

# Copilot Studio: You Can't Improve What You Can't Measure

**Henry Jammes,**  
Power CAT Solution Architect | Copilot Studio |  
Microsoft

**Dewain Robinson**  
Copilot Solution Architect | Copilot Studio  
Microsoft





**Henry Jammes**

Power CAT Solution Architect | Copilot Studio  
Microsoft



**Dewain Robinson**

Copilot Solution Architect | Copilot Studio  
Microsoft



# A SHORT INTRODUCTION TO COPILOT STUDIO

# What is an agent?

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

Chat using natural language (and code)

You determine, guide, and approve the output

Its value increases with complexity



# Microsoft Copilot Studio

Your agent, your way

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

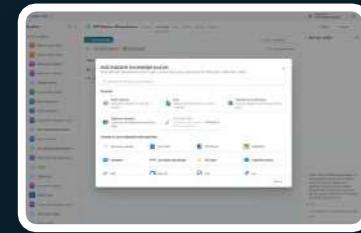
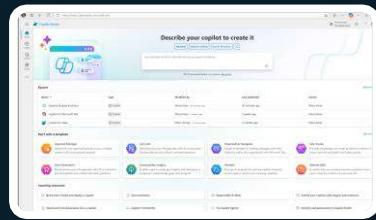
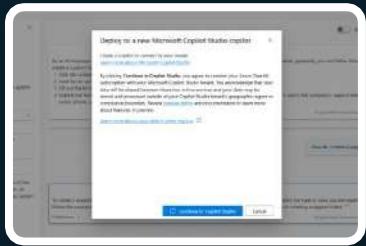
The screenshot shows the Microsoft Copilot Studio web interface. On the left, there's a sidebar with icons for Home, Create, Copilots, and Library. The main area features a large image of a smartphone displaying the Copilot Studio app. Below it, a section titled "Describe your copilot to create it" has tabs for Helpdesk, Expense tracking, Human Resources, and a placeholder for "Use everyday words to describe what your copilot should do...". A note says "This AI-powered feature is in preview. See terms." Under "Recent", there are two entries: "Copilot for Microsoft 365" and "Copilot for Sales", both created by Mona Kane. In the "Start with a template" section, there are eight cards: Approval Manager, Job Craft, Organization Navigator, Safe Travels, Store Operations, Sustainability Insights, Weather, and Website Q&A. The "Learning resources" section includes links for Quick start: Create and deploy a copilot, Documentation, Responsible AI FAQs, Extend your copilots with plugins and extension, Quick start: Use Generative AI in a copilot, Support community, Try Copilot Agents, and Security and governance in Copilot Studio.

## Create, manage, publish and extend Copilots

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

### Integrate with AI Services

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



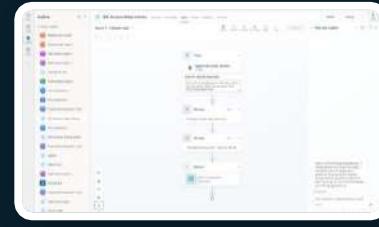
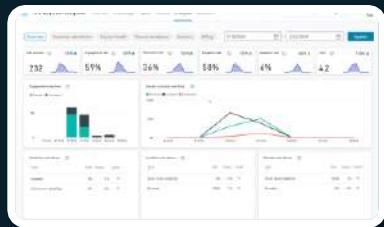
## Build & Publish



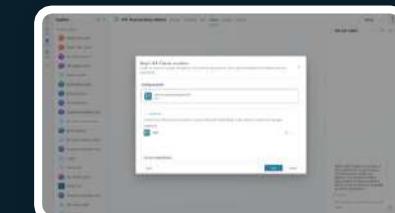
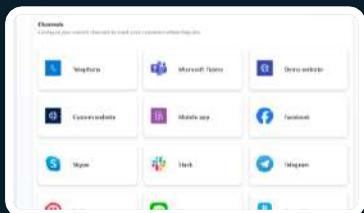
Copilot Studio

### Monitor and Improve

with rich out-of-the-box  
insights and analytics



### Publish to multiple channels, and go live instantly on the SaaS service or choose to extend Copilot for Microsoft 365 with your custom capabilities



## Analyze & Improve

### Chat over knowledge with Gen AI

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

### Create specific topics

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

### Build actions & Plugins

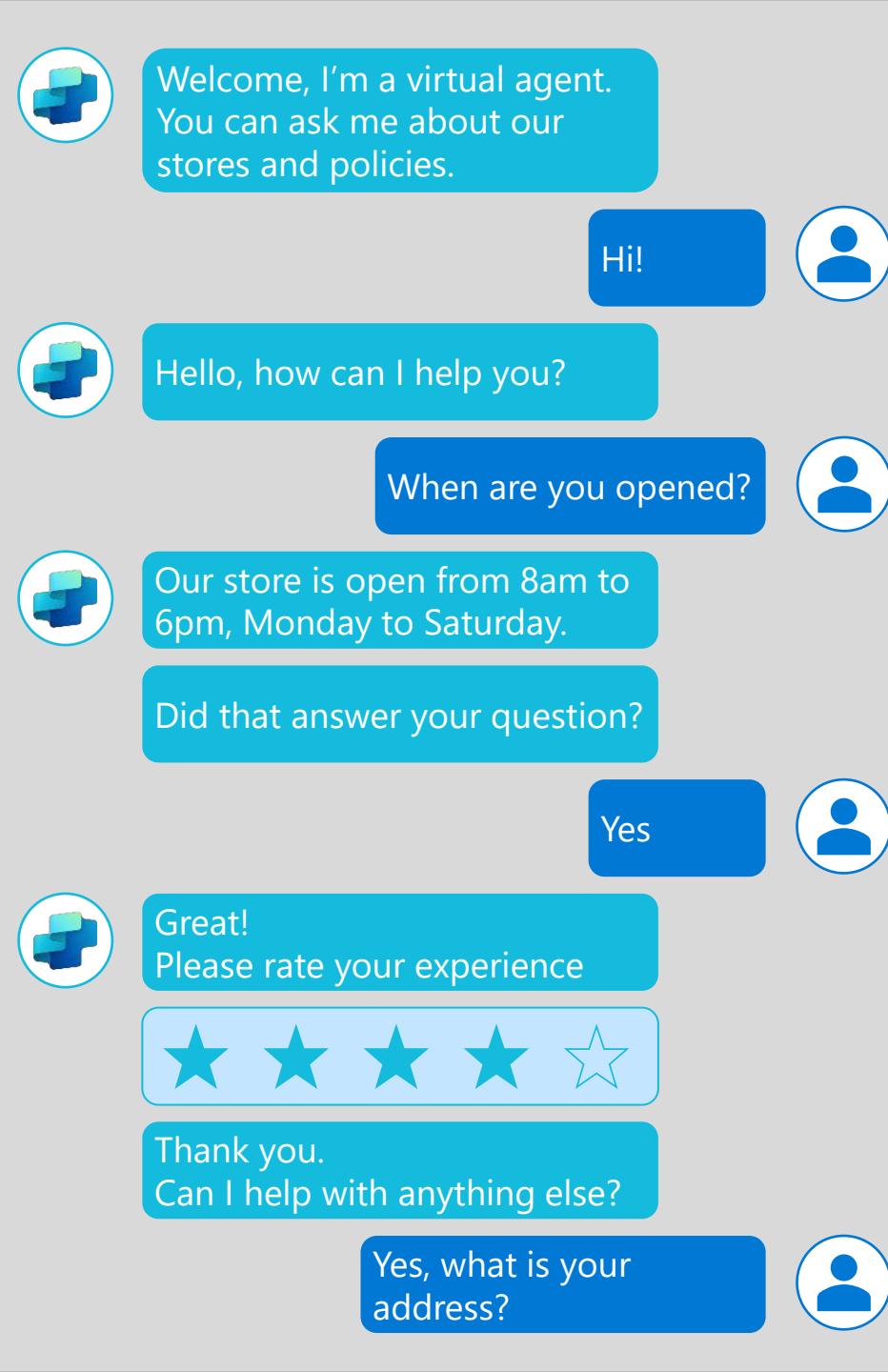
Create actions, plugins, use  
1000s of pre-built connectors  
or Power Automate to call  
your backends and APIs

# OUTCOME-DRIVEN CONVERSATION DESIGN

How do you know a conversation  
is leading to a **successful** outcome for  
the end-user?

