RAG Innovator Lab: User's Guide

Powerful AI Apps

on Open and Lake-centric Data Foundation

Contents

Overview

Agenda

No	Task	Duration	Description
1	Microsoft Fabric Overview	60 minutes	
2	Microsoft Fabric Hands-On Lab	60 minutes	
3	Azure AI Foundry Overview	60 minutes	
4	AI Application Development Hands-On Lab	30 minutes	(Optional) Evaluation, +30 minutes
5	Wrap-up and Q&A	30 minutes	

Pre-configured User Settings

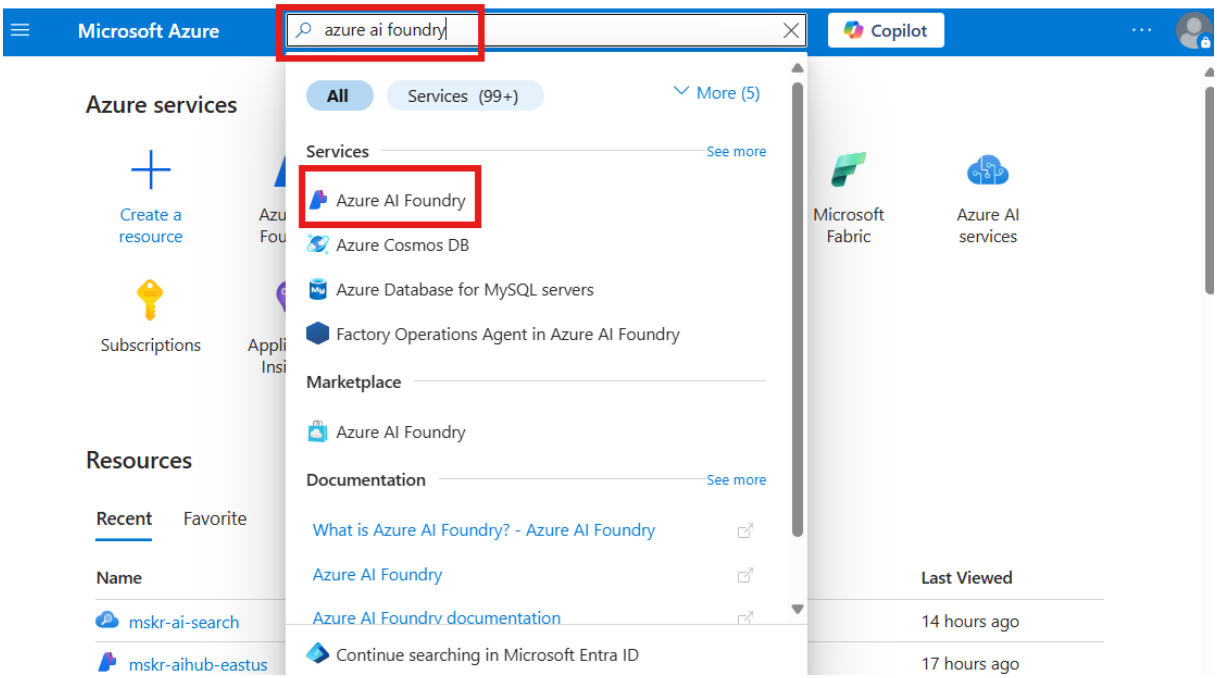
Assignee	Fabric Workspace	Azure OpenAI	AI Search	AI Foundry AI Hub	AI Foundry Compute	Custom Suffix
참석자1	__WS_Lab_1	lab-aoai-eus-1	lab-search-1	lab-hub-1	lab-compute-1	user1
참석자2	__WS_Lab_1	lab-aoai-eus-1	lab-search-1	lab-hub-1	lab-compute-1	user2
참석자3	__WS_Lab_2	lab-aoai-eus-2	lab-search-2	lab-hub-2	lab-compute-2	user3
참석자4	…					

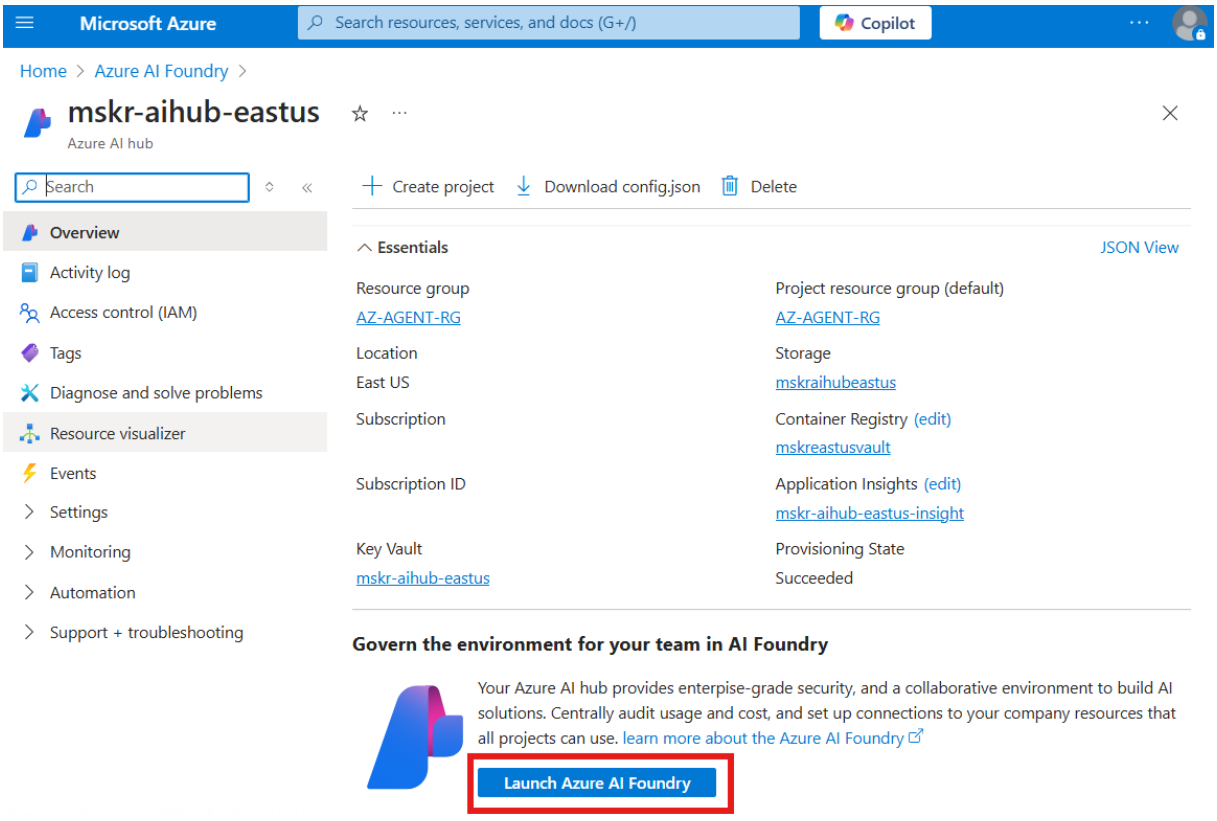
* Fabric, Azure 리소스는 참석자 간 공유 가능하지만, Custom Suffix는 참석자 간 unique해야 합니다.

