimidate blank topic canvas.

Copilot is also used to update existing messages and questions to suggest variations, making the copilot more friendly and natural.

Copilot also helps create adaptive cards that summarize the information gathered in multiple questions and variables.

Why do we ask these questions?

- In addition to the traditional method for creating and editing topics, Copilot Studio lets you create and iterate on topics by describing what you want and having AI build it for you, significantly decreasing the time it takes to create and update copilot topics.
- While AI helps get faster to results, it's important to check that the generated content works as intended and meets expectations.

Useful resources:

- [Create \(and edit\) with copilot](#)
- [Topic authoring with copilot \(video\)](#)

Other AI requirements

✔ Detail other expected AI capabilities for your project

Do you have other AI requirements for your copilot project?

Example answer (you can delete this)

Yes, AI capabilities are expected to cover these scenarios:

- Sentiment analysis to escalate the conversation to a live agent if the tone is too negative.
- Transcripts analysis to suggest new topics or new trigger phrases for existing topics.
- Topic overlap detection to warn of risks of ambiguity between existing topics.
- AI Builder integration for image recognition and text extraction from uploaded pictures.

Why do we ask these questions?

- Understand if other existing or custom AI capabilities are to be leveraged in the project.


Useful resources:

- [AI features for Teams and Classic bots](#)

Integrations & Channels

-
- Integrations
 - Power Automate and HTTP requests
 - Handling cloud flow timeouts examples
 - Deployed channels
 - Chat clients
 - Escalation management

Integrations

 Overview of all integrations with other systems and expected volumes

Connected system	Called by	Called with	Expected daily volume	Expected peak	Details / purpose
Example answer (you can delete this)					
ServiceNow	Service Desk KB topic	Cloud flow	5,000	1,000	The user query is used as an input to run a query on ServiceNow knowledge base using the standard connector. A JSON of results is returned to Copilot Studio for generative actions custom data.
Contoso website	Conversational Boosting topic	Generative answers	1,000	100	https://contoso.com/en-us/FAQ with /en-us/ being dynamically set based on the user locale variable.
Internal Directory API	Conversation Start topic	Cloud flow	15,000	5,000	Use of the on-premises data gateway to connect to an internal API.
Weather API	Weather topic	HTTP	500	100	GET request to a REST-based API
Teams	Meeting Booking topic	Plugin	1,500	300	Create a Teams meeting based on conversation inputs.
Bot Framework Skill	Travel topic	Skills	2,000	400	Leveraging existing Bot Framework Skill to book travels.
SAP	Every 24h	Cloud flow	20,000	-	Daily batch synchronization of the product catalogue from SAP to Dataverse.

Why do we ask these questions?

- Understanding the various systems that the copilot, or its underlying environment, will connect to, the purpose of these integrations, and the expected volumes.

Useful resources:

- [Use Power Automate cloud flows](#)
- [Perform HTTP requests](#)
- [Use plugin actions](#)
- [Use Bot Framework skills](#)

Power Automate and HTTP requests

✓ Leveraging the 1,000+ connectors to extend your copilot

How are you optimizing your cloud flows and HTTP requests to make sure they run as fast as possible?

Example answer (you can delete this)

When connecting to the target system, the query filters the results and only returns the columns that will be used.

Do you call cloud flows that don't have any connector as part of their actions? If yes, why?

Example answer (you can delete this)

No.

How do you handle cloud flow timeouts and limits (e.g., throttling)?

Example answer (you can delete this)

- Parallel branch in cloud flows to return a result before 100s.
- Each connector limit is assessed.
- Power Platform requests are accounted for.

Why do we ask these questions?

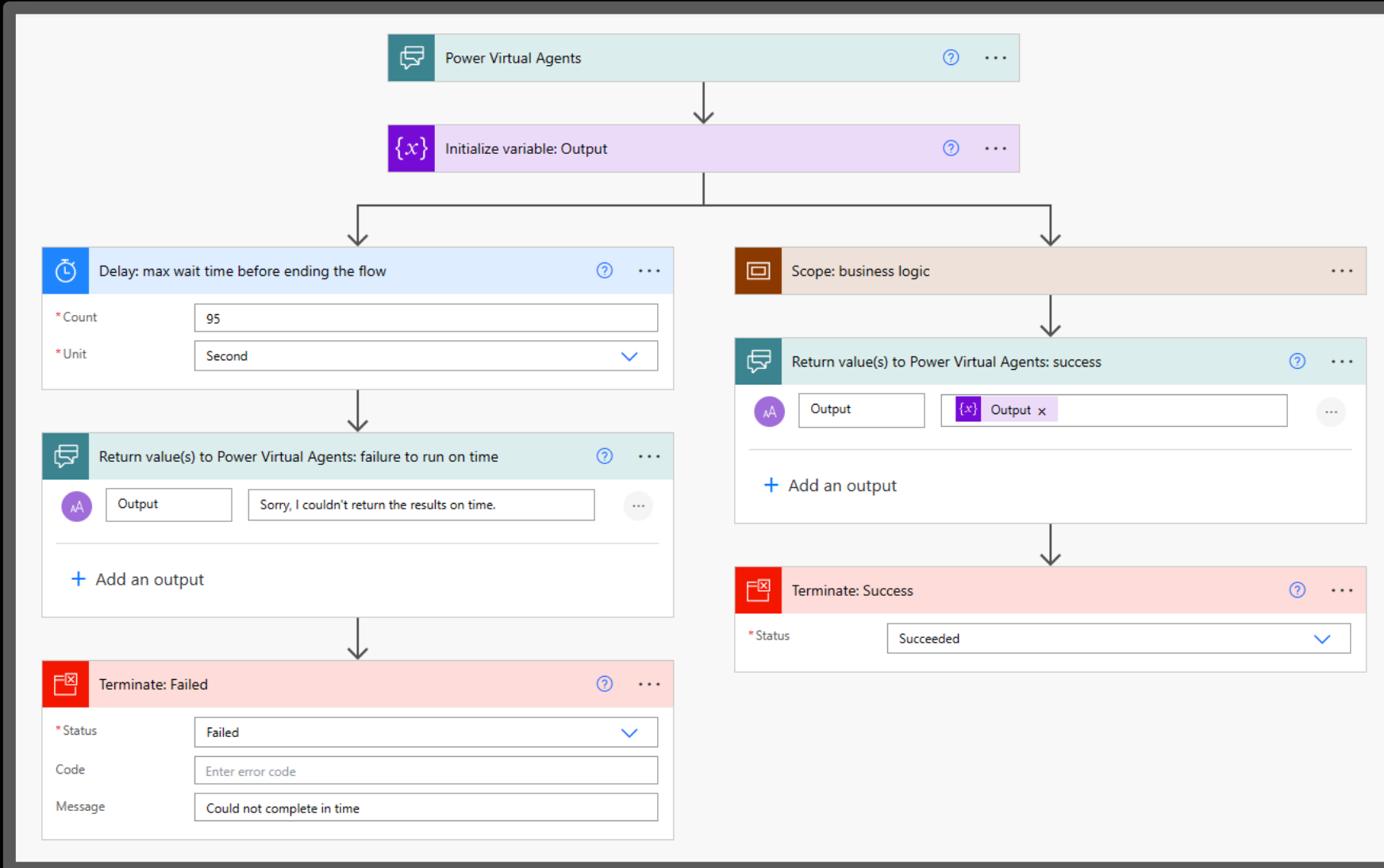
- Power Automate offers great integration capabilities, with more 1,000 native connectors or ways to create your own custom connectors to your APIs.
- For a good end-user experience, cloud flows triggered from Copilot Studio must execute quickly so that the user doesn't have to wait too long for the copilot to answer.
- Cloud flows triggered from Copilot Studio have a maximum of 100 seconds to return to Copilot Studio before they time out.
- You can make HTTP requests directly from Copilot Studio, to avoid invoking cloud flows and trim a few milliseconds.