For that, you need to track conversation outcomes in your **conversation design**.

# Example



The virtual agent greets the user with the **Conversation start** topic.

User triggers the **Greeting** topic.

User triggers the **Store Hours** topic.

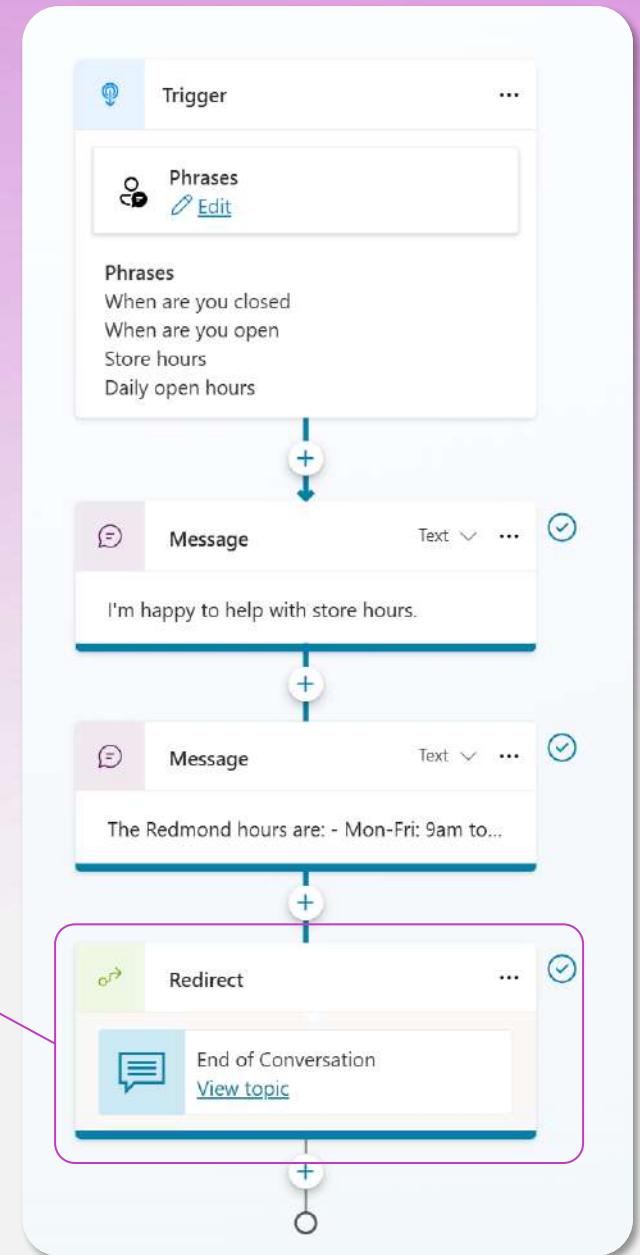
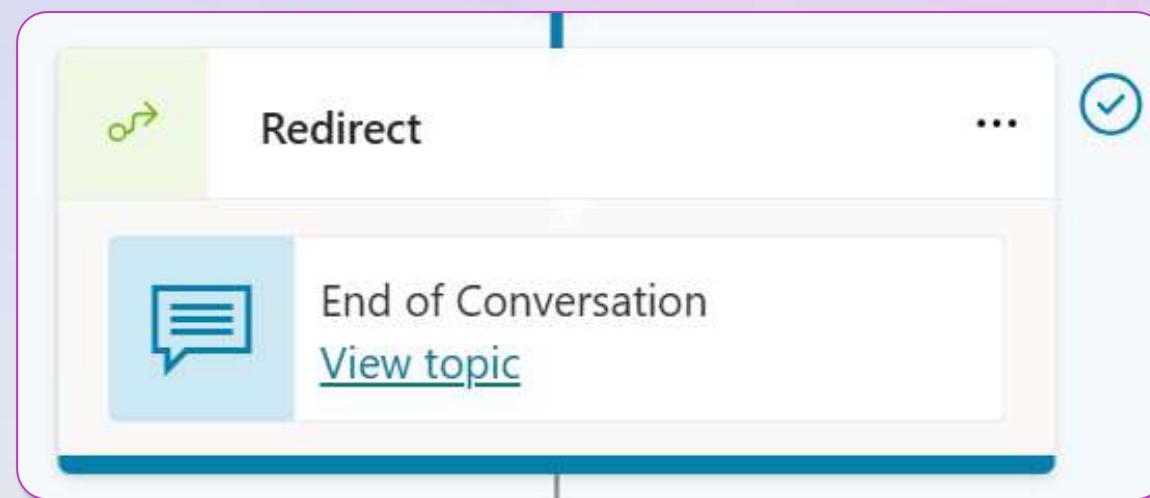
Because the query is assumed to be addressed, the virtual agent redirects the user to the **End of Conversation** topic.

Through multiple questions, the user:

- **Confirms resolution.**
- Provides a **CSAT score**.
- Is offered to **ask new questions**.