Prerequisites: Prepare AI Foundry Access and Setup

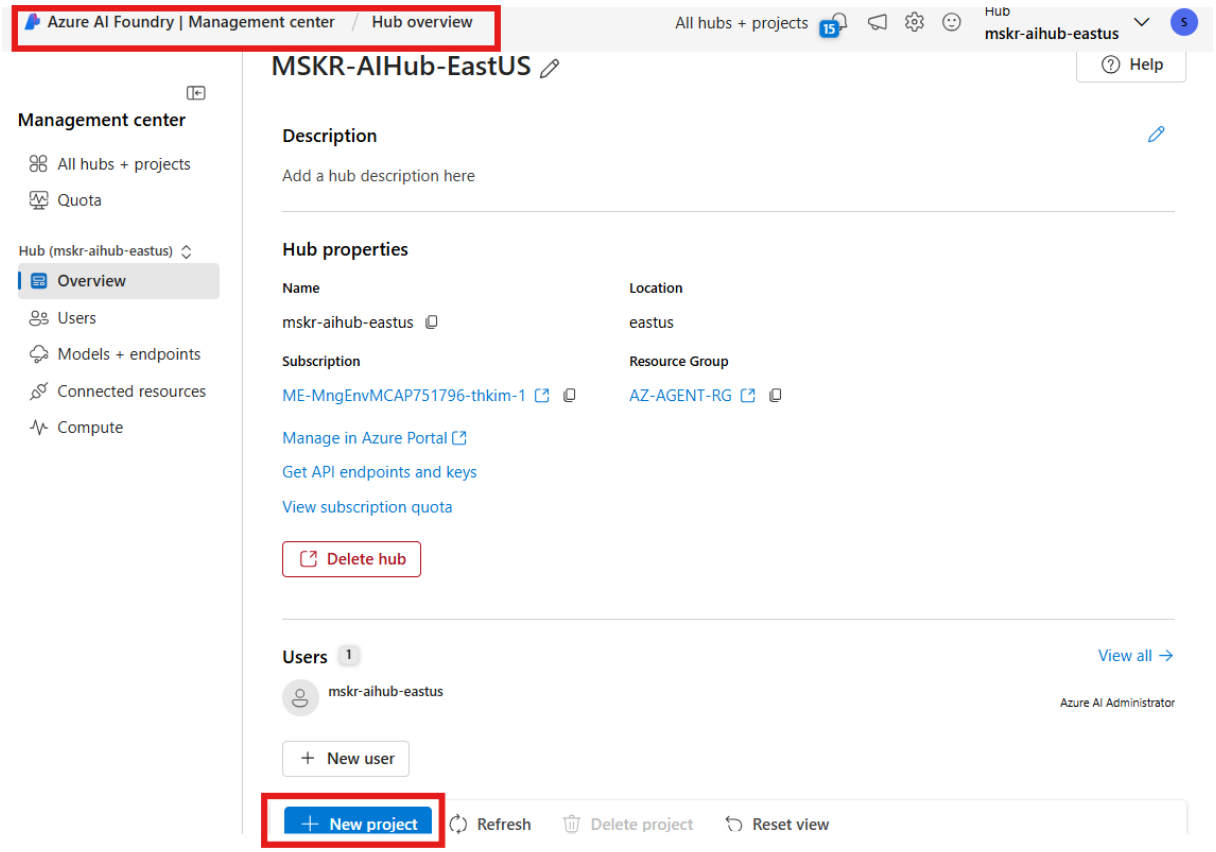
Provision AI Project

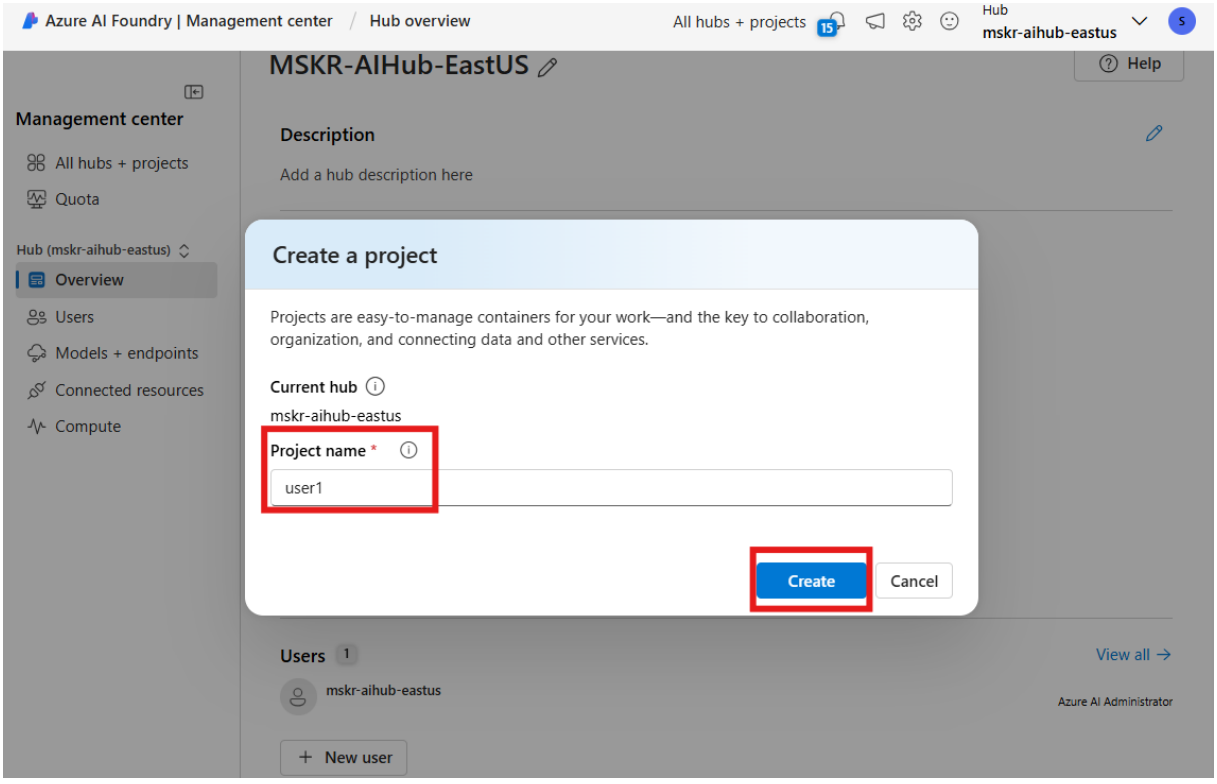
1. Azure Portal에 접속해 배정받은 AI Hub 리소스를 선택합니다.