Useful resources:

- [Call a flow as an action](#)
- [Perform HTTP requests](#)
- [Power Automate Standards by MVP Matthew Devaney](#)

Handling cloud flow timeouts example #1

✓ Return results to the copilot before the 100th second



Why do we share this example?

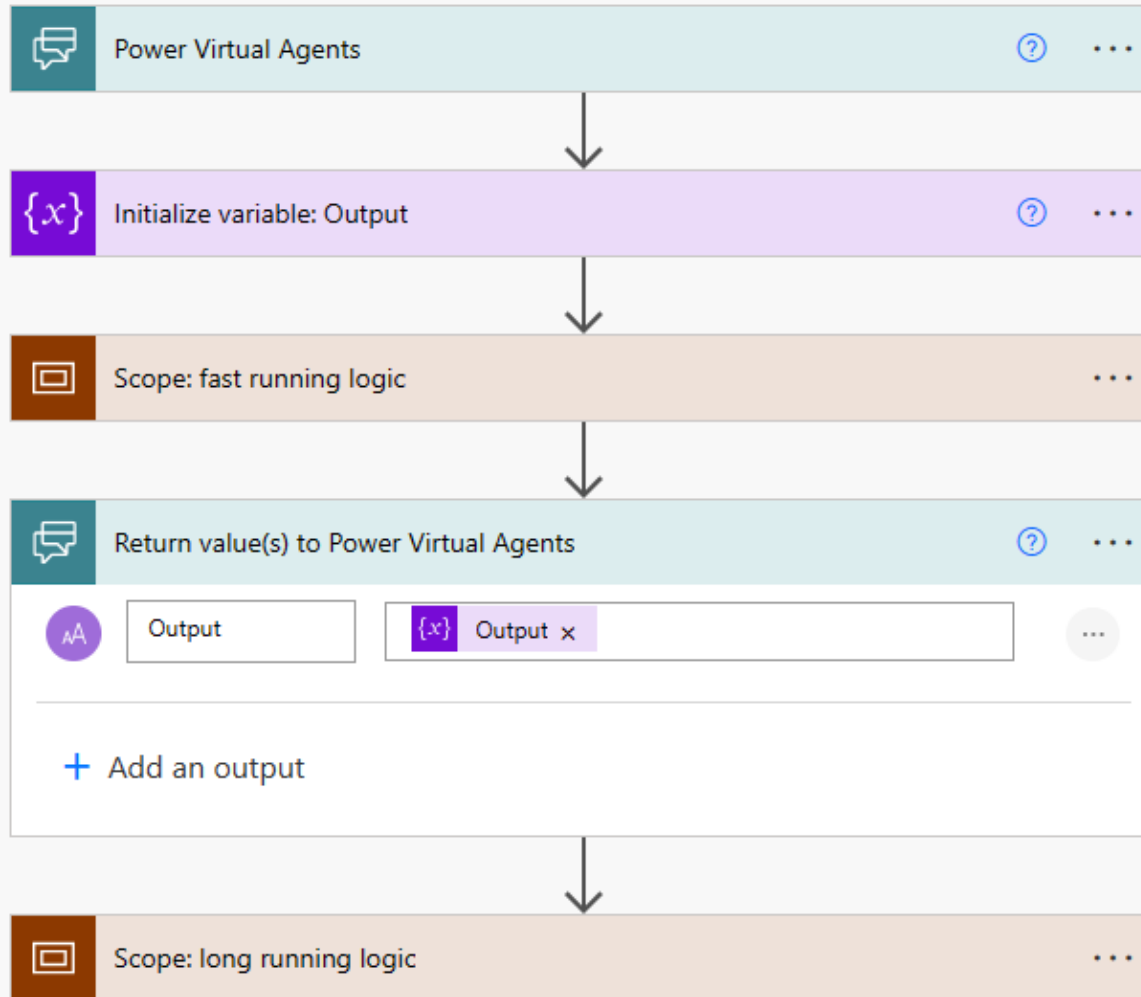
- By adding a parallel branch that returns a result to Copilot Studio within 100s, you can more gracefully handle the timeout on the copilot side.

Useful resources:

- [FlowActionTimedOut error code](#)
- [MVP Matt Jimison's example of handling cloud flow timeouts in Copilot Studio](#)

Handling cloud flow timeouts example #2

✔ Return results to the copilot and let long-running logic continue



Why do we share this example?

- Cloud flows can continue to run after outputs are returned to Copilot Studio.
- Place long running logic after the “Return value(s) to Copilot Studio” step.
- When deploying to Microsoft Teams, the ‘send proactive message’ message feature could be used to have the bot send the message to the user, once it is ready on the Power Automate side.

Useful resources:

- [FlowActionTimedOut](#) error code
- [Notify bot users in Teams with proactive messages](#)

Deployed channels

✓ Overview of all integrations with other systems and expected volumes

Channel	Expected Volume	Details
Example answer (you can delete this)		
Teams	5,000 / month	Deployed as a Teams app to all employees.
Direct Line – Intranet website	10,000 / month	Deployed using the standard custom website integration.
Direct Line – WhatsApp	2,000 / month	Azure relay copilot + channel adapter connecting to Infobip
Direct Line – Voice	5,000 / month	Voice channel through 3 rd -party AudioCodes

Why do we ask these questions?

- Copilot Studio deploys to multiple channels natively (Teams, websites, Facebook, Power Pages, Power Apps canvas apps, etc.).
- For advanced scenarios such as custom apps or web clients, customers can integrate with the copilot's Direct Line API.
- The Direct Line API supports WebSocket, and efficient streaming method to push messages to clients, whereas the HTTP GET interface enables clients to explicitly request messages.
- For additional channels, customers would typically setup a relay copilot to use Azure Copilot Services channels and potentially use community channel adapters.

Useful resources:

- [Publish a copilot to various channels](#)
- [Key concepts in the Bot Framework Direct Line API 3.0](#)
- [Add a Copilot Studio bot to Azure Copilot Service channels](#)
- [Available Azure Copilot Service channels](#)
- [Channels and Community / Botkit Adapters](#)

Deployed clients

- ✔ List the different clients the for the copilot

Client	Details	Client Customization requirements	Message formatting requirements
Example answer (you can delete this)			
Default Web Chat canvas	V1 MVP	Simple CSS and JavaScript styling options	Simple markdown 1.5 adaptive cards
Bot Framework Web Chat	V2 for advanced styling and controls	Logo & brand identity	Simple markdown 1.6 adaptive cards
Genesys Cloud chat client	V3 to add handoff	Logo & brand identity	Simple markdown 1.6 adaptive cards
Teams	For internal teams	N/A	Simple markdown 1.6 adaptive cards
SMS		N/A	Text only

Why do we ask these questions?

- There are many available options to deploy a copilot interface to users.
- Depending on the channel and client, not all Markdown and adaptive card formats may be supported.

Useful resources:

- [Configure a copilot for a live website](#)
- [Customize the web chat canvas \(simple\)](#)
- [Customize and host your own chat canvas \(advanced\)](#)
- [Bot Framework Web Chat canvas](#)
- [Format your bot messages in Teams](#)
- [Markdown Guide](#)
- [Adaptive Cards](#)

Hand off to a live agent

✔ Copilot Studio lets you hand over copilot conversations to a human

Are you handing off conversations to Dynamics 365 Omnichannel for Customer Service?
(dedicated section)

Example answer (you can delete this)

No

Are you handing off conversations to another engagement hub solution? If yes, which one and how?

Example answer (you can delete this)

Yes, Genesys Cloud through a custom adapter.

What chat client are you placing in front of your end-users?

Example answer (you can delete this)

Genesys chat widget.

Why do we ask these questions?

- When requiring hand off to an agent, the chat client must be the one of connected engagement hub, that is placed at the front, and not Copilot Studio. The engagement hub integrates with Copilot Studio through APIs, and that's how it's able to take over a conversation from a copilot to escalate to a live agent.