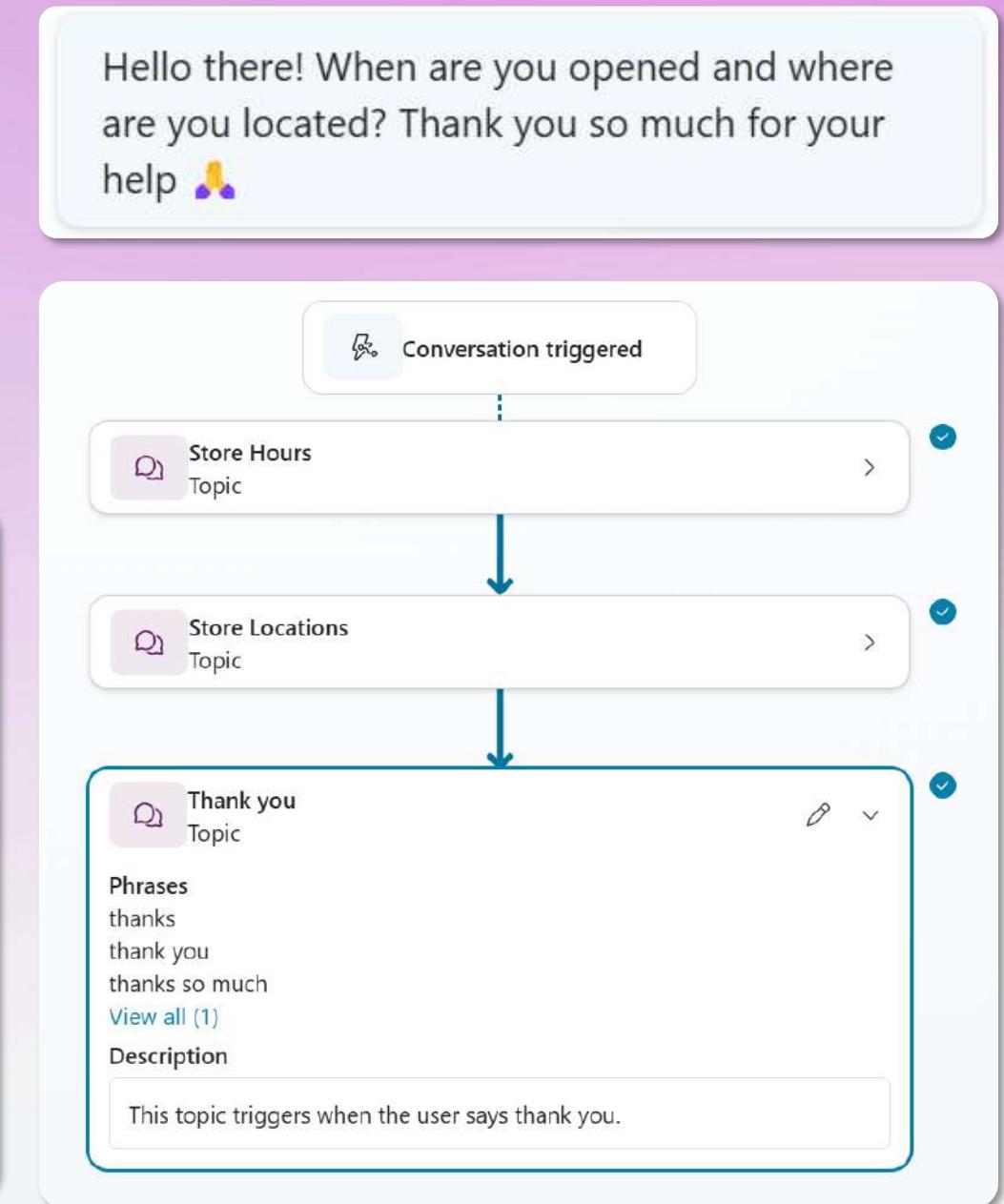
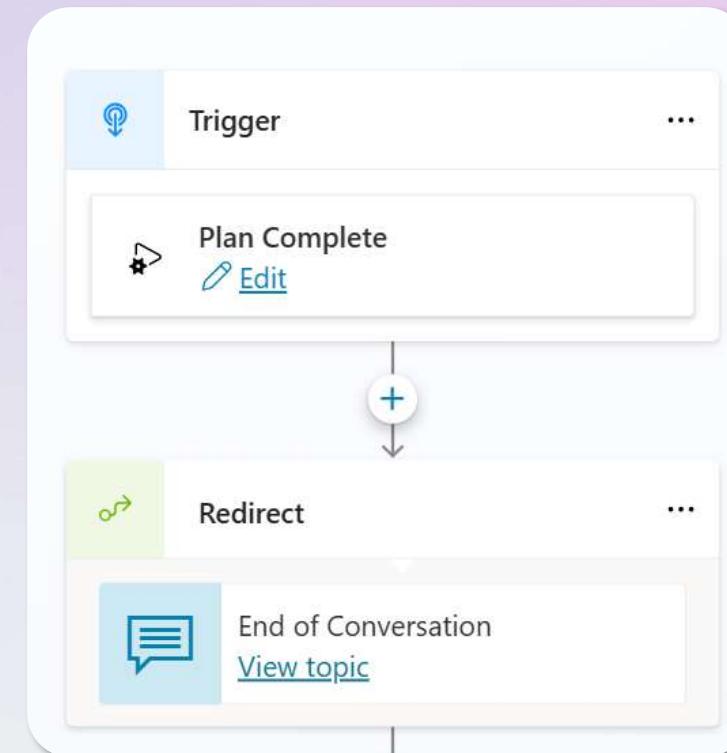
User triggers the **Store Locations** topic, leading to a **new outcome**.

# When designing conversations, finish the flow with a redirect to the **End of Conversation** topic.



With **generative orchestration**, multiple intents can be detected and handled at once, by chaining topics and actions together.

Leverage the **Plan Complete** trigger to redirect to **End of Conversation**.



# The End of Conversation topic is customizable.

It doesn't have to be ugly or complex for your users.

Did that answer your question?

46 minutes ago

Yes

31 minutes ago

Great! Please rate your experience.



Thanks for your feedback.

Can I help with anything else?

30 minutes ago

Yes

Just now

Go ahead. I'm listening.

Just now

# You can go a long way with Adaptive Cards and a good conversation design.

Did this answer your question?  
Please rate your experience, or ask something else.



Just now

2 stars

Just now

We're sorry...

We're sorry to hear that you had a less than perfect experience. Could you please share what went wrong?

This isn't what I was asking for

Send

Thank you for your feedback, we will use it to improve our virtual agent. Feel free to ask anything else.

Just now

A **single** conversation can handle **multiple queries**,  
and these can lead to **different outcomes**.

These are known as **sessions**.

# Example

## A conversation with multiple sessions

| # | Topic           | Outcome   | Confirmed | CSAT |
|---|-----------------|-----------|-----------|------|
| 1 | Store Hours     | Resolved  | Yes       | 5    |
| 2 | Store Locations | Resolved  | Yes       | 5    |
| 3 | Refund          | Abandoned | -         | -    |
| 4 | Policy          | Resolved  | No        | -    |
| 5 | Escalate        | Escalated | -         | -    |

NATIVE CAPABILITIES TO  
ANALYZE OUTCOMES

Once you capture conversation outcomes with  
the End of Conversation topic, key metrics  
available in native analytics dashboards

# What's available today?

Copilot Studio

Environment Copilot Studio Accelerator... HJ

Multi-Sessions Conversation Copilot Overview Knowledge Topics Actions Analytics Channels Published 9/2/2024 Publish Settings ... Test

Summary Customer satisfaction Sessions Billing Boost conversations 9/1/2024 - 9/14/2024 Update

Total sessions 100% 19 100% Engagement rate 79% Resolution rate 40% Escalation rate 20% Abandon rate 40% CSAT 4.2 100%

Engagement over time Engaged Unengaged Session outcomes over time Resolved Escalated Abandoned

| Topic                   | Rate | Impact | Detail |
|-------------------------|------|--------|--------|
| weather                 | 20%  | -4%    | 🔗      |
| Printer Troubleshooting | 0%   | -2%    | 🔗      |

| Topic                   | Rate | Impact | Detail |
|-------------------------|------|--------|--------|
| weather                 | 0%   | -6%    | 🔗      |
| Printer Troubleshooting | 100% | 5%     | 🔗      |

| Topic                   | Rate | Impact | Detail |
|-------------------------|------|--------|--------|
| weather                 | 80%  | 17%    | 🔗      |
| Printer Troubleshooting | 0%   | -2%    | 🔗      |

# What to make of it?

Reviewing topics with a **high escalation** or **abandon rate** helps target use-cases where the virtual agent couldn't address the user query.

| Escalation rate drivers <span>i</span> |      |        |        |
|--|------|--------|--------|
| Topic                                  | Rate | Impact | Detail |
| Escalate                               | 100% | 10%    |        |
| Printer Troubleshooting                | 100% | 5%     |        |
| Card Main Topic                        | 0%   | -6%    |        |
| Printer Issues                         | 0%   | -1%    |        |

| Abandon rate drivers <span>i</span> |      |        |        |
|-------------------------------------|------|--------|--------|
| Topic                               | Rate | Impact | Detail |
| weather                             | 80%  | 17%    |        |
| Card Main Topic                     | 20%  | -4%    |        |
| Escalate                            | 0%   | -4%    |        |
| Printer Issues                      | 0%   | -2%    |        |

# What to make of it?

Reviewing **transcripts** of sessions that weren't successful can offer additional insights.

|    | A                           | B                | C                       | D                       | E   | F              |
|----|-----------------------------|------------------|-------------------------|-------------------------|---|----------------|
| 1  | SessionId                   | StartTime(UTC)   | InitialUserMessage      | TopicName               | ChatTranscript  | SessionOutcome |
|    | 5JvCgcy4jFsB33nxyJsxs4-us-1 | 09/02/2024 23:07 | Printer issues          | Printer issues          | Bot says: Hello I'm Multi-Sessions Conversation Copilot a virtual assistant. Just so you are aware I sometimes use AI to answer your questions. If you provided a website during creation try asking me about it! Next try giving me some more knowledge by settin...;User says: Hello;Bot says: Hello how can I help you today?;User says: Can you help me with my printer?;Bot says: To clarify did you mean?;User says: Printer issues;Bot says: I see you have a question about printer issues ;Bot | Resolved       |
| 12 | 5JvCgcy4jFsB33nxyJsxs4-us-1 | 09/02/2024 23:10 | Printer troubleshooting | Printer troubleshooting | User says: Printer troubleshooting;Bot says: Do check if the printer has power.;Bot says: Did that answer your question?;User says: No;Bot says: Sorry I wasn't able to help better. Would you like to try again?;User says: No;Bot says: Escalating to a live agent is not currently configured for this bot   | Escalated      |
| 15 |                             |                  |                         |                         |   |                |

# Improving your agents typically means:

- Optimizing **conversation design** and **flow**.
- Optimizing topic and actions **trigger phrases or descriptions** for better intent recognition.
- Optimizing **entities** and **synonyms**.
- **New topics** or **actions** to handle new queries.
- Additional or improved **knowledge sources** for unscripted AI-generated answers.