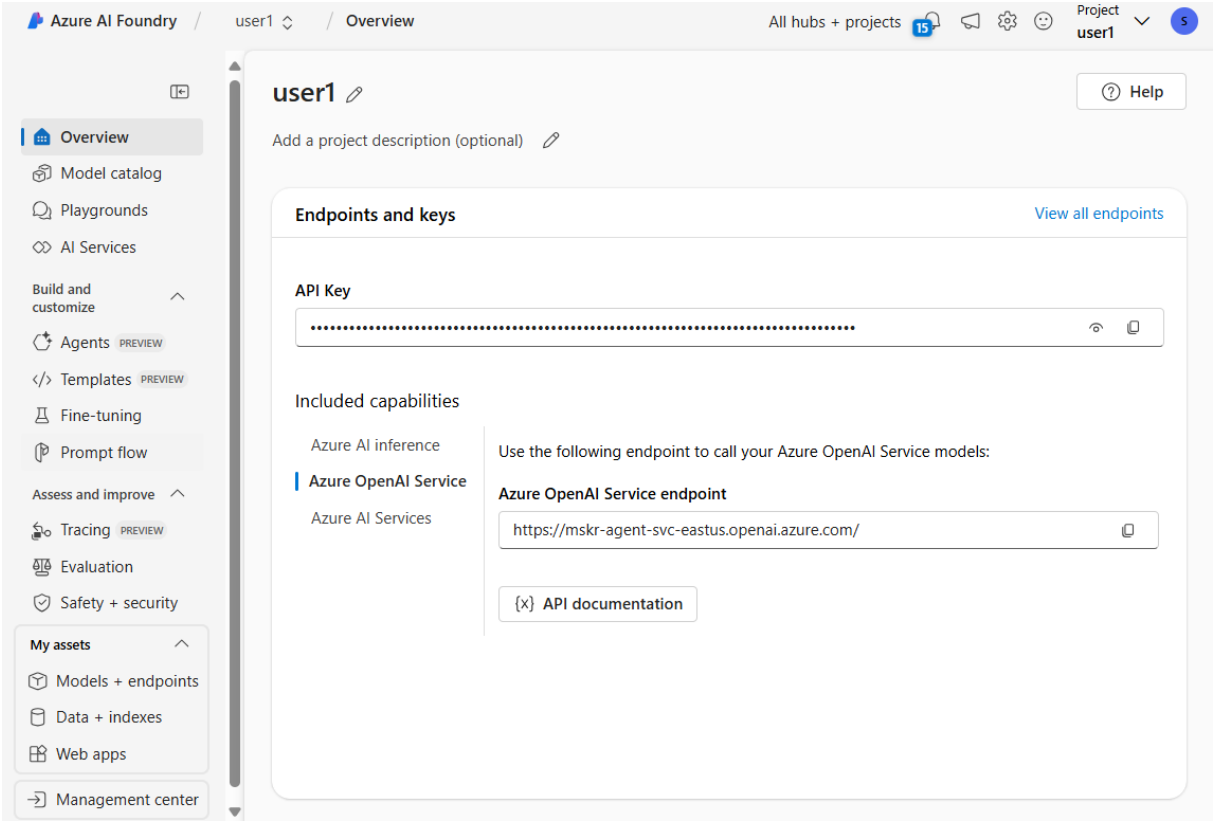




2. Custom Suffix의 이름으로 새로운 Project를 만듭니다.



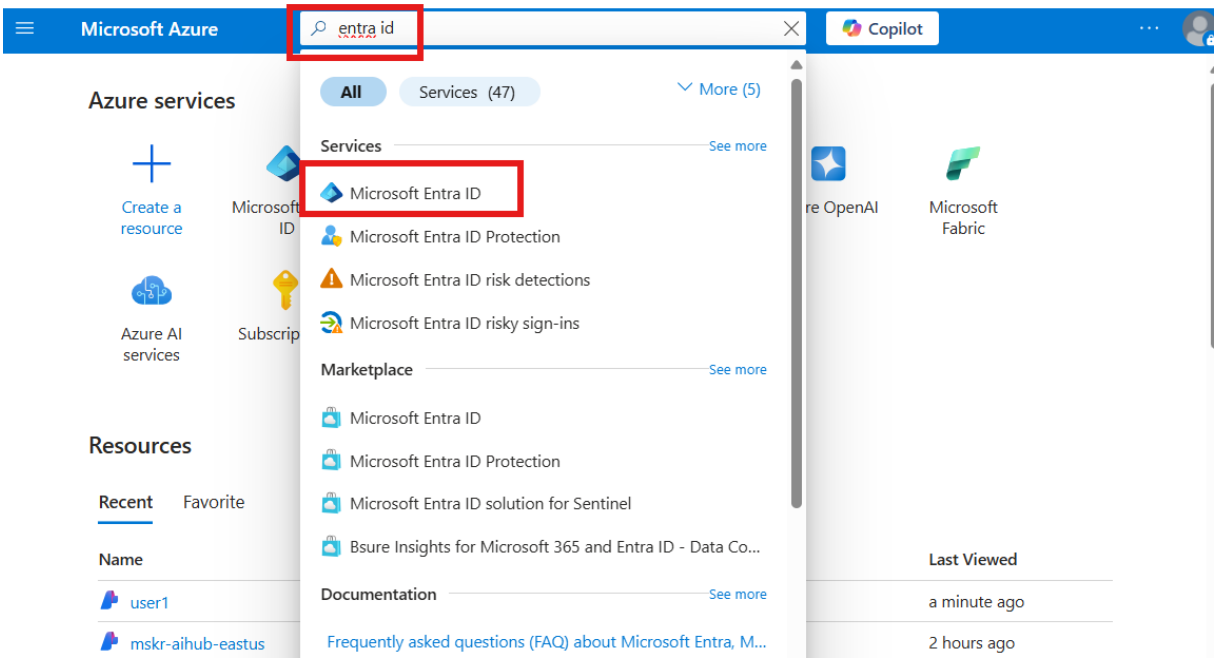




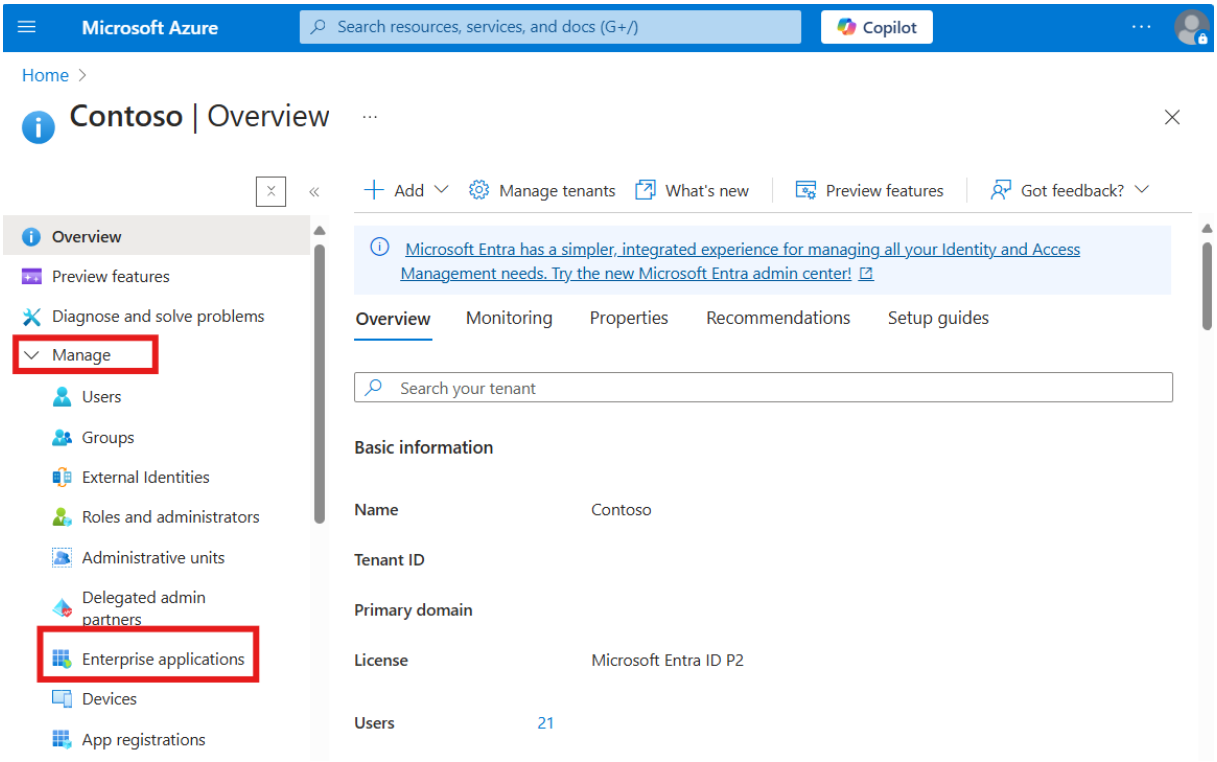
Assign Roles

1. 생성한 AI Project의 Application ID를 확인하고 CLI를 업데이트합니다.

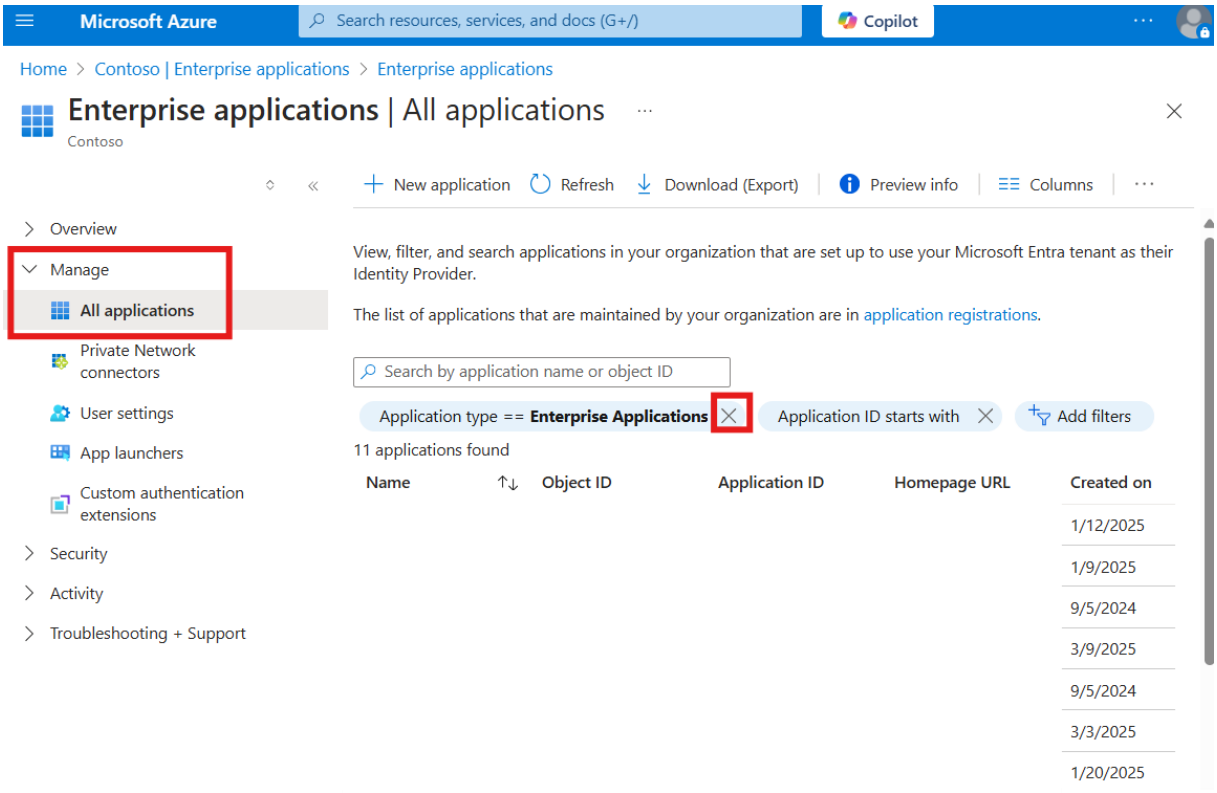
   A. Azure Portal에서 Entra ID 서비스로 이동합니다.