A full hand off to an engagement hub follows this pattern:

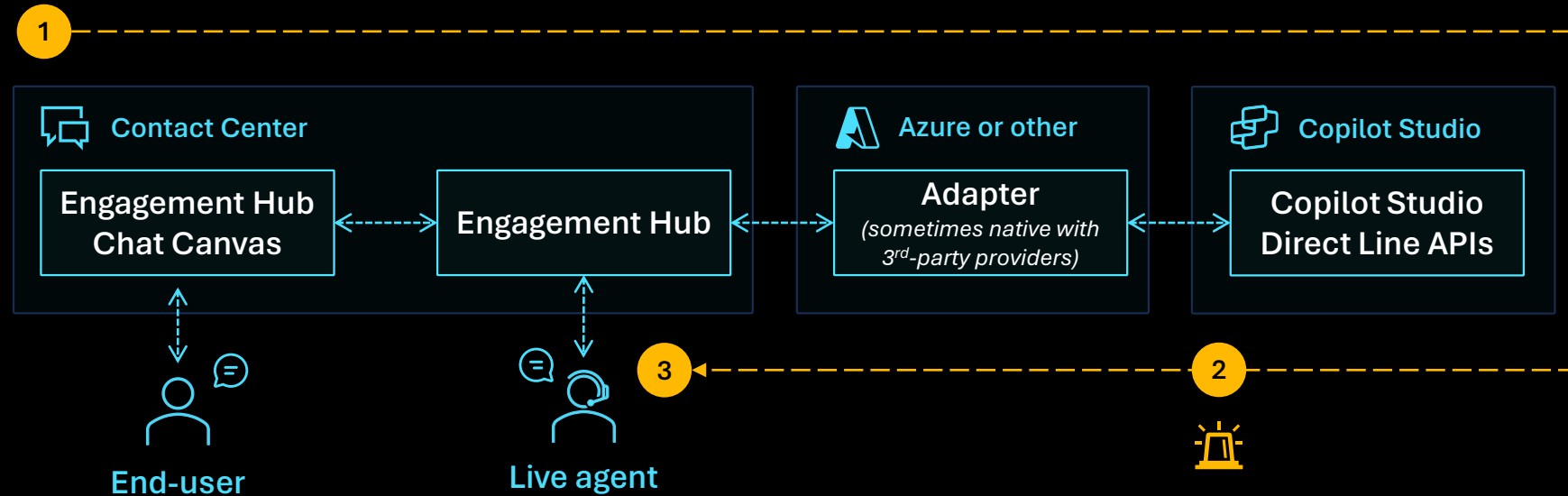
1. An end-user interacts with the engagement hub's chat canvas.
2. The engagement hub routes the incoming chat through routing capabilities.
3. A custom adapter relays the incoming chat messages to the Copilot Studio copilot.
4. Once the end user triggers hand-off, Copilot Studio starts hand-off with full chat context.
5. The custom adapter intercepts the hand-off message and context, and seamlessly routes the conversation to an agent.
6. The end user's chat is handed off to an agent who can resume the conversation.

Useful resources:

- [Hand off with Dynamics 365 Omnichannel for Customer Service](#)
- [Hand off conversations to any generic engagement hub](#)

Handing off to a live agent

✔ Pattern 1: “bot-as-an-agent” Engagement Hub at the front, Copilot Studio at the back



1 End-user chats with the Engagement Hub chat canvas and an adapter relays messages with a Copilot Studio copilot through the Direct Line APIs.

2 When an escalation event is sent by the Copilot Studio copilot APIs, the Engagement Hub takes over the conversation.

3 A live agent can resume the chat with the end-user.

Why do we share this?

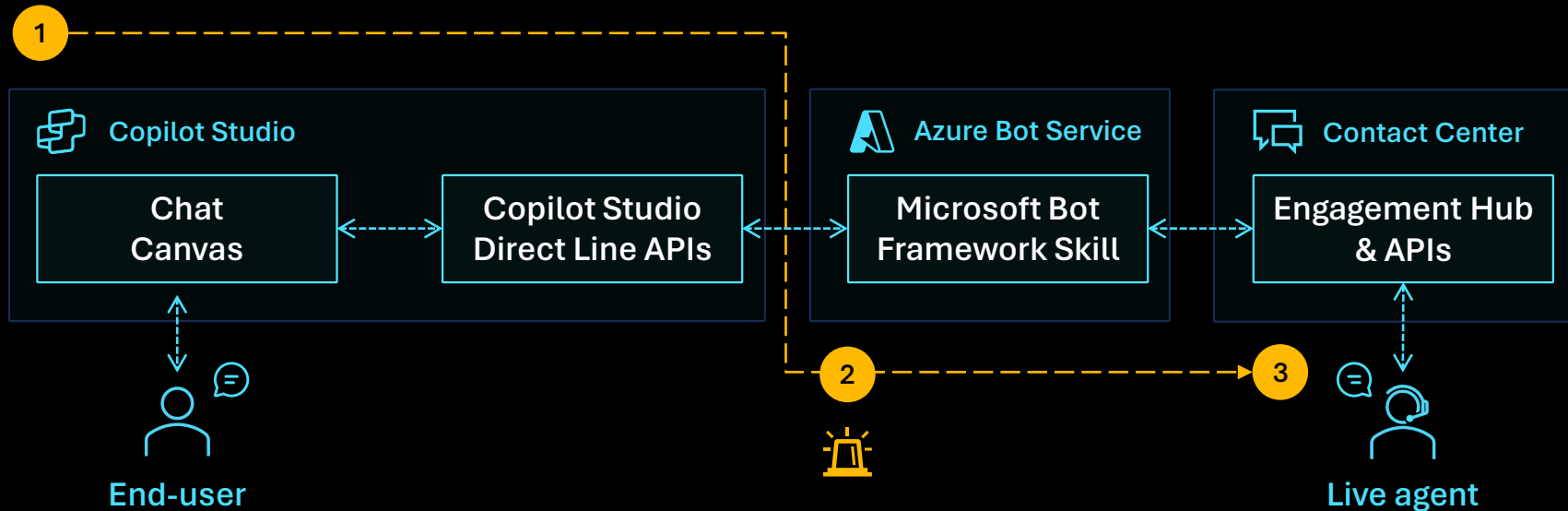
- The preferred and most standard approach to handle hand off with a live agent is to use the engagement hub chat canvas at the front.
- This is how most standard integrations with 1st-party or 3rd-party contact centers are done and made available.
- **Pros:**
 - ✓ Easier to setup without too much overhead (frontend configuration).
 - ✓ Agent messaging and capability maintain full fidelity.
 - ✓ Existing engagement hub capabilities (agent takeover, sentiment analysis, supervisory...) continue to work as-is.
- **Cons:**
 - × No ability to intercept human agent messages upon escalation (unless engagement hub supports an API).
 - × No whisper mode (unless engagement hub supports it)
 - × No control over user experience of responses (messages, cards) emitted from bot.

Useful resources:

- [Hand off to a live agent](#)
- [Hand off with Dynamics 365 Customer Service, Genesys, LivePerson, Salesforce Einstein Bot, ServiceNow.](#)
- [Hand off conversations to any generic engagement hub](#)

Handing off to a live agent

✔ Pattern 2: “bot-in-the-loop” Copilot Studio at the front, Engagement Hub at the back



- 1 End-user chats with the Copilot Studio copilot through the chat canvas (the standard one or a custom one that integrates with Copilot Studio Direct Line APIs).
- 2 When an escalation event occurs, Copilot Studio triggers a Bot Framework skill.
- 3 The skill relays messages back and forth between the Contact Center live agent and the end-user through the Engagement Hub APIs

Why do we share this?