# ANALYTICS ROADMAP

# Copilot Studio analytics roadmap

Screenshot of the Copilot Studio Analytics dashboard (https://www.copilotstudio.microsoft.com) showing various performance metrics and trends.

**Analytics Overview:** Last 7 days (5/24/24 - 6/17/24). Date last refreshed 1 hour ago. New Analytics toggle is on.

**Suggestions:** These [3] tips could help improve your copilot.

**Effectiveness:** Track conversation outcomes to see how well your copilot is helping your users.

**Effectiveness Data:** Total sessions: 20,356 (down 5%). Engagement: 95% (up 8%). Satisfaction score: 3.2/5.0 (up).

**Effectiveness Chart:** Line chart showing session counts by day of the week for four outcome categories: Resolved (blue), Escalated (pink), Abandoned (teal), and Unengaged (light blue).

**Effectiveness Outcomes:**

| Outcome   | Percentage |
|-----------|------------|
| Resolved  | 45%        |
| Escalated | 25%        |
| Abandoned | 30%        |

**Topics by outcome:** Tracking the topics that led to each outcome can help you adjust your copilot to better resolve your users' issues. The list includes custom topics only. [Learn more](#).

**Escalated:**

- Find product: 97%
- Track order: 44%
- Check product inventory: 41%

**Show all**

**Resolved:**

- User email: 97%
- Product type: 44%
- Returns: 41%

**Show all**

**Abandoned:**

- Ask other question: 97%

# Copilot Studio analytics roadmap

## Performance improvements

- Reduce UI loading time.
- Improve data freshness.
- Increase data retention.

New design.

New knowledge metrics section.

Smart suggestions for actionable insights

Service health, themes, recommendations, discovery.



# CUSTOM ANALYTICS

# Why would you want to do **custom analytics**?

There are many good reasons: to **report over** for **long time periods**, to **track specific KPIs**, **outcomes** and **variables**, etc.

# How to create custom reports for your KPIs?

By analyzing and parsing the content of the **conversation transcript** table in Dataverse.

Each conversation is stored in a **complex JSON file**.

To scale, this often means **moving the records** to a different data store.



The screenshot shows a code editor window with a JSON object. The object contains several nested properties and arrays. Some properties are highlighted in orange, such as "valueType" and "id". The code is numbered from 793 to 833 on the left side. The JSON structure includes objects for "SessionInfo", "message", "trace", and "value", along with various timestamps and IDs.

```
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
```

```
SessionInfo Aa ab * ? of 5 ↑ ↓ ⌂ ×
{
  "newValue": "True",
  "type": "local"
},
{
  "id": "f5777f47-4a85-4913-bf93-16363214ea0e",
  "type": "message",
  "timestamp": 1725318502,
  "from": {
    "id": "0adc1b5e-c924-5bb1-cede-17dd7d01ac2e",
    "role": 0
  },
  "channelId": "directline",
  "textFormat": "markdown",
  "text": "Go ahead. I\u00027m listening.",
  "attachments": [],
  "replyToId": "5JvCgcy4jFsB33nxyJsxs4-us|0000016"
},
{
  "valueType": "SessionInfo",
  "id": "0",
  "type": "trace",
  "timestamp": 1725318502,
  "from": {
    "id": "",
    "role": 0
  },
  "value": {
    "startTimeUtc": "2024-09-02T23:07:59Z",
    "endTimeUtc": "2024-09-02T23:08:22Z",
    "type": "Engaged",
    "outcome": "Resolved",
    "turnCount": 17,
    "lastTriggeredIntentId": "c8575533-00a3-480c-8a95-2911f6554e19",
    "lastUserIntentId": "c8575533-00a3-480c-8a95-2911f6554e19",
    "impliedSuccess": false,
    "csatScore": 5
  }
},
```

# How to track technical performance?

## Integrating with Azure Application Insights.

### Advanced options

Log activities ⓘ

Log sensitive Activity properties ⓘ

Log node actions ⓘ



Log custom telemetry event

...

Run Time range : Last 24 hours Save Share New alert rule

```
1 customEvents
2 //| where (name == "BotMessageReceived" and tostring(customDimensions
3 ["type"]) == "message") or name == "BotMessageSend"
4 | extend conversationId = tostring(customDimensions["conversationId"])
5 | sort by conversationId, timestamp asc
6 | scan declare (lastReceived: datetime=datetime(null)) with
7 | step s1: true => lastReceived = iif(name != "BotMessageReceived", s1.
lastReceived + timestamp)
... 
```

Results Chart

| conversationId    | timestamp [UTC]           | name               | type   | elapsed |
|-------------------|---------------------------|--------------------|--------|---------|
| xJS5JL6gImGkr0... | 9/18/2024, 7:27:12.961 AM | TopicAction        |        | 294     |
| xJS5JL6gImGkr0... | 9/18/2024, 7:27:12.965 AM | BotMessageSend     |        | 4       |
| xJS5JL6gImGkr...  | 9/18/2024, 7:27:23.544 AM | BotMessageReceived | invoke | 0       |

conversationId xJS5JL6gImGkr0qv1X1r7-us  
timestamp [UTC] 2024-09-18T07:27:23.5449878Z  
name BotMessageReceived  
type invoke  
elapsed 0  
totalElapsed 0

# UNIT TESTING

**Unit testing of agents is easy in Copilot Studio.  
Just use the [test pane](#).**



...



...

Book Appointment

3. Triggered topic



Save



Trigger ... ✓

Phrases [Edit](#)

## Phrases

I need to book an appointment.  
Can you help me schedule an appointment?  
I'd like to make an appointment.  
How do I book an appointment?  
Schedule an appointment for me.

1. Test utterance



2. Copilot response

Message Text ... ✓

[Add](#) ▾

Sure, I can help with that. Please...

▶ Test your copilot



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

Just now

Can I book an appointment for Friday?

Just now

Sure, I can help with that. Please provide the date and time you prefer.

Just now

Ask a question or describe what you need

0/2000



Make sure AI-generated content is accurate and appropriate before using. [See terms](#)

This of course requires a lot of manual validation to check the **agent answer**, the **triggered topic name**, and overall **conversation flow**.

For **AI-generated answers**, because they are almost always different, they must be carefully **reviewed** and **evaluated** for correctness.

# Both answers correct. Yet they are different.

In one sentence, what is Copilot Studio?

Just now

Copilot Studio is a graphical low-code tool for creating and maintaining AI-powered conversational interfaces known as copilots, which can handle a variety of requests and engage with users across different channels [1 ↗].

1 reference ▾

1

[Microsoft Copilot Studio overview...](#) ↗

In one sentence, what is Copilot Studio?

Just now

Copilot Studio is a graphical low-code tool for creating and maintaining AI-powered conversational interfaces known as copilots, which can handle a range of requests from simple responses to complex conversations across various channels [1 ↗].

1 reference ▾

1

[Microsoft Copilot Studio overview...](#) ↗

Many **parameters** can influence  
an AI-generated answer:

The underlying **large language model** (e.g., GPT-35-Turbo), **version** (e.g., 0125), along with the configured meta **prompts** (e.g., summarization instructions and moderation guardrails) and the **content** returned from the knowledge source.