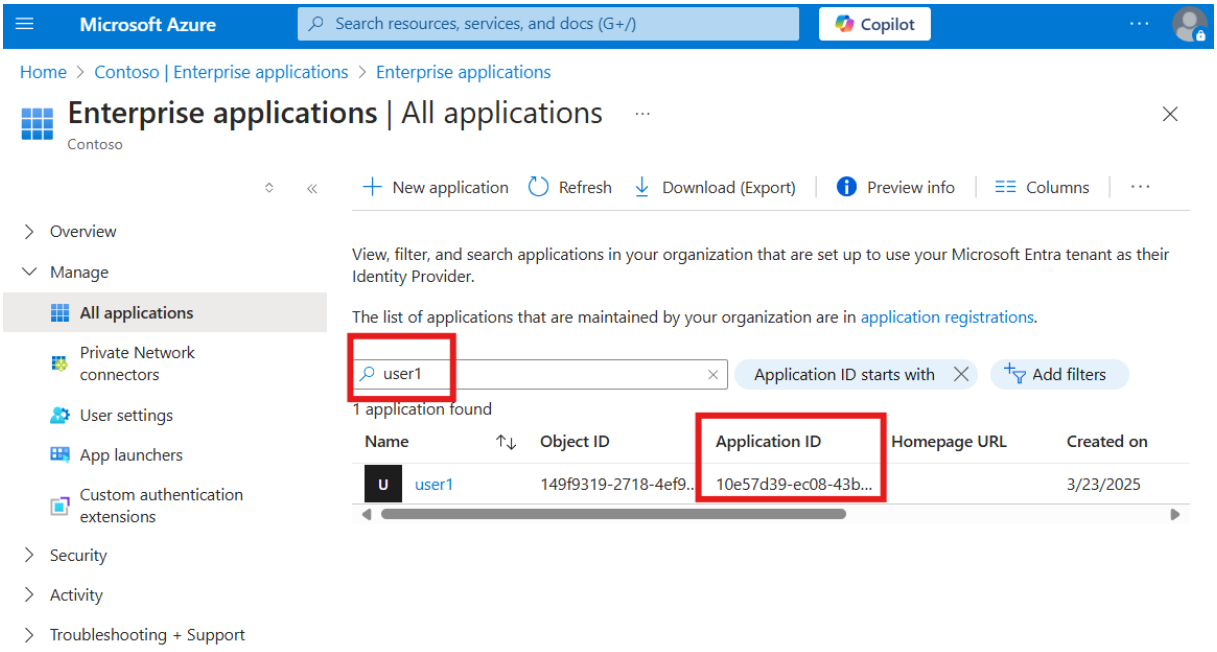
B. Enterprise Applications를 선택합니다.



C. All Applications를 선택하고, Application Type 필터를 해제합니다.



D. 생성한 AI Project의 이름을 검색하고, Application ID를 복사합니다.



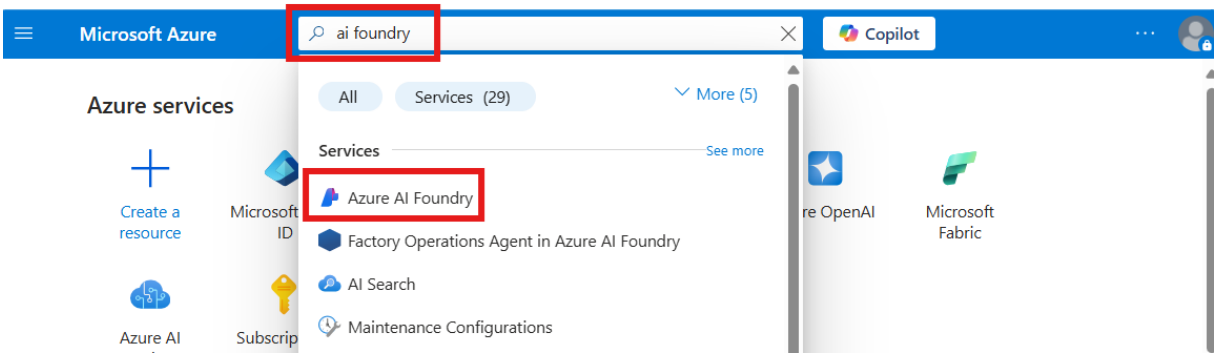
E. 클론된 폴더의 AIFoundry\RBAC_CLI 내 user-role-assign.azcli의 ai_project_assignee 정보를 업데이트합니다.

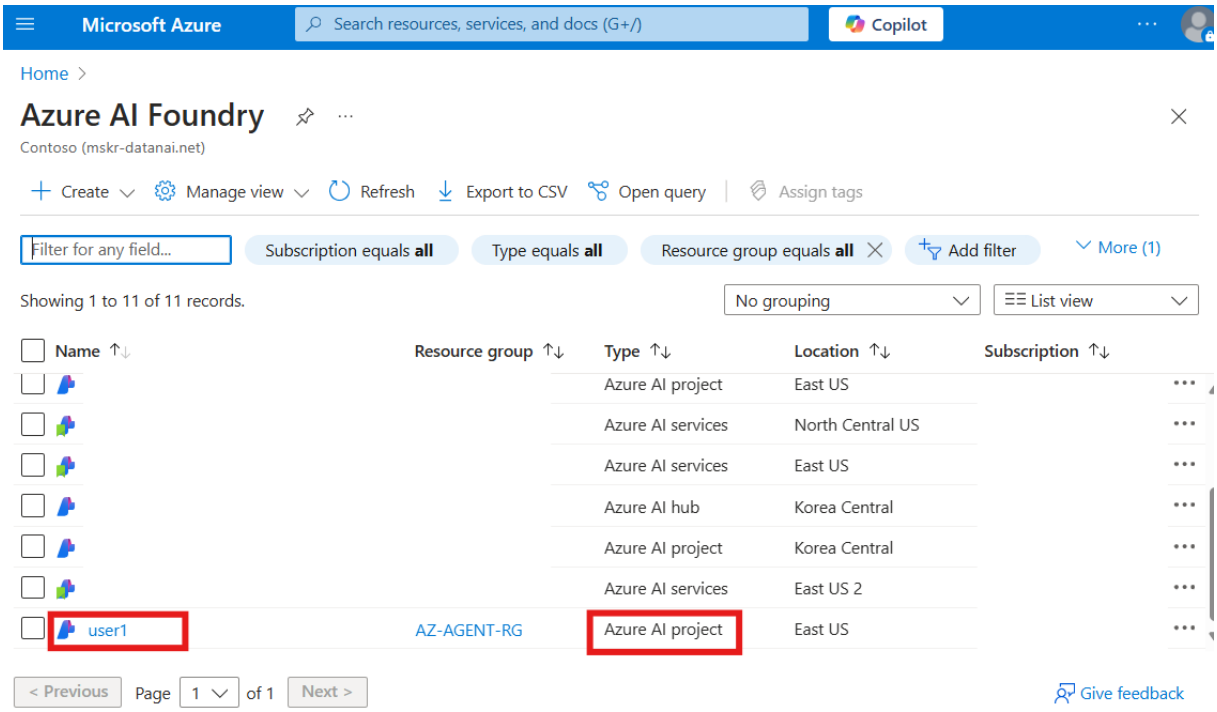
AIFoundry > RBAC_CLI > user-role-assign.azcli

```
# Open terminal in Azure portal after login
# Assign roles for PromptFlow with following commands



$ai_project_assignee = ''
$ai_project_workspace_resource = ''

$storage_account_resource = ''
$acr_assignee = ''
$compute_instance = ''

```

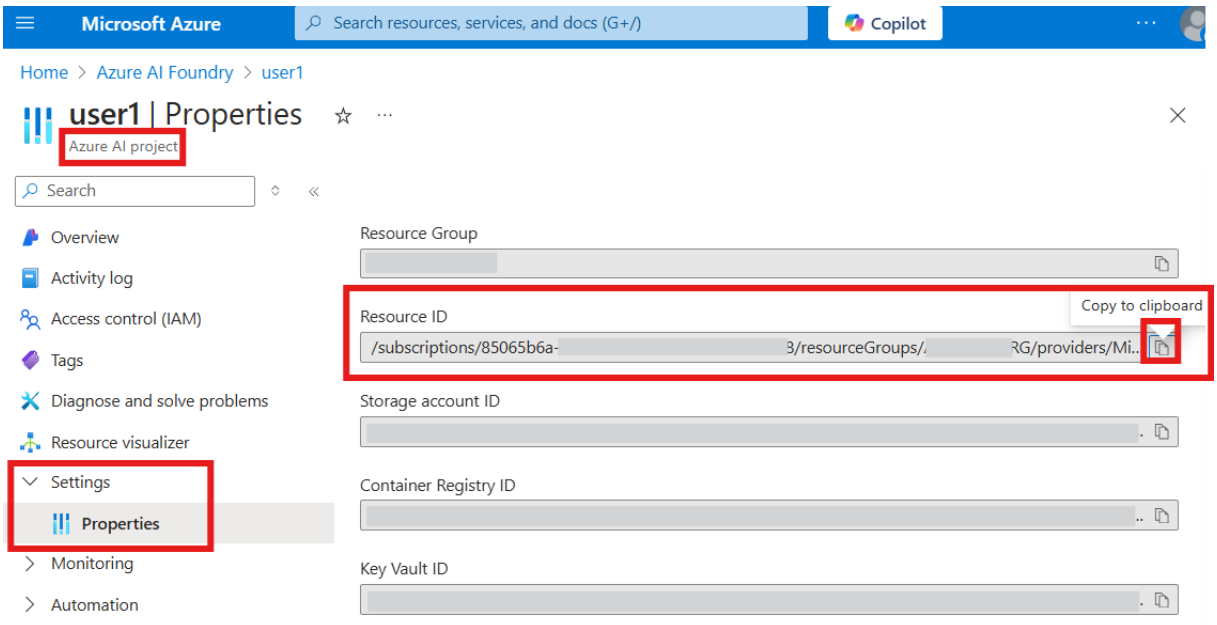
2. 생성한 AI Project의 Workspace ID를 확인하고 CLI를 업데이트합니다.