- An alternative but more custom and complex approach is to use Copilot Studio at the front and integrate through an engagement hub APIs through a skill.
- **Pros:**
 - ✓ Copilot Studio is always in the loop, including agent messages.
 - ✓ Full control over how responses (messages, cards, etc.) emitted from the bot are shown to end-user.
 - ✓ Agent may be assisted in whisper mode (i.e., agent assist).
 - ✓ Opportunity for the bot to do skill-based routing to the right agent.
- **Cons:**
 - × Engagement hub must be sufficiently extensible to support this pattern.
 - × Many hops between systems.
 - × Requires a pro-dev and PaaS approach for the Bot Framework skill.
 - × Heavy overhead and integrations of the hub with Copilot Studio.
 - × Live agent is limited to using a compatible chat canvas.
 - × No ability for channel provider to customize their agent messages to show up in the canvas.
 - × Agent takeover / supervisor capabilities likely aren't possible.

Useful resources:



- [Hand off to a live agent](#)
- [Use Microsoft Bot Framework skills](#)

Security, monitoring & governance

-
- Security and administration controls
 - Copilot security configuration
 - Data loss prevention policies
 - Tenant and environment security
 - Authentication & channel security
 - Single sign-on
 - Monitoring
 - Governance
 - Compliance

Security and administration controls

✓ Available security, governance and AI controls at the tenant, environment, or copilot level

	 Copilot Studio features controls	 Copilot Studio maker access controls
TENANT	<ul style="list-style-type: none"> ✓ Data loss prevention (DLP) policies can be enforced for Copilot Studio and applied to all environments to: <ul style="list-style-type: none"> • Allow/block unauthenticated usage. • Allow/block individual channels. • Allow/block individual connectors. • Allow/block connecting to skills. ✓ Allow/block publishing of bots that use generative AI. ✓ Allow/block M365 Copilot Conversational Plugins. 	<ul style="list-style-type: none"> ✓ Allow/block access to the Copilot Studio authoring web app (i.e., copilotstudio.microsoft.com) can be controlled by: <ul style="list-style-type: none"> • Assignment/removal of a Copilot Studio User License to an individual user or to a security group. • Allow/block self-service trials for the tenant. ✓ Allow/block access to the Copilot Studio Teams app (i.e., aka.ms/PVAforTeams)
ENVIRONMENT	<ul style="list-style-type: none"> ✓ Data loss prevention(DLP) policies can be scoped to exclude specific environments to allow required capabilities (e.g., allow unauthenticated usage). ✓ For environments outside of the U.S. region: <ul style="list-style-type: none"> • Allow/block generative AI features. • Allow/block the public data source (i.e., Bing). 	<ul style="list-style-type: none"> ✓ Allow/block environment access with security groups. ✓ Allow/block permissions to create, read, update or delete bots through security roles. Dataverse tables: <ul style="list-style-type: none"> • Chatbot (<i>bot</i>) • Chatbot Subcomponent (<i>botcomponent</i>) • Conversation Transcript (<i>conversationtranscript</i>)
COPILOT	<ul style="list-style-type: none"> ✓ Individual bots can be exempted from DLP policies. ✓ Enable/disable dynamic chaining at bot level ✓ Enable/disable generative answers at bot level. ✓ Enable/disable intelligent topic authoring at user level. ✓ Set bot authentication (None – if allowed, Only for Teams/Power Apps, and Manual – i.e., app registration) ✓ Enforce Web Channel security at bot level. 	<ul style="list-style-type: none"> ✓ Share/unshare bots with other users for authoring. ✓ Users with a System Administrator or System Customizer roles can read and update all bots and transcripts.

Why do we share this?

- Data loss prevention (DLP) policies can be configured at tenant-level to apply either to all environments, only specific environments, or all environments except specific ones. DLP policies can enforce authentication, disable specific channels (e.g., Direct Line, Teams, Omnichannel, Facebook), block specific connectors, or block the use of skills. Individual copilots can be exempted from DLP policies.
- The publishing of copilots using generative AI features such as generative answers can be completely blocked at the tenant level.
- For environments outside of the U.S. region, generative AI features can be blocked at the environment-level. The public data source (Bing) can also be individually disabled.

Useful resources:

- [Security and governance in Copilot Studio](#)
- [Enforce DLP policies for Copilot Studio](#)
- [Exempt a copilot from DLP policies](#)
- [Assign Copilot Studio user licenses](#)
- [Control user access to environments](#)
- [Manage app permissions in Microsoft Teams to block the PVA Teams app](#)
- [Disable self-service trial sign-ups: set AllowAdHocSubscriptions to \\$false](#)
- [Share a copilot for collaborative authoring](#)
- [Configure web channel security](#)

Copilot security configuration

✔ Detail your security requirements

Do you have a list of security requirements for your copilot?

How are you going to implement these?

Example answer (you can delete this)

Enterprise copilots:

- Users need to be authenticated using Azure Active Directory.
- Users need to be automatically signed in.
- IPs need to be allow-listed in the corporate firewall.
- Need to integrate with internal, on-premises, APIs.
- Integrations must use service principals.
- Secrets must be stored in key vaults.
- Only specific apps should be able to call the copilot as a skill.

Citizen-developer copilots:

- The unauthenticated configuration need to be blocked on all copilots and all channels.
- Copilot usage (inventory and conversations) need to be monitored at the tenant level).

Why do we ask these questions?

- Understanding your key requirements around security and making sure they are mapped with existing features.
- Identify alternatives for scenarios not natively covered.

Useful resources:

- [Copilot Studio security](#)

Data loss prevention policies

- ✓ Enforcing security settings as well as exceptions at a global level

How do you ensure that internal confidential data doesn't get available externally?

Example answer (you can delete this)

Strong governance of connectors and authentication settings, especially for generative answers data sources.

Do you plan to use DLP policies to enforce Copilot Studio allowed channels and authentication settings?

Example answer (you can delete this)

Yes, by default no unauthenticated copilots are allowed and only the Microsoft Teams channel can be used, to make sure they're internal.

Do you plan to use DLP policies to enforce allowed or blocked Power Automate connectors or HTTP requests?

Example answer (you can delete this)

Yes, with Copilot Studio, only related business applications such as Dynamics 365, ServiceNow and SharePoint are allowed. Endpoint filtering will be used for allowed HTTP APIs.

Why do we ask these questions?

- Administrators can govern copilots in their organization by using data loss prevention (DLP) policies with specific Copilot Studio connectors.
- Channels (Teams, Direct Line, Facebook, Omnichannel), unauthenticated use, and Bot Framework skills can be allowed or blocked.

Useful resources:

- [Security and governance in Copilot Studio](#)
- [Apply data loss prevention policies to copilots](#)
- [Connector endpoint filtering](#)

Tenant and environment security

✔ Securing access to data

Do you enforce strong requirements for authentication?

Multi-factor authentication and conditional access.

How do you control privileged users access to sensitive data?

Just-in-time access to production.

How do you manage secrets?

Azure Key Vault and Environment Variables

Do you use service principals or managed identities?

Service principals for integrations and deployments.

Do you have network security requirements?

Some internal APIs are only accessible in a VNET.

Do you audit and review access periodically?

Yes, every month. Roles only granted through groups.

Why do we ask these questions?

- Copilot conversation transcripts are automatically stored in Dataverse.
- These can contain sensitive customer information that need to be protected.

Useful resources:

- [Power Platform security](#)
- [Microsoft Entra multifactor authentication](#)
- [Microsoft Entra Conditional Access](#)
- [Use Azure Key Vault secrets](#)
- [Just-In-Time Access Management](#)
- [Manage application users](#)
- [Manage group teams](#)
- [Customer Lockbox](#)

Authentication & channel security

✓ Authentication requirements

Is the copilot working unauthenticated, authenticated, or both?

Example answer (you can delete this)

External copilot: hybrid authentication. The public website can start chat sessions and handle FAQ questions unauthenticated, but specific operations require the user to authenticate

If yes, what is the authentication service?

Example answer (you can delete this)

Internal copilot: “Manual” (AAD) or “Only for Teams and Power Apps”. External copilot: “Manual” (Generic OAuth2)

Do you need to enforce web and Direct Line channel security?