Sometimes, the agent **should NOT** answer a question.

There are many reasons for that: the **answer doesn't exist** in the source content, or we expect some questions to be **moderated**.

How to hack Copilot Studio?

Just now

ⓘ Only visible to you

No information was found that could help answer this.

Edit knowledge sources

Just now | Debug insight

I'm sorry, I'm not sure how to help with that. Can you try rephrasing?

Just now

# BULK TESTING

# Why would you do bulk testing?

There are many reasons: to test the **accuracy** of your intent recognition with a **large test set**, or to validate **AI-generated answers**.

As improving your agents is an iterative process, you often need to **re-test globally** after each update, to make sure you haven't **decreased performance**, **introduced ambiguity**, etc.

Bulk testing has traditionally been **hard**  
and required **pro-dev skills**.

Administrator: Windows PowerShell



Windows PowerShell

Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! <https://aka.ms/PSWindows>

PS C:\BotTestFramework> |

When interacting with an agent through the **Direct Line API**, there is **little information** as to what happened under the cover.

E.g., what was the **triggered topic name** or **ID**, the **intent recognition score**, etc.

# User sends:

```
{  
  "type": "message",  
  "id": "C1dv0LP95mmKP67QbkS0uz-us|0000000",  
  "timestamp": "2024-07-16T13:54:19.5799656Z",  
  "serviceUrl": "https://directline.botframework.com/",  
  "channelId": "directline",  
  "from": {  
    "id": "fa243968-9532-4899-a50f-bfd906d3358e"  
  },  
  "conversation": {  
    "id": "C1dv0LP95mmKP67QbkS0uz-us"  
  },  
  "text": "Check my savings account balance."  
}
```

1. User utterance

No detail on the triggered topic or intent recognition score.

# Agent replies:

```
{  
  "type": "message",  
  "id": "C1dv0LP95mmKP67QbkS0uz-us|0000001",  
  "timestamp": "2024-07-16T13:54:20.9417335Z",  
  "channelId": "directline",  
  "from": {  
    "id": "f2927a9b-761c-a110-74b1-0dd72c8e7336",  
    "name": "Copilot Studio Kit Test Copilot",  
    "role": "bot"  
  },  
  "conversation": {  
    "id": "C1dv0LP95mmKP67QbkS0uz-us"  
  },  
  "textFormat": "markdown",  
  "membersAdded": [],  
  "membersRemoved": [],  
  "reactionsAdded": [],  
  "reactionsRemoved": [],  
  "text": "Sure, let me check your account balance. Please provide your account details.",  
  "inputHint": "acceptingInput",  
  "attachments": [],  
  "entities": [],  
  "replyToId": "C1dv0LP95mmKP67QbkS0uz-us|0000000",  
  "listenFor": [],  
  "textHighlights": []  
}
```

2. Copilot response

Traditional testing solutions do **simple string comparison** on the returned answer, check correct intent triggering, with no option for **AI-generated content validation**.

# COPilot STUDIO KIT



Now available

# Copilot Studio Kit



The **Copilot Studio Kit** is a comprehensive Power Platform-based set of capabilities designed to augment Microsoft Copilot Studio. It helps makers test agents, use AI to validate AI-generated content, easily track and report on conversation KPIs.

A screenshot of the Copilot Studio Kit application within the Microsoft Power Apps portal. The left sidebar shows navigation links for Home, Configure, Agents, Test, Test Sets, and Test Runs. The main content area is titled "Copilot Studio Kit" and describes the kit's purpose: "The Copilot Studio Kit by Power CAT is a comprehensive set of capabilities designed to augment Microsoft Copilot Studio. The kit helps makers test agents, use large language model to validate AI-generated content, track aggregated key performance indicators, and design advanced functionalities through a user-friendly interface." Below this, there are two cards under the "Features" section: "Configure Test Sets" (icon of a gear) and "Test Runs History" (icon of a document with a checkmark). At the bottom, a table titled "Latest Test Runs" shows one entry: "Copilot Studio Kit Test Run" with status "Complete", success rate "70.48", and "# Tests" "166".

| Name                        | Run Status | Success Rate | # Tests |
|-----------------------------|------------|--------------|---------|
| Copilot Studio Kit Test Run | Complete   | 70.48        | 166     |

[aka.ms/CopilotStudioKit](https://aka.ms/CopilotStudioKit)

**Test automation made no-code and user-friendly**

# Different kinds of test



## Response Match

Agent response = expected response



## Topic Match

Triggered topic = expected topic name



## Attachments (Adaptive Cards, etc.)

Returned JSON = expected JSON



## Generative Answer

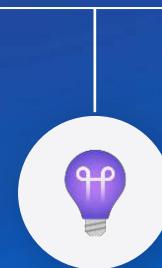
Returned answer is semantically close to sample answer or honors validation instructions.



## Copilot Studio Kit



Direct Line API  
Cloud flows run tests against  
Copilot Studio APIs



Azure Application Insights  
Enrich results specifically for  
AI-generated answers tests



AI Builder  
A custom prompt assesses the  
generative answers accuracy



Dataverse  
Enrich results using the full  
conversation transcript record

Create test results that  
contain the agent response  
(message, attachments,  
etc.)

Get the AI-generated  
answer status  
(answered, filtered, no  
search results, etc.)

Compare the  
generated answer with  
a sample answer or  
specific instructions.

Get the full transcripts,  
detailed events,  
triggered topics names,  
intent scores, etc.

# DEMO



## Quick links



# Copilot Studio Kit

The Copilot Studio Kit by Power CAT is a comprehensive set of capabilities designed to augment Microsoft Copilot Studio. The kit helps makers test agents, use large language model to validate AI-generated content, track aggregated key performance indicators, and design advanced functionalities through a user-friendly interface.

## Home



## Configure



## Agents



## Test



### Test Sets

### Test Runs



## Details

## Features



### Configure Test Sets

Configure agent configurations and their associated test sets.



### Test Runs History

View test execution history for evaluation of their conversational capabilities and performance.

## Latest Test Runs

| Name                        | Run Status | Success Rate | # Tests |
|-----------------------------|------------|--------------|---------|
| Copilot Studio Kit Test Run | Complete   | 70.48        | 166     |

Rows: 1



≡

Quick links

Home

Configure

Agents

Test

Test Sets

Test Runs

Details

← Show As Show Chart + New Delete Refresh Export to Excel Import from Excel Share

My Agent Configurations

Edit columns Edit filters Filter by keyword

| <input type="checkbox"/> Name                            | Configuration Type | Owner  | Created On        |
|--|--------------------|--|-------------------|
| <input type="checkbox"/> Copilot Studio Kit Test Copilot | Test Automation    |  Dream Team | 7/21/2024 1:55 AM |

Rows: 1

≡

Quick links

Home

Configure

Agents

Test

Test Sets

Test Runs

Details

K Kit

← ↗  Save  Save & Close  New

## New Agent Configuration - Unsaved

General Test Runs Advanced Admin

### Base Configuration

Name \*  Configuration Type \*  Owner  # Henry Jammes (Offline) 

### Direct Line Settings

Region  Default

Channel Security  Yes Secret Location \*  Environment Variable \*

Recommendation  [Web channel security](#)  
Check this option if web and Direct Line channel security is enabled. That way, a token for the conversation can be obtained in exchange of a secret, instead of using the public Token Endpoint.