A. Azure Portal에서 AI Foundry 서비스로 이동한 후, 생성한 AI Project를 선택합니다.





B. AI Project의 Properties 항목에서 Resource ID를 복사합니다.



C. 클론된 폴더의 AIFoundry\RBAC_CLI 내 user-role-assign.azcli의 ai_project_workspace_resource 정보를 업데이트합니다.

AIFoundry > RBAC_CLI > user-role-assign.azcli

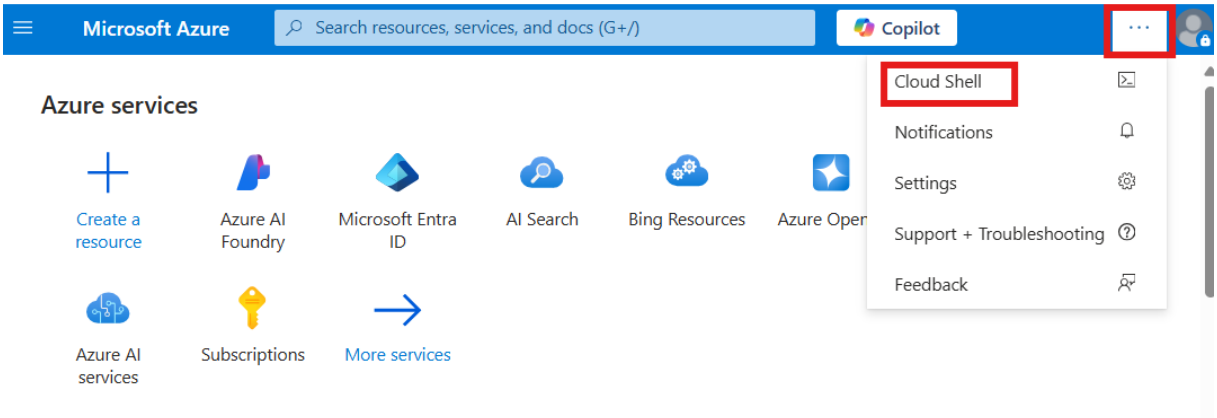
```
# Open terminal in Azure portal after login
# Assign roles for PromptFlow with following commands



$ai_project_assignee = ''
$ai_project_workspace_resource = ''

$storage_account_resource = ''
$acr_assignee = ''
$compute_instance = ''


```

3. 배정받은 리소스의 ID를 확인해 클론된 폴더의 AIFoundry\RBAC_CLI 내 user-role-assign.azcli를 업데이트합니다.

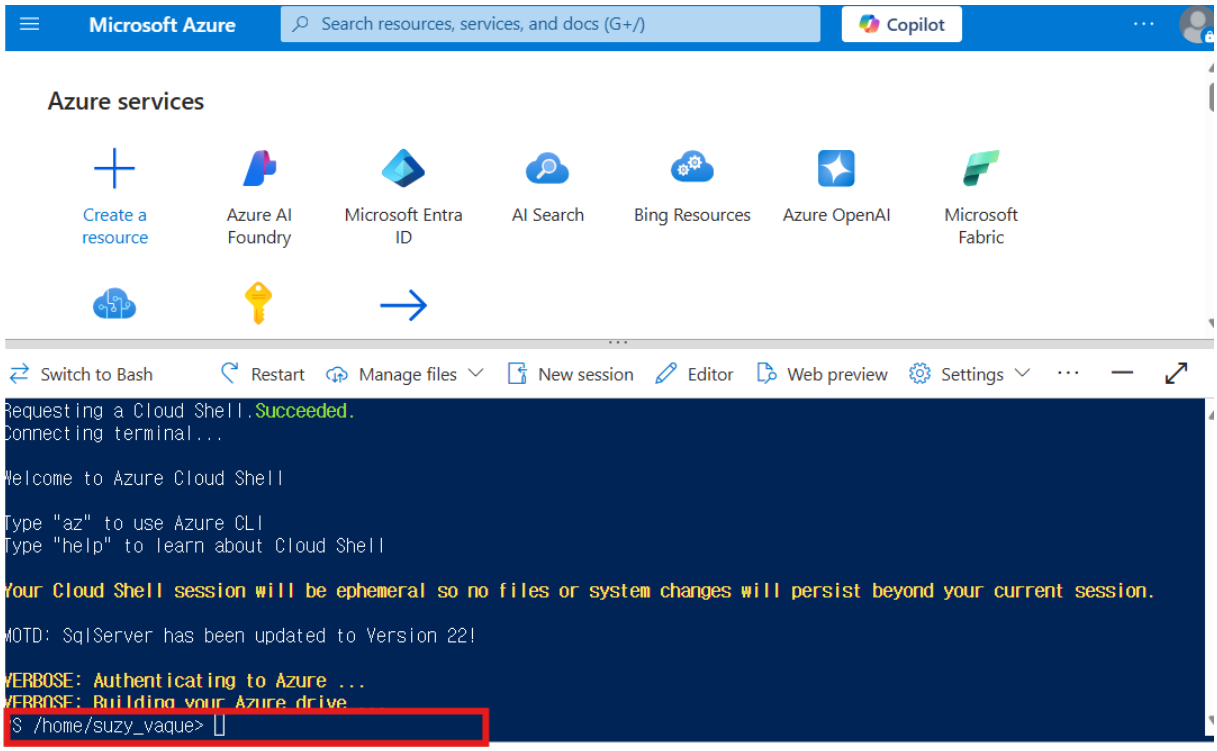
AIFoundry > RBAC_CLI > user-role-assign.azcli

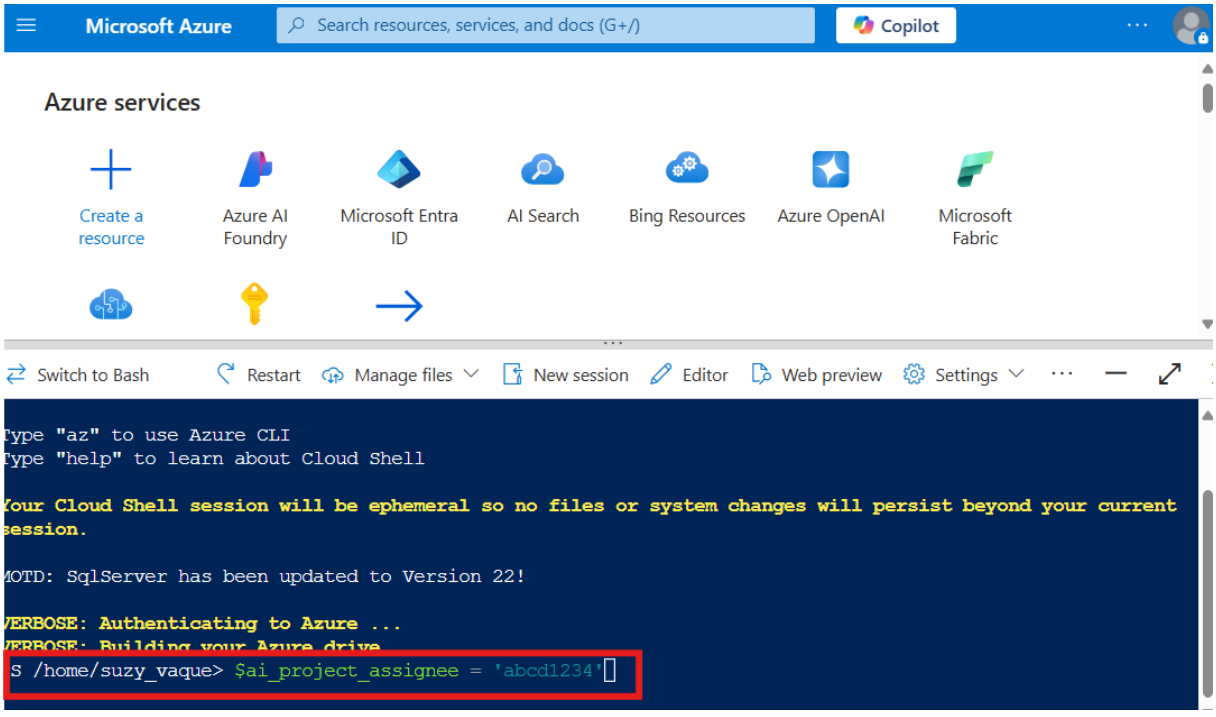
```
# Open terminal in Azure portal after login
# Assign roles for PromptFlow with following commands



$ai_project_assignee = ''
$ai_project_workspace_resource = ''

$storage_account_resource = ''
$acr_assignee = ''
$compute_instance = ''


```

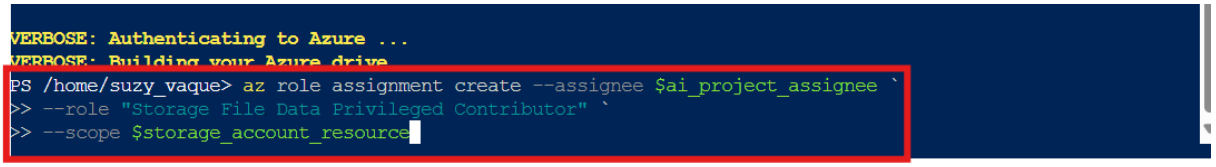
4. Azure Portal에서 Cloud Shell을 실행합니다.