Example answer (you can delete this)

Internal copilot: yes.
External copilot: no.

Why do we ask these questions?

- Authentication allows users to sign in, giving your copilot access to a restricted resource or information. Users can sign in with Microsoft Entra ID, or with any OAuth2 identity provider such as Google or Facebook.
- With Direct Line-based security, you can enable access only to locations that you control by enabling secured access with Direct Line secrets or tokens.

Useful resources:

- [Configure user authentication](#)
- [Configure user authentication with Microsoft Entra ID](#)
- [Configure web channel security](#)

Single sign-on (SSO)

✔ Automatic signing in authenticated users

Do you have SSO requirements?

How are you going to implement these?

Example answer (you can delete this)

Internal copilot:

- Teams users and intranet website users must be automatically signed in (AAD v2).
- SSO must also work on the internal SharePoint intranet.

External copilot:

- SSO with Generic OAuth 2.

Why do we ask these questions?

- Copilot Studio supports single sign-on (SSO), which means copilots can sign the user in.
- SSO needs to be implemented on your web pages, mobile applications.
- For Microsoft Teams, SSO is seamless if you select the “Only in Teams” authentication. It can also be configured with Manual AAD v2, but in this case the Teams app must be deployed as a zip file, not with the 1-click Teams deployment from Copilot Studio.

Useful resources:

- [Configure SSO with Entra ID for web/apps](#)
- [Configure SSO with Entra ID for Teams](#)

Monitoring

✓ Capture telemetry, measure performance, and setup alerts with Application Insights

How do you plan to technically monitor your copilots?

Example answer (you can delete this)

Azure Application Insights to log activities, errors, and custom events. Also used to monitor Generative Answers and details on failed or moderated answers.

How do you plan to technically monitor your integrations and cloud flows?

Example answer (you can delete this)

Azure Application Insights for cloud flows.

Do you have other technical monitoring requirements and plans?

Example answer (you can delete this)

Yes, the copilot client load time on the web page will be monitored with our web analytics tools.

Why do we ask these questions?

- In addition to the native analytics features within Copilot Studio, you can send telemetry data to Application Insights.
- You can send custom events to Azure Application Insights.

Useful resources:

- [Capture Copilot Studio telemetry with Azure Application Insights](#)
- [Set up Application Insights with Power Automate](#)

Azure Application Insights example

Monitoring generative answers results and moderation

The screenshot displays the Azure Application Insights interface. At the top, there's a search bar with 'Generative Ans...' and a '+'. Below it, the 'Run' button is visible along with 'Time range: Custom'. The main area shows a Kusto query in the editor, which is also displayed in a larger text box on the right. The query filters for 'GenerativeAnswers' and projects various fields including cloud_RoleInstance, name, timestamp, conversationId, topic, message, result, feedback, and Summary, ordered by timestamp descending.

```
1 customEvents
2 | where name == "GenerativeAnswers"
3 // | where cloud_RoleInstance == "MS Learn Chatbot" // Bot Name
4 | extend cd = todynamic(customDimensions)
5 | extend conversationId = tostring(cd.conversationId)
6 | extend topic = tostring(cd.TopicName)
7 | extend message = tostring(cd.Message)
8 | extend result = tostring(cd.Result)
9 | extend SerializedData = tostring(cd.SerializedData)
10 | extend Summary = tostring(cd.Summary)
11 | extend feedback = tostring(todynamic(replace_string(SerializedData,"$", "")).value)
12 | project cloud_RoleInstance, name, timestamp, conversationId, topic, message, result, feedback, Summary
13 | order by timestamp desc
```

The results table below the query shows the following data:

cloud_RoleInstance	name	timestamp [UTC]	conversationId	topic	message	result	Summary
MS Learn Chatbot	GenerativeAnswers	11/22/2023, 4:52:30.856 PM	302287e9-9ca7-44e2-860c-985...	Conversational boosting	What is Microsoft Copilot Studio? Does i...	Answered	Microsoft Copilot Studio combines natural language u...
Microsoft Copilot Studio Demo	GenerativeAnswers	11/21/2023, 9:15:28.783 AM	ANGtaAVlu21EFxrJegV3r-us	SharePoint or OneDrive for Business	What is the patient discharge process?	Answered	The patient discharge process involves cohesive joint w...
MS Learn Chatbot	GenerativeAnswers	11/20/2023, 3:31:36.750 PM	9094d8ba-d29d-46cf-bdea-645...	Conversational boosting	What is Copilot Studio?	Answered	Microsoft Copilot Studio is a platform that allows you t...
MS Learn Chatbot	GenerativeAnswers	11/20/2023, 3:14:22.060 PM	9094d8ba-d29d-46cf-bdea-645...	Conversational boosting	What is Power Virtual Agents?	Answered	Power Virtual Agents is a platform that allows you to cr...
MS Learn Chatbot	GenerativeAnswers	11/20/2023, 3:13:04.886 PM	98787c69-ccca-45ab-878d-73f...	Conversational boosting	microsoft store	Filtered by High Content Moderation	
Microsoft Copilot Studio Demo	GenerativeAnswers	11/18/2023, 6:53:26.307 PM	FoeHwLeXvYxYFqS2ye8vZ-us	SharePoint or OneDrive for Business	What is the patient discharge process?	Answered	The patient discharge process involves cohesive joint w...
Microsoft Copilot Studio Demo	GenerativeAnswers	11/18/2023, 6:51:51.553 PM	e933479a-bb63-4221-a9e7-c4e...	SharePoint or OneDrive for Business	What is the patient discharge process?	Answered	The patient discharge process involves cohesive joint w...
MS Learn Chatbot	GenerativeAnswers	11/18/2023, 12:18:38.603 AM	E4gKZ9mQMWCJ2l0GAtLv3f-us	Conversational boosting	what is copilot?	No Search Results	
MS Learn Chatbot	GenerativeAnswers	11/16/2023, 11:28:00.962 PM	9juuiz7RIISFdYf6asqSKS-us	Conversational boosting	copilot studio	Filtered by Open AI	
MS Learn Chatbot	GenerativeAnswers	11/16/2023, 11:27:43.621 PM	9juuiz7RIISFdYf6asqSKS-us	Conversational boosting	what is copilot?	No Search Results	
Microsoft Copilot Studio Demo	GenerativeAnswers	11/16/2023, 9:36:46.399 PM	Cqz1Ukrp9Mml2WH03V4vtv-us	SharePoint or OneDrive for Business	What is the patient discharge process?	No Search Results	
MS Learn Chatbot	GenerativeAnswers	11/15/2023, 12:43:05.627 AM	ee81d804-feb6-4bac-acdf-b0d...	Conversational boosting	What is Power Apps?	Answered	Power Apps is a suite of apps, services, and connectors ...
MS Learn Chatbot	GenerativeAnswers	11/15/2023, 12:42:43.349 AM	ee81d804-feb6-4bac-acdf-b0d...	Conversational boosting	what is power automate?	Answered	Power Automate is a service that allows you to automa...

Governance

✔ How to you govern Copilot Studio at scale?

Do you plan to use the Center of Excellence Starter Kit to monitor your copilots in your tenant?

Example answer (you can delete this)

Yes, the central CoE team already uses the Starter Kit to monitor Power Apps and Power Automate. Copilot Studio monitoring will allow track usage across copilots and environments.

Do you have specific governance requirements?

Example answer (you can delete this)

As capacity is purchased globally and pooled at the tenant level, tenant-wide analytics are required to implement a charge-back model.

Do you allow users to create copilots using Copilot Studio for Teams?

Example answer (you can delete this)

No, Copilot Studio for Teams is blocked as an app in Microsoft Teams. Trials of Copilot Studio are blocked, and users need a license to create copilots.

Do you allow users to create Copilot Studio trials?

Why do we ask these questions?

