

Automate agent evaluation using the Power Platform API.

Copilot Studio provides users with tools to continuously evaluate agent performance by running automated tests based on predefined test suites using the Power Platform REST API.

Copilot Studio provides users with tools to continuously evaluate agent performance by running automated tests based on predefined test suites using the Power Platform REST API. Using the REST API, you can programmatically enable agent evaluation as part of the development process, such as during agent updates, release validation, or regression testing.

Automating the assessment process helps you:

1. Verify agent quality after making changes.
2. Run periodic performance tests on agents in production or staging environments.
3. Integrate testing agents into CI/CD pipelines.
4. Detect regression errors in agent behavior early in the development lifecycle.

Prerequisites

1. You have the Bot ID and Environment ID for the target agent.
2. A test suite is created in Copilot Studio for your target agent.
3. A user access token issued by Microsoft Entra ID (OAuth 2.0).
 1. Obtain the access token using the client ID of the application subscription with the appropriate scope provided by the Power Platform API.

Overview of running assessments using REST API

To run an assessment using the Power Platform API, follow these general steps:

1. Fulfill the prerequisites.
2. Find and retrieve the ID of the test suite you want to use.
3. Run the evaluation.
4. Retrieve results using the evaluation run ID.

When the request is successful, the evaluation process will run asynchronously and produce results that you can review in Copilot Studio.

API operations for automating assessments

Copilot Studio supports REST API operations that you can use to trigger programmatic evaluation of your agent using an existing test toolkit.

How to obtain the test agent toolkit

1. **Endpoint** :GET
`https://api.powerplatform.com/copilotstudio/environments/{EnvironmentId}/botversion=2024-10-01`
2. **Purpose** : To retrieve an array of test kit IDs and other details for a specific agent.
3. **Response** : Returns a list of `value`test toolkit names with the following information:
 1. `auditInfoTimestamps` and user IDs are used to create and modify each test kit.
 2. `displayName`: Name of the test kit.
 3. `id`: ID of the test kit.
 4. `description`Description of the testing toolset.
 5. `state`: The state of the test suite. An available test suite has the state `Active`.
 6. `totalTestCases`The number of test cases in the test toolkit.

Get details of the test agent toolkit.

1. **Endpoint** :GET
`https://api.powerplatform.com/copilotstudio/environments/{EnvironmentId}/botversion=2024-10-01`
2. **Purpose** : To retrieve details for a specific test suite, using the test suite ID.
3. **Response** : Returns information about an item in the response array.

Start evaluating agents

1. **Endpoint**: POST
`https://api.powerplatform.com/copilotstudio/environments/{EnvironmentId}/botversion=2024-10-01`
2. **Purpose** : To run an evaluation for a test suite using its test suite ID. You can also include a user profile to authenticate the connection during the evaluation run. Use this `mcsConnectionId`to specify the user profile. If you don't add it `mcsConnectionId`to your call, the evaluation will run without authentication.
3. **Response** : Returns the following information:
 1. `runId`: The ID of the evaluation run. Use this ID to retrieve evaluation details.
 2. `lastUpdatedAt`: The time the status was last updated.
 3. `executionState`: The status of the run while the evaluation process is underway.
 4. `state`: Current state of the run.
 5. `totalTestCases`: The total number of test cases in the test suite used for evaluation.
 6. `testCasesProcessed`: Total number of test cases evaluated up to the last update.

Retrieve information on the agent's test runs.

1. **Endpoint** :GET
`https://api.powerplatform.com/copilotstudio/environments/{EnvironmentId}/botversion=2024-10-01`

2. **Purpose** : To retrieve an array of all previous runs.
3. **Response** : Each item in the array contains the same values as those found in Get agent test run details.

Get details about the agent's test run.

1. **Endpoint** :GET
`https://api.powerplatform.com/copilotstudio/environments/{EnvironmentId}/botversion=2024-10-01`
2. **Purpose** : To retrieve detailed information from a single assessment for use runId in your target assessment run.
3. **Response** : Returns the following information:
 1. **id**: The ID of the evaluation run. Use this ID to retrieve evaluation details.
 2. **environmentId**: Your agent environment ID.
 3. **cdsBotId**: Target agent ID.
 4. **ownerId**: The user ID has started the evaluation process.
 5. **testSetId**: ID of the test kit used for evaluation.
 6. **state**: Status of the evaluation process.
 7. **startTime**: Time when the assessment begins.
 8. **endTime**: The end date of the evaluation, if applicable.
 9. **name**: Name of the evaluation process.
 10. **totalTestCases**: Total number of test cases in the test toolkit.
 11. **mcsConnectionId**: The Copilot Studio connection ID of the user profile used for the assessment process. Returns null if no user profile is connected.
 12. **testCasesResults**: A list of test cases during the evaluation process. Includes:
 1. **testCaseId**: Test case ID.
 2. **state**: The test case is complete.
 3. **metricsResultsDetails** and results for each test method used for the test case. Including:
 1. **typeTesting** methods.
 2. **result**The final test results for this test case. These include:
 1. **dataDetails** of the results. The exact values depend on the testing method. Learn more in the Power Platform API documentation. For a general quality test, the feedback includes the following:
 1. **abstention** Did the agent answer the question?
 2. **relevance** Is the answer relevant or not?
 3. **completeness** Is the answer complete or not?
 3. **status**: The state of the test case.
 4. **errorReason** If an error occurs, what is the cause of the error?
 5. **aiResultReason** AI's explanation of the test case results.

Use Microsoft Studio Connector ID for evaluation.

To begin evaluating the agent, you can optionally add your Microsoft Studio Connector ID to the call as a user profile for the evaluation process. To find `mcsConnectionId` yours:

1. Access Power Automate .
2. Open the **Connections** page .
3. **Select Microsoft Copilot Studio** connection .

4. Copy mcsConnectionId from URL: .

`./connections/shared_microsoftcopilotstudio/{mcsConnectionId}/details`

You finished reading the article "**Automate agent evaluation using the Power Platform API.**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.