5. 터미널에 user-role-assign.azcli의 각 커맨드를 실행해 Role Assign을 완료합니다.







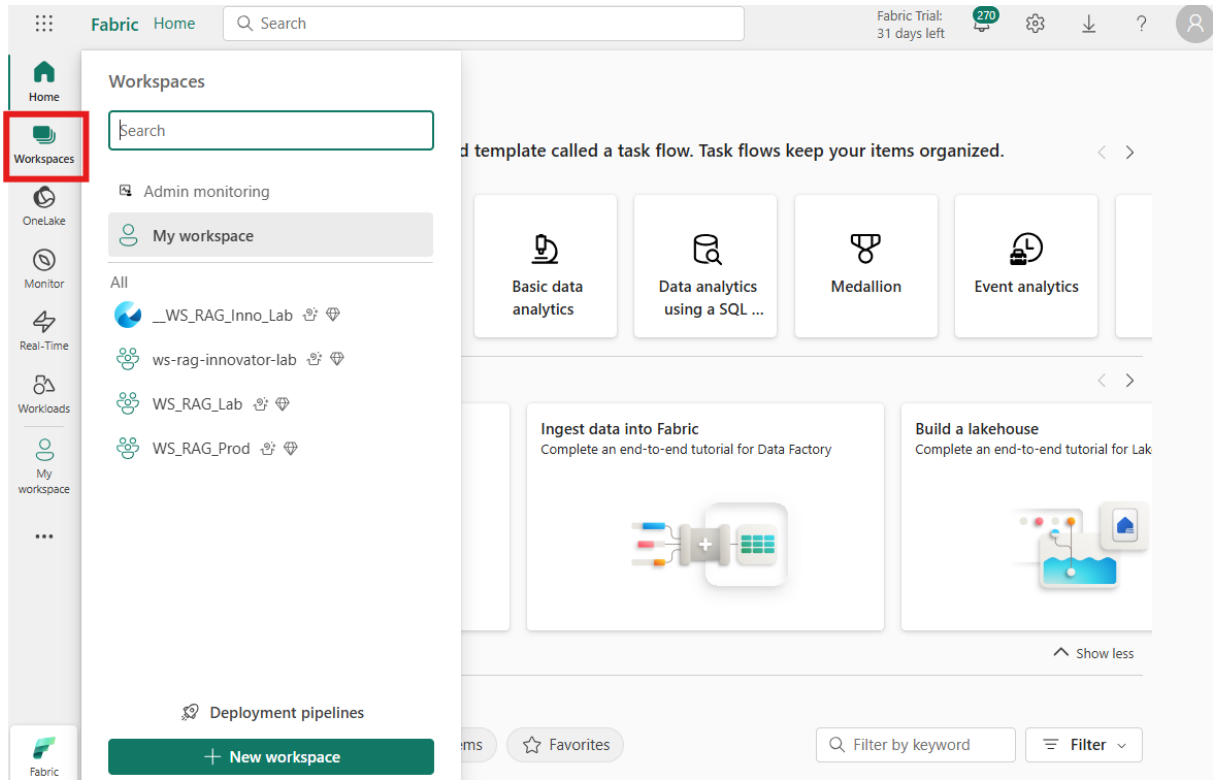
Get Started with Fabric

Import Fabric Notebook

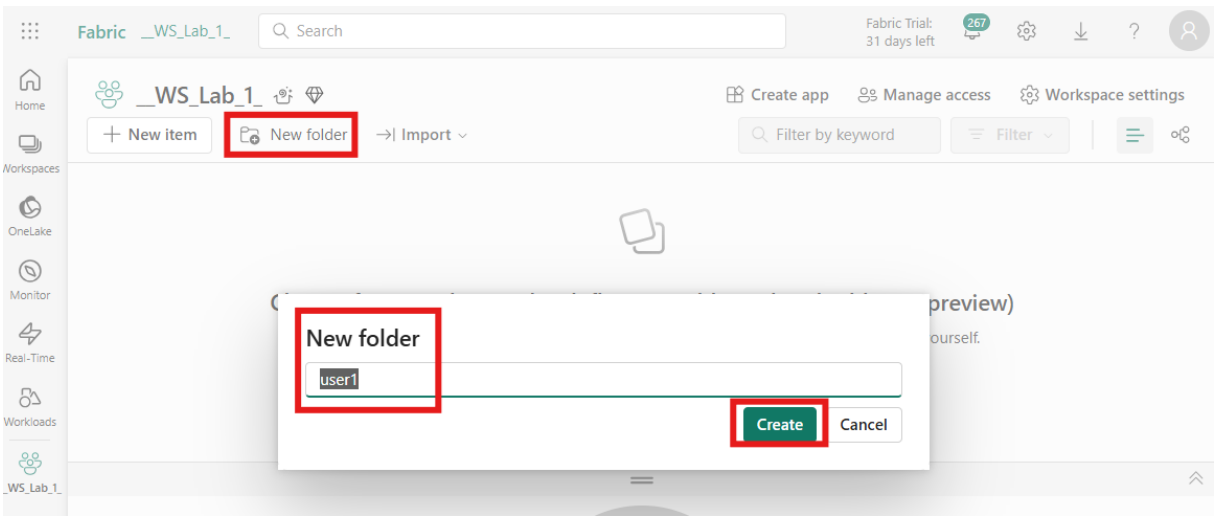
1. GitHub Repository를 로컬로 Clone하기 위해, 로컬에서 터미널을 열고 다음 커맨드를 실행합니다.

```
git clone https://github.com/KRSTUDataAI/rag-innovator-lab.git
```

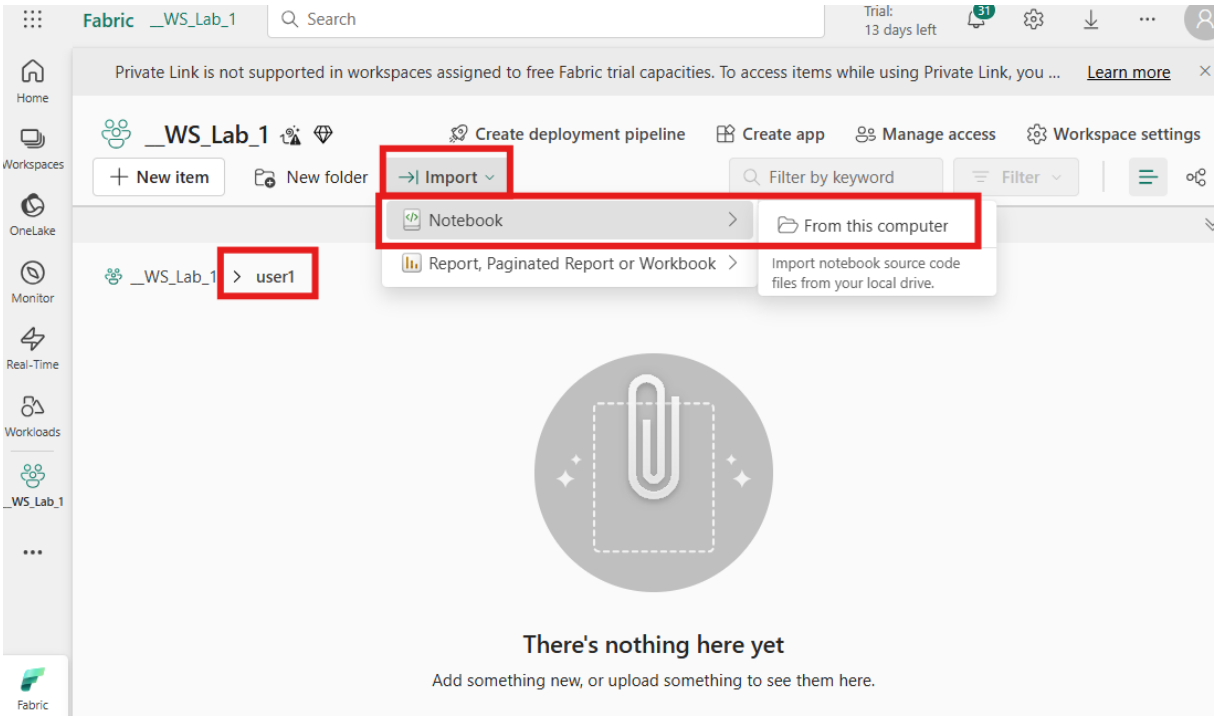
6. Fabric Portal에서 배정받은 Workspace를 선택합니다.