- As customers expand their use of Microsoft Power Platform to enable citizen developers across their organizations, they also seek capabilities to govern and monitor usage.
- The Microsoft Power Platform CoE Starter Kit is a collection of components and tools that are designed to help organizations develop a strategy for adopting and supporting Power Platform.

CoE Starter Kit features for Copilot Studio

- Keep track of the number of conversations per copilot over time, with daily aggregates.
- Allocate a number of conversations per environment
- Receive capacity email alerts when environment add-ons get near or over their allocated number of conversations.
- Get an overview dashboard of Copilot Studio usage in a tenant over time as well as an all-up conversation consumption versus allocation.

Useful resources:

- [Center of Excellence Starter Kit with Copilot Studio: Govern, Monitor and Manage](#)

Compliance

- ✓ Understand how regulatory requirements impact your project

What the regulatory requirements and norms you need to comply with?

Example answer (you can delete this)

GDPR compliance.

Do you have hard requirements as to where the data must be located?

Example answer (you can delete this)

Data location and processing must stay within the European Union boundary.

Do you have other compliance requirements?

Example answer (you can delete this)

- Transcripts must be kept for 2 years in a compliant location.
- Sensitive data must be obfuscated from transcripts

Why do we ask these questions?

- Microsoft complies with data protection and privacy laws applicable to cloud services. Our compliance with world-class industry standards is verified.
- Environments can be created in specific regions, even if different from the region the tenant resides in.
- By default, conversation transcripts are only kept for 30 days in Dataverse. This can be updated on per-environment basis.

Useful resources:

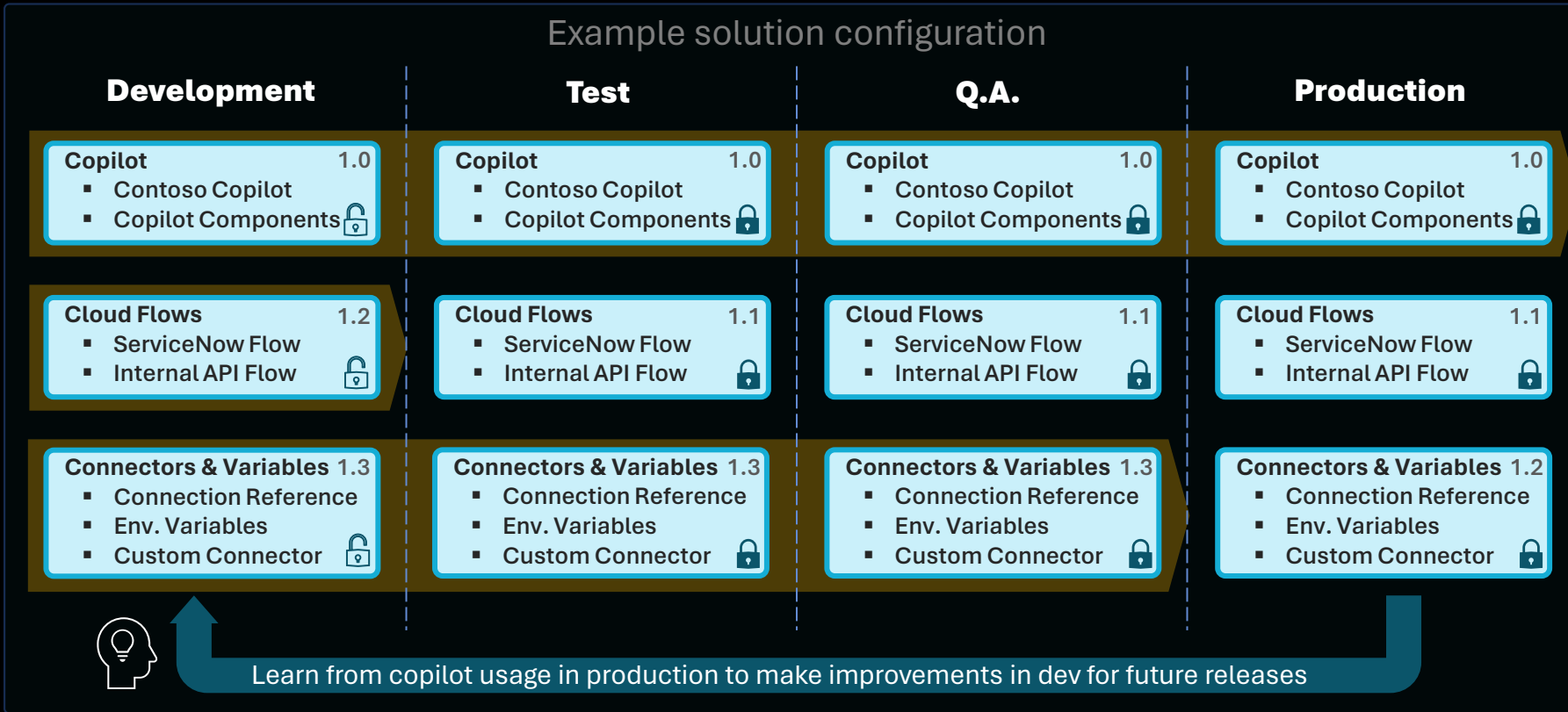
- [Copilot Studio compliance offerings](#)
- [Copilot Studio GDPR compliance](#)
- [Copilot Studio Data Locations](#)
- [Managing compliance in the cloud](#)
- [Service Trust Portal](#)
- [Change the default retention period for conversation transcripts](#)
- [Move data across geographic locations for generative AI features outside United States](#)

Application Lifecycle Management

-
- Application Lifecycle Management example
 - ALM strategy
 - Environments
 - Capacity
 - Test strategy
 - Training
 - Run & maintenance

Application Lifecycle Management

✓ Safely deploy customizations across environments and keep track of changes.



Why is ALM important?

- Copilot Studio ALM is based on the broader Power Platform ALM principles.
- They use the concept of solutions to package and deploy customizations.

What are ALM golden rules?

- Work in the context of solutions
- Create separate solutions only if you need to deploy components independently.
- Use a custom publisher and prefix.
- Use environment variables for settings and secrets that change across those.
- Export and deploy solutions as managed, unless setting up a dev environment.
- Don't do customizations outside of dev.
- Consider automating ALM for source control and automated deployments.

What needs to be set manually?

- Azure Application Insights integration.
- Deployed channels.
- Some security settings.

Useful resources:

- [Use environment variables](#)
- [Power Platform ALM](#)
- [Power Platform pipelines](#)
- [Azure DevOps Power Platform Build Tools](#)
- [GitHub Power Platform Actions](#)
- [ALM Accelerator for Power Platform](#)
- [Test Framework](#)



Automate

Evaluate options from simple to advanced continuous integration & delivery (CI/CD)



1. Manual deployment of solutions
2. Automated deployment with user-friendly pipelines in Power Platform (no source control)
3. Automated deployment + source control with Azure DevOps or GitHub Power Platform tools.
4. Use of the ALM Accelerator for #2 + #3 and advanced deployment options for both no-code and pro developers.
5. Automated testing of the copilot as part of the deployment process.



Test

ALM Strategy

✓ Detail your CI/CD processes

How to you deploy your customizations, variables and connections?

Example answer (you can delete this)

Deployment are done using Azure DevOps pipelines and managed solutions. Environment variables are used for settings and secrets. Connections are configured by a service account.

How do you organize your solutions?

Example answer (you can delete this)

A single solution with the copilot, cloud flows, connection references, environment variables.
A single publisher, "Contoso" (cto_) is used.

Do you source control your customizations?

Example answer (you can delete this)

Azure DevOps. Pull requests require peer reviews before customizations are merged into the main branch and deployed to downstream environments.

Why do we ask these questions?

- It's key to have a healthy ALM process to avoid production issues and catch and fix any regression early.