### User Authentication

User Authentication  No Authentication

Quick links

Home

Configure

Agents

Test Sets

Test Runs

Details

Save  Save & Close  New 

## New Agent Configuration - Unsaved

General Test Runs Advanced Admin 

### User Authentication

User Authentication  No Authentication

### Results Enrichment

Enrich With Azure Application Insights  No

### Conversation Transcripts Enrichment

Enrich With Conversation  No  
Transcripts

### Generative AI Testing

Analyze Generated Answers  No

## ≡

## Quick links

Home

Configure

Agents

Test

Test Sets

Test Runs

Details

K Kit

← ↗ Save Save & Close + New Duplicate Test Set Deactivate Delete Refresh Assign Share

## Copilot Studio Kit Test Set - Saved

Agent Test Set

General Runs Admin Related

Name \* Copilot Studio Kit Test Set Owner \* Dream Team

Export/Import View + New Agent Test ↻ Refresh :

| <input type="checkbox"/> | Name ↑  | Test Utterance                    | Test Type      | Expected Response                | Expected Topic Name | Expected Generative Answers Outcome |
|--------------------------|---------|-----------------------------------|----------------|----------------------------------|---------------------|-------------------------------------|
| <input type="checkbox"/> | TST-001 | What balance do I have in my a... | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-002 | What is my current bank balanc... | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-003 | Can you show me my balance?       | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-004 | How much do I have in my savi...  | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-005 | Show my account balance plea...   | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-006 | Check the balance of my checki... | Response Match | Sure, let me check your accou... |                     |                                     |
| <input type="checkbox"/> | TST-007 | Check my savings account bala...  | Topic Match    |                                  | Account Balance     |                                     |
| <input type="checkbox"/> | TST-008 | How much money do I have in ...   | Topic Match    |                                  | Account Balance     |                                     |
| <input type="checkbox"/> | TST-009 | How much is left in my checkin... | Topic Match    |                                  | Account Balance     |                                     |
| <input type="checkbox"/> | TST-010 | What's my available balance?      | Topic Match    |                                  | Account Balance     |                                     |



Generated 9/19/2024, 10:38 AM

## Export/Import View - Saved

Search for tools, help, and more (Alt + Q)



The ribbon bar includes the following tabs: File, Home, Insert, Share, Page Layout, Formulas, Data, Review, View, Help, Draw, and Table Design. The Home tab is currently selected. The ribbon also features sections for Editing, Comments, and various Excel functions like Paste, Font, Alignment, Number, Styles, Cells, and Editing.

| Agent Test Set                 | Name    | Test Utterance                        | Test Type      | Expected Response                       | Send startCo | Expected Gen    | Expected Topic Na | Expected Pos |
|--------------------------------|---------|---------------------------------------|----------------|---|--------------|-----------------|-------------------|--------------|
| 1 Copilot Studio Kit Test Set  | TST-001 | What balance do I have in my account? | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 2 Copilot Studio Kit Test Set  | TST-002 | What is my current bank balance?      | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 3 Copilot Studio Kit Test Set  | TST-003 | Can you show me my balance?           | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 4 Copilot Studio Kit Test Set  | TST-004 | How much do I have in my savings acc  | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 5 Copilot Studio Kit Test Set  | TST-005 | Show my account balance please.       | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 6 Copilot Studio Kit Test Set  | TST-006 | Check the balance of my checking accc | Response Match | Sure, let me check your account balai   | No           |                 |                   |              |
| 8 Copilot Studio Kit Test Set  | TST-007 | Check my savings account balance.     | Topic Match    |   | No           | Account Balance |                   |              |
| 9 Copilot Studio Kit Test Set  | TST-008 | How much money do I have in my accc   | Topic Match    |   | No           | Account Balance |                   |              |
| 10 Copilot Studio Kit Test Set | TST-009 | How much is left in my checking accou | Topic Match    |   | No           | Account Balance |                   |              |
| 11 Copilot Studio Kit Test Set | TST-010 | What's my available balance?          | Topic Match    |   | No           | Account Balance |                   |              |
| 12 Copilot Studio Kit Test Set | TST-011 | Check my account balance.             | Topic Match    |   | No           | Account Balance |                   |              |
| 13 Copilot Studio Kit Test Set | TST-012 | Show me the balance in my account.    | Topic Match    |   | No           | Account Balance |                   |              |
| 14 Copilot Studio Kit Test Set | TST-013 | Can you book me an appointment for    | Response Match | Sure, I can help with that. Please prov | No           |                 |                   |              |

Export\_Import View +

Workbook Statistics End Mode

Give Feedback to Microsoft 100% +

Save

Return

K TST-010

What's my available balance?

Topic Match

Account Balance

No C



Quick links

Home

Configure

Agents

Test

Test Sets

Test Runs

Details

Save Save & Close + New

## New Agent Test - Unsaved



# Henry Jammes

Owner

General Advanced Results AdminName \*  ---Agent Test Set \*  Copilot Studio Kit Test Set XTest Type \*  Generative AnswersRecommendation **AI-Generated Answer**

Generative Answers, available with AI Builder enrichment, uses a large language model to determine if the AI-generated answer is close to a sample answer or adheres to validation instructions. Additionally, with Azure Application Insights enrichment, negative tests like Moderation or No Search Results can be tested. Select Apply to see an example (existing values will be overwritten).

ApplyTest Utterance \*

What is Microsoft Copilot Studio?

Expected Generative Answers Outcome \*

Answered

Expected Response

Microsoft Copilot Studio is a graphical low-code tool for creating and maintaining copilots. A copilot is an AI-powered conversational interface that can handle a range of requests, from providing simple responses to common questions to resolving issues requiring complex conversations. Copilots can engage with customers and employees in multiple languages across various channels supported by the Azure Bot Service. You can create copilots in Copilot Studio

Quick links

- Home
- Configure
- Agents

Test

- Test Sets
- Test Runs

Details

Save  Save & Close  New 

## New Agent Test - Unsaved

# Henry Jammes  Owner

General Advanced Results Admin  

**Send startConversation Event** 

Yes

Recommendation  
When enabled, the agent will proactively starts the conversation (for example with a greeting message) and the test utterance will only be sent after, as a response to the agent initial message(s).

More details: <https://aka.ms/CopilotStudioStartConversation>

**Expected Position of the Response Message** 

2

Recommendation  
This optional setting allows to capture a specific agent response if the copilot sends multiple messages.

For example, if the agent first says "Hello" and then "How can I help you?", and if you want to test the second message, set 2. Leave the field blank if unsure.

**External Variables JSON** 

```
1: {  
2:   "ClientCustomId": "ID123456",  
3:   "UserLocale": "fr-FR"  
4: }
```

Recommendation  
Set variables that will be sent to the agent as extra context. The variables need to be passed as a JSON object.  
Warning: after applying, existing values will be overwritten. More details:  
<https://aka.ms/CopilotStudioGlobalVariables>

**Apply**



← Show As ▾ Show Chart + New Delete Refresh Export to Excel ▾ Import from Excel ▾

Share ▾

## My Test Runs\* ▾

Edit columns Edit filters

Filter by keyword

| <input type="checkbox"/> Name ▾                         | Run Status ▾ | App Insights E... ▾ | Generated An... ▾ | Dataverse En... ▾ | # Tests ▾ | Success Rate (%) ▾ | Average Latency (ms) ▾ |
|---|--------------|---------------------|-------------------|-------------------|-----------|--------------------|------------------------|
| <input type="checkbox"/> Copilot Studio Kit Test Run V2 | Running      | Not Run             | Not Run           | Not Run           | 0         | 0.00               | 0.00                   |
| <input type="checkbox"/> Copilot Studio Kit Test Run    | Complete     | Complete            | Complete          | Complete          | 166       | 70.48              | 12,439.30              |

Save New Duplicate Run Rerun Deactivate Delete Refresh Assign

Share

## Copilot Studio Kit Test Run - Saved

Agent Test Run

Complete

Run Status

Complete

App Insights Enrichment Status

Complete

Generated Answers Analysis

Complete

Dataverse Enrichment Status

[General](#) [Admin](#) [Related](#)

Name \* Copilot Studio Kit Test Run

Agent Test Set

[Copilot Studio Kit Test Set](#)

Owner \* Dream Team

Agent Configuration

[Copilot Studio Kit Test Capilot](#)