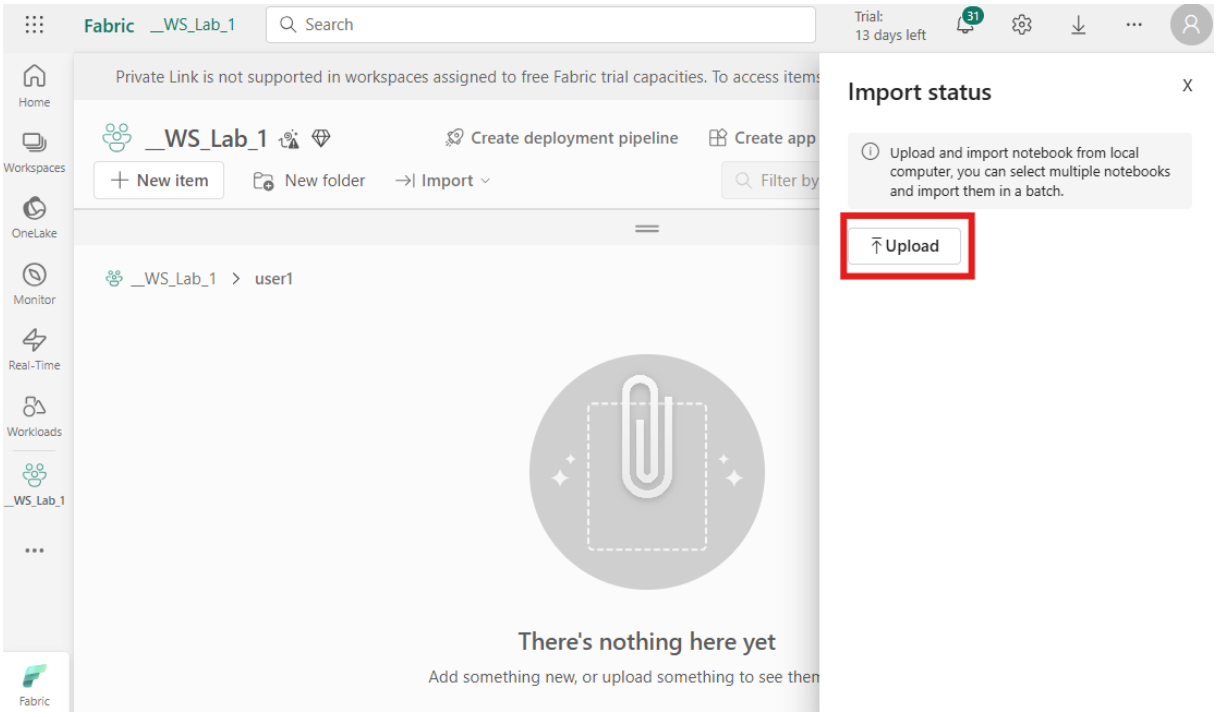
7. 자신의 Custom Suffix 이름으로 폴더를 생성합니다.



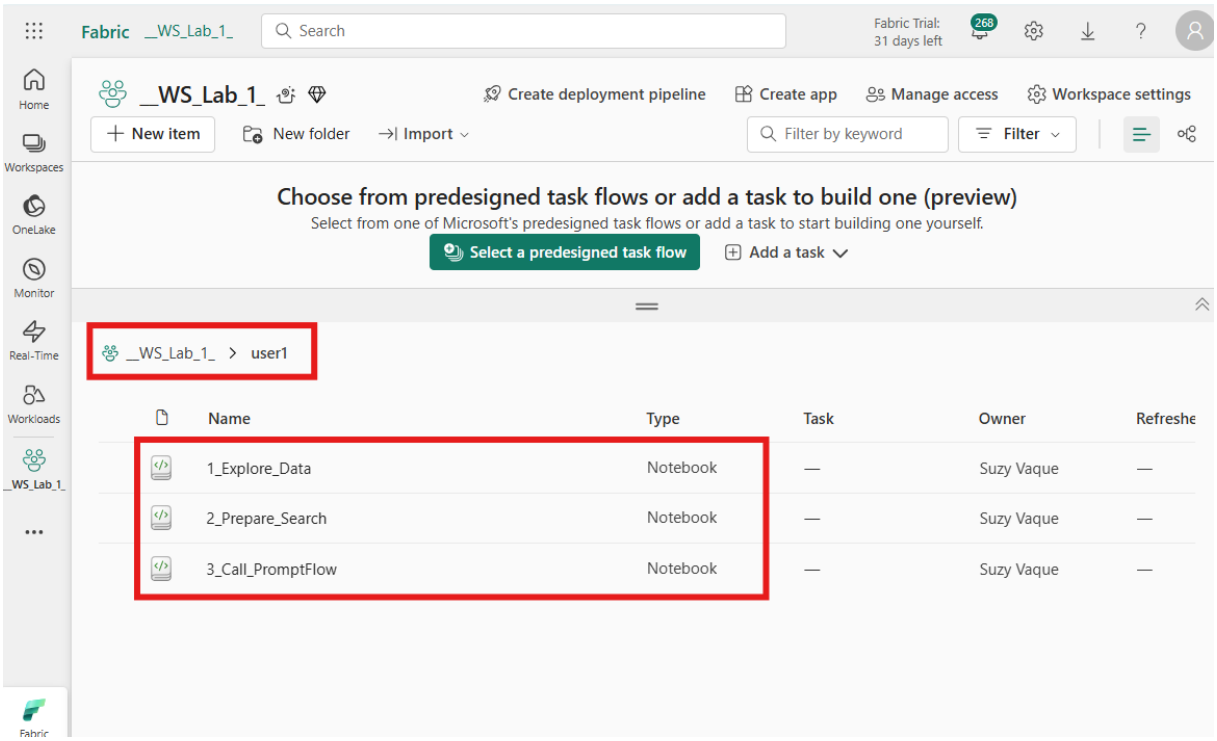
8. 생성한 폴더에 접속한 상태에서 Import Notebook을 진행합니다.



9. 클론된 폴더의 Fabric\Notebooks 내 1_Explore_Data.ipynb, 2_Prepare_Search.ipynb, 3_Call_PromptFlow.ipynb 세 개 노트북을 모두 업로드합니다.