Useful resources:

- [Power Platform ALM](#)
- [Use segmented solutions](#)

Environments

- ✔ List the different environments and copilots for your project

Tenant ID

8a235459-3d2c-415d-8c1e-e2fe133509ad

Name	Type	Region	Required apps	Environment Id	Bot(s) and Bot Id(s)
Example answer (you can delete this)					
Contoso DEV	Sandbox	Europe	D365 + OC Voice	7f8448e4-b056-e6f9-a235-ddec0aa85c0d	Contoso Copilot cb093eb6-ec05-ee11-8f6e-000d3a5c61dd
Contoso TEST	Sandbox	Europe	D365 + OC Voice	47378d94-f56c-eed0-b8cd-89d17d7e31b3	Contoso Copilot d182abc5-7335-ee11-bdf4-000d3a37044a
Contoso QA	Sandbox	Europe	D365 + OC Voice	20bb12c3-1a5c-e8ff-b0f7-4f32adef8616	Contoso Copilot 64c03bd0-04c4-499a-8fa6-77de5d1d27aa
Contoso	Production	Europe	D365 + OC Voice	c53bf002-7923-4d1c-b5ae-3265093d59e1	Contoso Copilot 65d3b807-8cd7-4b58-afae-768ecd914db7

Why do we ask these questions?

- Knowing your tenant, environment, and copilot ids can be useful for proactive monitoring.
- Understanding how you setup your environments is also important in relation with your ALM processes.
- The region can be meaningful for releases and availability of certain features.
- It's a best practice to limit access to an environment to the specific members of a Microsoft Entra ID security group.

Useful resources:

- [Create and use environments](#)
- [Control user access to environments](#)

Licensing & capacity

✔ Aligning product, user licenses & add-ons to technical and business requirements

Product	Quantity	Capacity	Comments
Example answer (you can delete this)			
Copilot Studio User	50	50 users	For Copilot Studio bot authors
Copilot Studio Messages	100	2,500,000 messages / month	To allocate per environment
Dataverse Database Storage	50	50 GB	+ 10 GB Database base capacity
Dataverse File Storage	5	5 GB	+ 20 GB Database base capacity
Dataverse Log Storage	5	5 GB	+ 2 GB Database base capacity
Power Platform Requests	1	50,000 requests / 24h	+ 250,000 requests / 24h base capacity
Power Automate Process	4	4 processes / month	To run flows with service principals
Azure OpenAI	1	Standard	1 gpt-35-turbo model
Azure AI Search Service	1	Basic Tier	1 Replica, 1 Partition, 1 Search unit
Azure Storage	1	StorageV2	RA-GRS
Azure Application Insights			

Why do we ask these questions?

- Understanding what you have purchased in terms of capacity for the project.
- Potential impact of design on capacity and licensing. E.g., keeping conversation transcripts will grow database storage.
- Additional services, such as Azure OpenAI and AI Search (if used as the Azure OpenAI 'on your data' data source in generative answers), Data Lake (if used to store conversation transcripts) or Application Insights (to log additional telemetry) require an Azure subscription.
- Power Automate cloud flows owned by a service principals require a Power Platform Process license. If owned by a user, they're covered for premium connectors and come with a base capacity of 250,000 requests per month.

Useful resources:

- [Assign user licenses and manage access](#)
- [Quotas in Copilot Studio](#)
- [Licensing overview for Power Platform](#)
- [Copilot Studio pricing](#)
- [Power Automate pricing](#)
- [Power Platform Licensing Guide](#)
- [Azure OpenAI Service pricing](#)
- [Azure AI Search pricing](#)
- [Azure Data Lake pricing](#)
- [Azure Application Insights pricing](#)

Test strategy

✓ Test plan and non-functional requirements

What is your test strategy for your copilots and integrations?

Example answer (you can delete this)

Use of the Test Framework to bulk test user utterances and validate that the appropriate topic triggers or that the first 'did you mean' option is the correct one in 90% of the time.

Integrations will be tested as part the full end-to-end conversation testing with the Test Framework scale test capability.

Do you have non-functional requirements?

Example answer (you can delete this)

Chat widget on the website should load and start the conversation in less then 5 seconds when clicked.

Cloud flows triggered to return information to the user need to return the desired data in a maximum of 10s.

Why do we ask these questions?

- Understand your plans to validate that your copilots and integrations work as expected.

Useful resources:

- [Automate testing of your bots with the Test Framework sample solution](#)

Analytics & KPIs

-
- Engagement and outcomes
 - Analytics strategy
 - Optimization strategy

Engagement and outcomes

✓ How do you track engagement rates & session outcomes?

Example answer (you can delete this)

- Every conversation path, including Generative Answers ones, end with the End of Conversation topic.
- End of Conversation topic is customized to make the resolution validation as simple as possible for the end-user.
- Conversation outcomes are also set at specific node levels (in the YAML).

The screenshot shows a chatbot conversation with four numbered steps:

- 1**: Bot: "Hello, I'm Retail Bot. How can I help?" (Just now). User: "When are you open?" (Just now). Bot: "I'm happy to help with store hours." (Just now). Bot: "The Redmond hours are: Mon-Sat: 9am to 6pm, Sun: Closed." (Just now).
- 2**: User: "Did that answer your question?" (Just now). User: "Yes" (Just now).
- 3**: Bot: "Great! Please rate your experience." (Just now). A 5-star rating scale is shown with stars 1 through 5.
- 4**: Bot: "Thanks for your feedback." (Just now). Bot: "Can I help with anything else?" (Just now). User: "Yes" (Just now).

Why do we ask these questions?

- Tracking conversation engagement and outcomes is crucial to **measure** the **copilot performance metrics** and spot areas for improvements in the analytics dashboard.
- A **conversation** with a Copilot Studio copilot can generate one or multiple **sessions**.
- A conversation can have multiple sessions when a user has new questions after an initial conversation completed (**4**).
- Sessions are either **Unengaged** or **Engaged**. Unengaged sessions always have an outcome of **None**.
- A session is engaged (**1**) by either triggering a **custom topic** or the **Escalate** topic.
- Engaged sessions outcome can be either:
 - ✓ **Abandoned**
 - ✓ **Resolved**
 - ✓ **Escalated**
- It's important to end conversation with the **End of Conversation** topic (**2**) so that the end-user can confirm their issue was resolved or not (and potentially escalate).
- The **Confirmed Success** topic displays a **CSAT survey** (**3**) to capture a 0-5 score.
- A 'conversationOutcome' can also be set at the node level in the code editor view.

Useful resources:

- [Key concepts – Analytics](#)
- [Measuring copilot engagement](#)
- [Measuring copilot outcomes](#)

Analytics strategy

- ✓ Detail your strategy to monitor your copilot key performance indicators

Do you have an analytics strategy?

Do you plan to develop your own custom analytics?

If you are going to create your own reports, please detail how and where you plan to store the data.

Example answer (you can delete this)

- Initially, only a few set of KPIs will be defined for the copilot success, essentially the engagement rate, resolution rate and the defection rate.
- In phase 1, assess the native dashboards, even though they're limited to 30-day of data.
- In phase 2, develop a Power BI report consuming the copilot and Conversation Transcript data directly from Dataverse. This will also be the opportunity to enrich copilot data with other business metrics such as sales or website traffic.
- In phase 3, move data to cheaper long-term storage option (e.g., Azure Data Lake) and update the Power BI report to point to these storage points.

Why do we ask these questions?

- Copilot Studio provides comprehensive out-of-the-box analytics that allow customers to understand a copilot's usage and key performance indicators.
- Customers can view reports related to:
 - ✓ Performance and usage.
 - ✓ Customer satisfaction.
 - ✓ Session information.
 - ✓ Topic usage.
 - ✓ Billed sessions.
- However, there are often scenarios where you'll need to create or use custom analytics. For example, you may need to:
 - ✓ Share analytics with non-makers or users.
 - ✓ Report on conversation transcripts data for a period longer than the default last 30 days.
 - ✓ Design a report not covered by out-of-the-box analytics.