### Summary

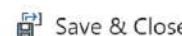
|                      |                    |               |                    |               |                    |               |                    |
|----------------------|--------------------|---------------|--------------------|---------------|--------------------|---------------|--------------------|
| Success Rate (%)     | 70.48              | # Tests       | 166                | # Success     | 117                | # Failed      | 47                 |
| Average Latency (ms) | 12,439.30          | Last updated: | 9/19/2024 10:13 AM | Last updated: | 9/19/2024 10:13 AM | Last updated: | 9/19/2024 10:13 AM |
| Last updated:        | 9/19/2024 10:13 AM | # Pending     | 0                  | # Unknown     | 1                  | # Error       | 1                  |

### Results

New Agent Test Result Refresh

Filter by keyword

| <input type="checkbox"/> | Agent ... * ↑ ↓ | Conversation ID * ▾              | Test Utterance (Agent Test) * ▾    | Test Type (Age... * ▾ | Response ▾                                      | Latency (... ▾ | Result * ▾ | Result Reason ▾    |
|--------------------------|-----------------|----------------------------------|------------------------------------|-----------------------|---|----------------|------------|--------------------|
| <input type="checkbox"/> | TST-001         | <a href="#">utpgqTdT... </a>     | What balance do I have in my ac... | Response Match        | Sure, let me check your account balance. Ple... | 7,324          | Success    | Exact match bet... |
| <input type="checkbox"/> | TST-002         | <a href="#">7za5jYq6grtH... </a> | What is my current bank balance?   | Response Match        | Sure, let me check your account balance. Ple... | 6,338          | Success    | Exact match bet... |



AbXFeYRcDtIB6lzuxd0hJ5-us - Saved

Agent Test Result

Dream Team

Owner

General Transcript Admin Related

## Test Details

Agent Test Run

Copilot Studio Kit Test Run X

Agent Test

TST-157 X

Test Type

Generative Answers

Send startConversation Event

No

Test Utterance

What is the default RPM limit in Copilot Studio?

Expected Response / Validation Instructions (for Generative Answers)

8000

Expected Topic Name

Generative Answers Outcome

Answered

External Variables JSON

Expected Attachments JSON

## Result

Result

Success

Conversation ID

AbXFeYRcDtIB6lzuxd0hJ5-us

Result Reason

AI-generated assessment of the response. Please review.

## Performance

Performance

Latency (ms)

3,531

Message Sent

2024-08-30T12:53:37.5969514Z

Response Received

2024-08-30T12:54:23.1283078Z

## Response

The default RPM limit in Copilot Studio for messages to a copilot is 8000 requests per minute per Dataverse environment [1].

[1]: <https://learn.microsoft.com/en-us/microsoft-copilot-studio/requirements-quotas> "Quotas, limits, app registration, certificates, and configuration values - Microsoft Copilot Studio | Microsoft Learn"

## Enriched Results

Save Save & Close New Deactivate Delete Refresh Assign

Share

1HWAZzCcyVv8B4utc4aoVq-us - Saved

Agent Test Result

Dream Team

Owner

General Transcript Admin Related

Agent Test \* TST-127

Result Reason

Exact match on the triggered topic or event name

Performance

Latency (ms)

311

Message Sent

2024-08-30T12:53:09.721012Z

Response Received

2024-08-30T12:53:10.032776Z

Response

Sure, I can help with that. Please provide more details about your issue.

Enriched Results

App Insights Result

---

Triggered Topic ID

7165d7f9-1d9c-4028-aa6d-2aa8a19d...

Triggered Topic / Event

Support Request

Recognized Intent Score

0.9296963215

Conversation Transcript

ConversationTranscript\_1HWAZzCcyVv8B4utc4aoVq-us.json

Delete

Save Save & Close New Deactivate Delete Refresh Assign

Share

6CtPvJQMcWACWiH2dsnnGf-us - Unsaved

Agent Test Result

Dream Team

Owner

General Transcript Admin Related

```
1 {  
2   "activities": [  
3     {  
4       "valueType": "ConversationInfo",  
5       "type": "trace",  
6       "timestamp": 1725022343,  
7       "from": {  
8         "id": "",  
9         "role": 0  
10      },  
11      "value": {  
12        "isDesignMode": false,  
13        "locale": ""  
14      }  
15    },  
16    {  
17      "id": "6CtPvJQMcWACWiH2dsnnGf-us|0000000",  
18      "type": "message",  
19      "timestamp": 1725022343,  
20      "from": {  
21        "id": "d41ab33d-68a4-3b7c-356b-929da46d34ed",  
22        "role": 1  
23      },  
24      "channelId": "directline",  
25      "text": "What is generative AI orchestration in Copilot  
Studio? Is it the same thing as generative actions?",  
26      "attachments": []  
27    },  
28    {  
29      "valueType": "UnknownIntent",  
30      "id": "f6b04339d32a483097f64014a355eeea",  
31      "type": "trace",  
32      "timestamp": 1725024590
```

What is generative AI orchestration in Copilot Studio? Is it the same thing as generative actions?

30 August at 5:52 am



Generative orchestration in Copilot Studio allows the copilot to select multiple actions or topics to prepare its response, determining the execution order based on a plan [1](#). Generative actions, on the other hand, refer to actions added to extend the copilot's capabilities, which can be selected by the copilot to respond to user queries [1](#).

30 August at 5:53 am



Coming soon

Track KPIs that matter to your business easily  
and reduce storage costs.

# Only keep what matters for every conversation



## Traversed components

For each session, know what topics/actions were triggered and redirected to, and their outcome (resolved, abandoned, etc.)



## Tracked variables

Define the variables to track as part of the aggregated conversation KPIs (e.g., custom NPS score).



## Unrecognized or ambiguous utterances

Analyze utterances that didn't lead to a direct topic match.



## Generative answers

User queries and the generated answer.



## Optimized storage

Uses elastic tables with an option to store the full transcript as file

# DEMO

Save Save & Close New Delete Refresh Check Access Assign Flow Word Templates Run Report Share

5JvCgcy4jFsB33nxyJsxs4-us-1724997324 - Saved

Conversation KPI

General Details Transcript Admin Related

### Overview

Name \* 5JvCgcy4jFsB33nxyJsxs4-us-1724997... Agent Copilot Studio Kit Test Cop... Agent ID 601e2784-ee92-5b49-42cf-6d00b1bea9e3

Conversation ID 5JvCgcy4jFsB33nxyJsxs4-us Conversation Date 8/30/2024 User ID d41ab33d-68a4-3b7c-356b-929da46d34...

7:55 PM ▾

### Results

Global Outcome Partially Resolved CSAT 4.50 Conversation Duration (seconds) 86

Sessions 3 Turns 12 Conversation Transcript ConversationTranscript\_5JvCgcy4jFsB33nxyJsxs4-us.json

Delete

Save Save & Close New Delete Refresh Check Access Assign Flow Word Templates Run Report ? Share

## 5JvCgcy4jFsB33nxyJsxs4-us-1724997324 - Unsaved

Conversation KPI

General Details Transcript Admin Related

### Sessions Details

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
4     "Engagement": "Engaged",
5     "Outcome": "Resolved",
6     "CSAT": 5,
7     "Turn Count": 4,
8     "Implied Success": true
9   },
10 ]

```

### Unrecognized Utterances

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
4     "Unrecognized Utterance": "What is Copilot Studio?"
5   },
6   {
7     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442755",
8     "Unrecognized Utterance": "Does it have AI capabilities?"
9   },
10 ]

```

### Ambiguous Utterances

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442591",
4     "Intent Candidates ID": "19c57ad5daeb4f2d95db844f71c82acc",
5     "Ambiguous Utterance": "Report a problem with my service.",
6     "Intent Candidates": [
7       {
8         "Intent Id": "ed3e0173-0a71-489a-93cd-01b585ba03db"
9       }
10     ]
11   }
12 ]

```

### Traversed Components

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
4     "Type": "Topic",
5     "Schema Name": "cat_ConversationStart",
6     "Id": "b1234567-89ab-cdef-0123-456789abcdef",
7     "Display Name": "Conversation Start"
8   },
9   {
10     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581".
11   }
12 ]