Useful resources:

- [Custom analytics strategy](#)

Optimization strategy

- ✓ What is your plan to keep improving your copilot's performance and ROI?

Detail your strategy to regularly improve your copilot

Example answer (you can delete this)

Monthly review of the copilot performance

- Deflection rate
- Resolution rate
- Engagement rate
- Topics with low resolution
- Unrecognized utterances
- Analysis per channel

This review helps prioritize the backlog of copilot updates. For example, unrecognized utterances are used to either train existing topic or create new topic where there is demand.

Why do we ask these questions?

- Return on investment (ROI) and improved customer satisfaction (CSAT) are top priorities for the organizations that implement Copilot Studio copilots.
- Optimizing the copilot deflection rate is one of the top focus areas for organizations to achieve their business goals around ROI and CSAT, and to improve the copilot's overall performance. There are major indicators in Copilot Studio that help improve copilot performance, such as resolution rate, escalation rate, and CSAT.
- While the metrics continue to evolve, there are several things you can do as a copilot builder to improve the deflection rate of your copilot. In these articles, we cover the importance of deflection in conversational AI and general techniques/considerations that are universal for optimizing deflection for copilots.

Useful resources:

- [Deflection optimization](#)

Gaps & top requests

-
- Gaps
 - Top requests

Gaps

✔ What are the main feature gaps you have identified

Gap #1

Example answer (you can delete this)

No native first-party IVR without Dynamics 365 Omnichannel Voice

Gap #2

Example answer (you can delete this)

No native integration with WhatsApp

Gap #3

Example answer (you can delete this)

Pro-code is required for a SharePoint integration with single sign-on support.

Why do we ask these questions?

- Help us prioritize our future investments by flagging gaps and submitting feature requests.
- Detailing the business impacts for your organization helps build a business case for them.

Useful resources:

- aka.ms/CopilotStudioFeatureRequest

Top requests

✓ What are the top gaps, feature requests, and priorities?

Request #1

Example answer (you can delete this)

P1: native file upload support by end-users to upload pictures that are then analyzed with OCR.

Request #2

Example answer (you can delete this)

P1: support for right-to-left languages Arabic and Hebrew.

Request #3

Example answer (you can delete this)

Ability to disable adaptive card buttons once they have been actioned.

Why do we ask these questions?

- Help us prioritize our future investments by submitting feature requests and detailing the business impact they have for your organization by using the link below.

Useful resources:

- aka.ms/CopilotStudioFeatureRequest

Dynamics 365 Omnichannel for Customer Service *(optional)*

-
- Chat widget integration & customizations
 - Omnichannel hand-off
 - Unified Routing Configuration

Note: it is expected for Dynamics 365 Omnichannel for Customer Service to have an extended implementation review, typically with FastTrack teams. This chapter calls specific items which are important for the Copilot Studio and Omnichannel integration.

Chat widget integration & customization

✔ How are you modifying the chat widget in Dynamics 365 Omnichannel?

Please describe the planned use of the chat widget.

Do you plan to develop a custom version of the widget?

Do you have other client requirements for specific channels?

What your expectations for the widget load time and first message display?

How are you making sure the first message is fast?

Example answer (you can delete this)

- Use of the LCW v2 configured with the brand theme.
- Many context variables, including web page, user ID, locale, etc. must be passed from the website to the chat widget.
- Widget and first message should load in less than 5 seconds.
- Internal mobile app teams are also looking to integrate their native app with Omnichannel through APIs.

Why do we ask these questions?

- Within Dynamics 365 Omnichannel for Customer Service, there is a customized Copilot Studio chat widget which is utilized rather than the native Copilot Studio widget without Dynamics 365 Omnichannel. This customized out-of-the-box widget is a component that Dynamics 365 customers can utilize to get a greater number of options to tailor the widget without having to spend a lot of time on the configuration every single time.
- Some of these options include the design but also include functionality like business hours and even channel specific behavior like persistent chat experiences.
- Ensure you have reviewed these options within the chat widget
- Additionally, ensuring you take into consideration the connection experience between Copilot Studio and Omnichannel in the 'Agent Hand-off' area of the Copilot Studio portal, and track your environments based on your environment structure, as well as the copilot users within each environment

Useful resources:

- [Configure the chat widget](#)
- [Develop a custom live chat widget](#)
- [Omnichannel Live Chat Widget SDK](#)

Copilot Studio & Omnichannel hand-off

✔ Connect your Copilot Studio copilot to Dynamics 365 Omnichannel

Which topics within will link to the 'Escalate' topic?

Do you need to extend the 'Escalate' topic? If so, how and why?

Would you be utilizing both the 'Escalate' and 'Fallback' topic?

Have you considered what the waiting experience for a handover to Dynamics 365 looks like and any limitations of the default experience?

Example answer (you can delete this)

- Standard behavior will be maintained to escalate.
- Over the voice channel, DTMF key 0 will escalate to a live agent.
- Fallback topic will not be used over the voice channel but will be used on the website: generative answers will be used before the user can escalate.

Why do we ask these questions?

- You can directly use the 'Escalate' system topic to be able to escalate to a live agent in Dynamics 365 Omnichannel within Copilot Studio.
- Make sure you consider the experience of both the user and the agent when designing this functionality.
- Consider what data you wish the Dynamics 365 Omnichannel agent to receive from the copilot when a handover is successfully taken place e.g., data passed to the agent.

Useful resources:

- [Configure the Copilot Studio copilot](#)

Unified routing configuration

✓ How is your copilot being routed?

Have you considered the routing experience when a copilot escalates a conversation?

Which queue is it being routed to?

You may have a blueprint where the copilot is escalated but re-directed to another copilot, for example in another language. How are you differentiating this routing (e.g., conditions) managing those variables, and also managing a true escalate to human experience?

Example answer (you can delete this)

- Only one queue and support team.

Why do we ask these questions?

- Unified routing is critical because within Dynamics 365 Omnichannel this is how a copilot, or a user, is routed to another queue.
- Consider expanding on this slide by creating a routing diagram that can be used to describe the simple routing experience you are proposing on using.
- This would help identify any gaps or dependencies.

Useful resources:

- [Overview of unified routing](#)

Omnichannel Voice integration

✔ Using Copilot Studio as an Interactive Voice Response (IVR)

Do you plan to use Omnichannel Voice? If yes, please detail how.

Do you plan to have different conversation paths and formats, or even copilots based on the channel?

How are you planning to test it in real-life scenarios? (E.g., with background noises, etc.)

Do you have specific requirements for the voice channel? (E.g., DTMF, barge-in, silence detection...)

Example answer (you can delete this)

- Omnichannel Voice is used.
- Conversations and messages in general are designed to be shorter for the voice channel.
- Topics will have branching logic based on the channel.
- DTMF options are offered as much as possible, to navigate menus and to provide client IDs.
- Tests will be performed by group of pilot users testing in real-life conditions.
- Silence detection, barge-in, noise cancelation, are crucial for the voice channel success.

Why do we ask these questions?

- To route customers call to the best department, diagnose issues, collect information, and give recommendations, conversational IVR copilots speak to customers when they call in. Copilot Studio makes it easy to author IVR copilots and you can use the same copilots for other channels, like chat and voice.
- It's important that the implementation of Voice with Copilot Studio as an IVR is well planned and expectations appropriately set.

Useful resources:

- [Voice channel in Omnichannel](#)
- [Configure Copilot Studio bots for voice](#)
- [Use SSML to customize speech responses](#)

Thank you



Get started today



aka.ms/TryCopilotStudio



Learn more

Copilot Studio website

aka.ms/CopilotStudio

Blog

aka.ms/CopilotStudioBlog

Demo

aka.ms/CopilotStudioDemo

Product documentation

aka.ms/CopilotStudioDocs

Product guidance

aka.ms/CopilotStudioGuidance

Implementation guide

aka.ms/CopilotStudioImplementationGuide

Community page

aka.ms/CopilotStudioCommunity