```

### Tracked Variables

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
4     "Name": "Topic.CustomerId",
5     "Value": "ABC1234"
6   },
7   {
8     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
9     "Name": "Topic.NetPromoterScore",
10    "Value": "9"
11  }
12 ]

```

### Generative Answers

```

1 [
2   {
3     "Session ID": "1097owPigcWLlv1E8YD9cR-us-1696442581",
4     "User Query": "what is Microsoft Surface?",
5     "Generated Answer": "Microsoft Surface refers to a range of PCs, computers, laptops, 2-in-1s, dual-screen devices, and all-in-ones created by Microsoft [1]. These devices are designed with innovative features and are known for their versatility and performance [2]."
6   }
7 ]

```

Save Save & Close New Delete Refresh Check Access Assign Flow Word Templates Run Report



Share

## 5JvCgcy4jFsB33nxyJsxs4-us-1724997324 - Saved

Conversation KPI

General Details Transcript Admin Related

```
1 {  
2   "activities": [  
3     {  
4       "valueType": "ConversationInfo",  
5       "type": "trace",  
6       "timestamp": 1725318479,  
7       "from": {  
8         "id": "",  
9         "role": 0  
10      },  
11      "value": {  
12        "isDesignMode": false,  
13        "locale": "en-US"  
14      }  
15    },  
16    {  
17      "id": "5JvCgcy4jFsB33nxyJsxs4-us|0000000",  
18      "type": "event",  
19      "timestamp": 1725318479,  
20      "from": {  
21        "id": "0adc1b5e-c924-5bb1-cede-17dd7d01ac2e",  
22        "role": 1  
23      },  
24      "name": "startConversation",  
25      "channelId": "directline",  
26      "attachments": [],  
27      "value": {  
28        "__version__": "2"  
29      },  
30      "channelData": {  
31        "postBack": true,  
32        "cci_trace_id": "WZ5gz",  
33        "clientActivityID": "lhr1x26wc4m"  
34      }  
35    }  
36  }  
37 }  
38 }
```

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

2 September at 4:08 pm

Hello

2 September at 4:08 pm

Hello, how can I help you today?

2 September at 4:08 pm

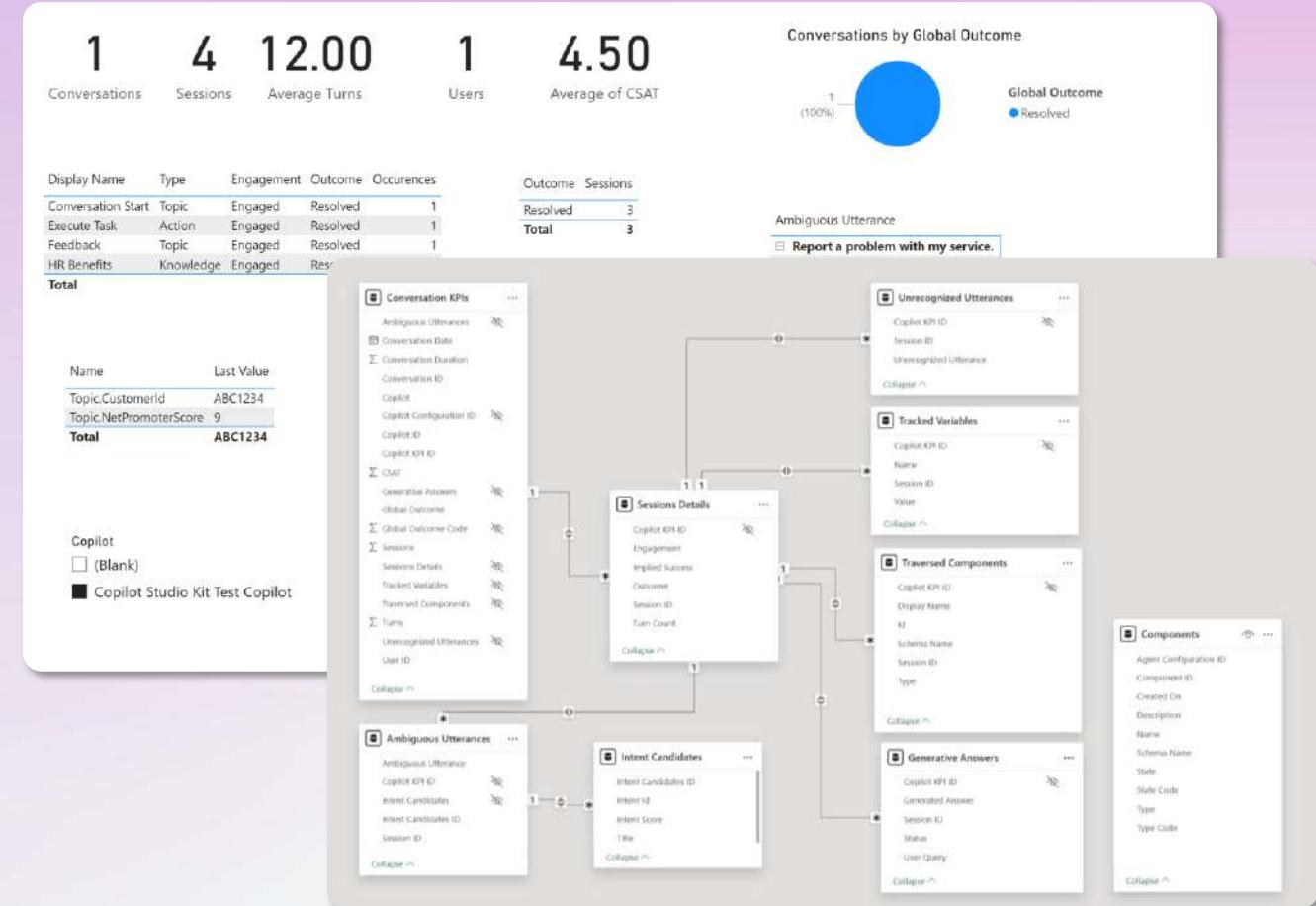
Can you help me with my printer?

2 September at 4:08 pm

To clarify, did you mean:

# Reporting on aggregated conversation KPIs

Don't start from scratch but from simple and easy to extend pre-built Power BI templates.



# COPilot Studio Kit Roadmap

Coming soon

## Aggregated KPIs

---

Surface simplified conversation outcome results (vs analyzing complex conversation transcripts).

Offer out-of-the-box Power BI reports based on these aggregates



Coming soon

## Adaptive Cards Gallery

---

Offer a gallery of pre-configured Adaptive Cards

Allow pro-devs to surface new Adaptive Cards for makers of their organization



Coming soon

## Multi-turn testing capabilities

---

Ability for have multiple back and forth interactions with the custom copilot vs single-turn today.



Coming soon

## Run tests in the context of a deployment pipeline

---

Define rules to validate if specific thresholds must be met before deploying a custom copilot to a downstream environment.



THANK YOU



The Power CAT team helps scale Copilot Studio knowledge gather from hundreds of customer engagements and build tools, for everyone

#### Guide

**Copilot Studio Implementation Guide**  
[aka.ms/CopilotStudioImplementationGuide](https://aka.ms/CopilotStudioImplementationGuide)

#### Documentation

**Copilot Studio Guidance**  
[aka.ms/CopilotStudioGuidance](https://aka.ms/CopilotStudioGuidance)

#### Tools

**Copilot Studio Kit**  
[aka.ms/CopilotStudioKit](https://aka.ms/CopilotStudioKit)

#### Training

**Copilot Studio Workshop**  
[aka.ms/CopilotStudioWorkshop](https://aka.ms/CopilotStudioWorkshop)

#### Samples

**Copilot Studio Code Samples**  
[aka.ms/CopilotStudioSamples](https://aka.ms/CopilotStudioSamples)