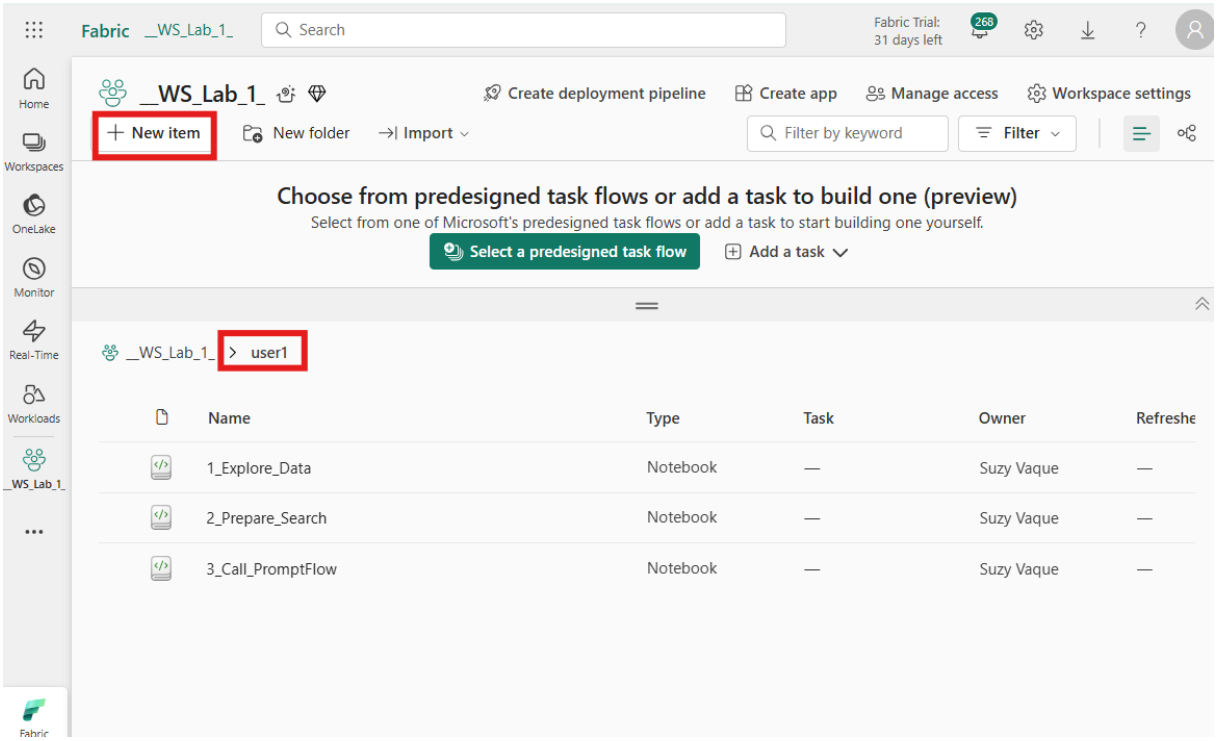


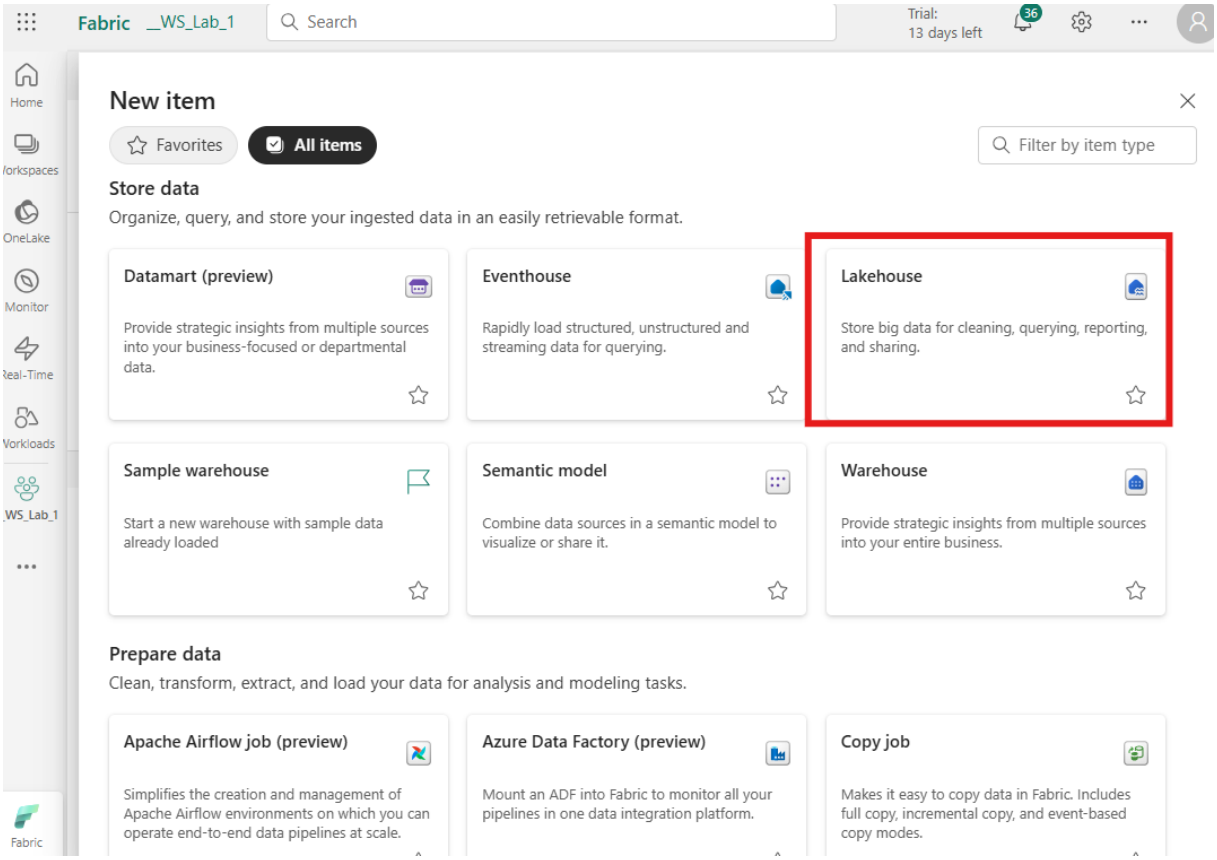
10. Import가 완료되면 Fabric Notebook을 사용할 수 있습니다.

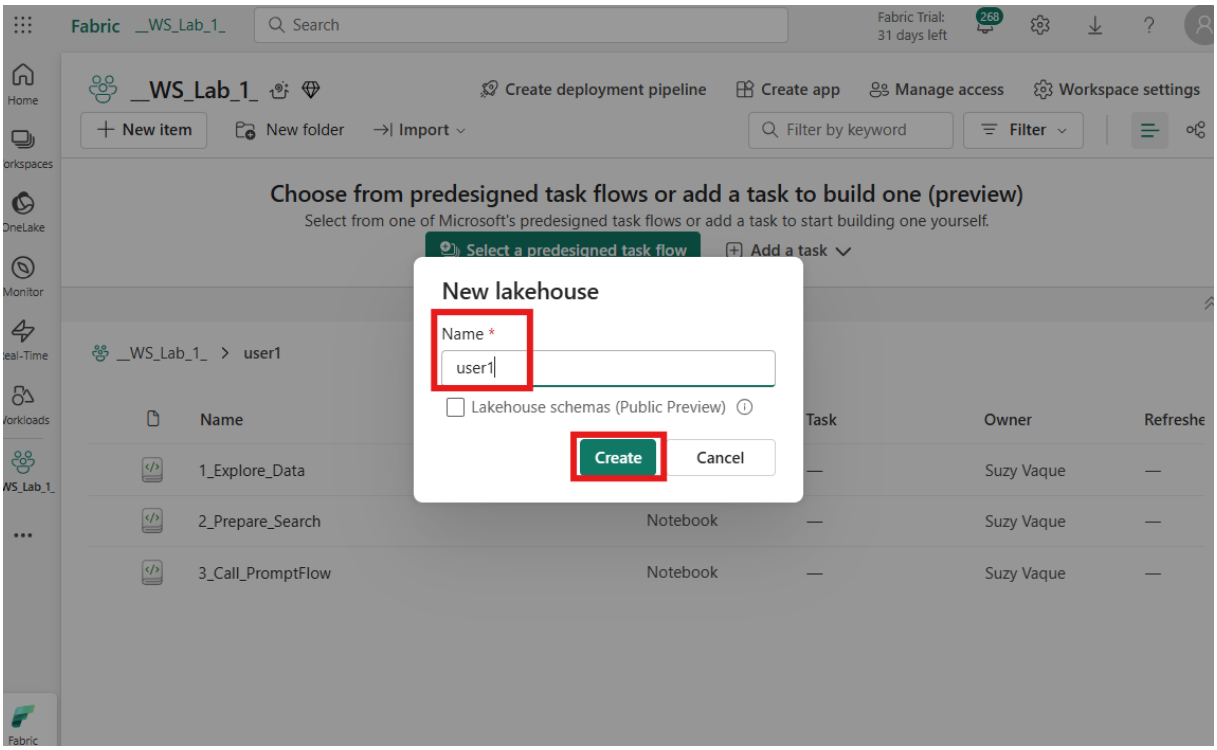


Create Lakehouse

1. 폴더에 들어온 상태에서, New Item을 선택해 자신의 Custom Suffix 이름으로 Lakehouse를 생성합니다.

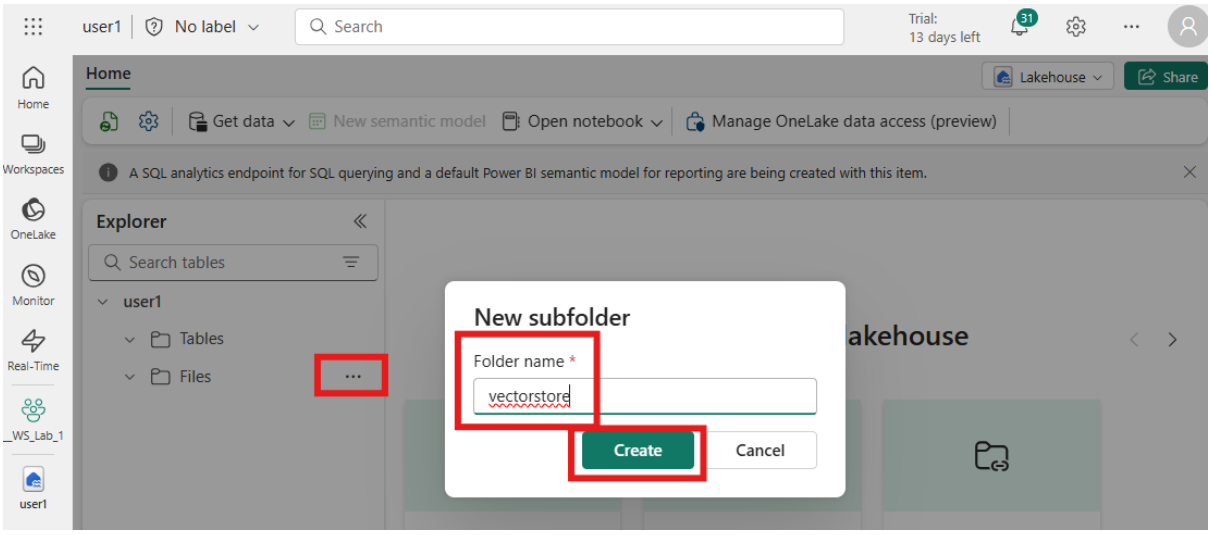




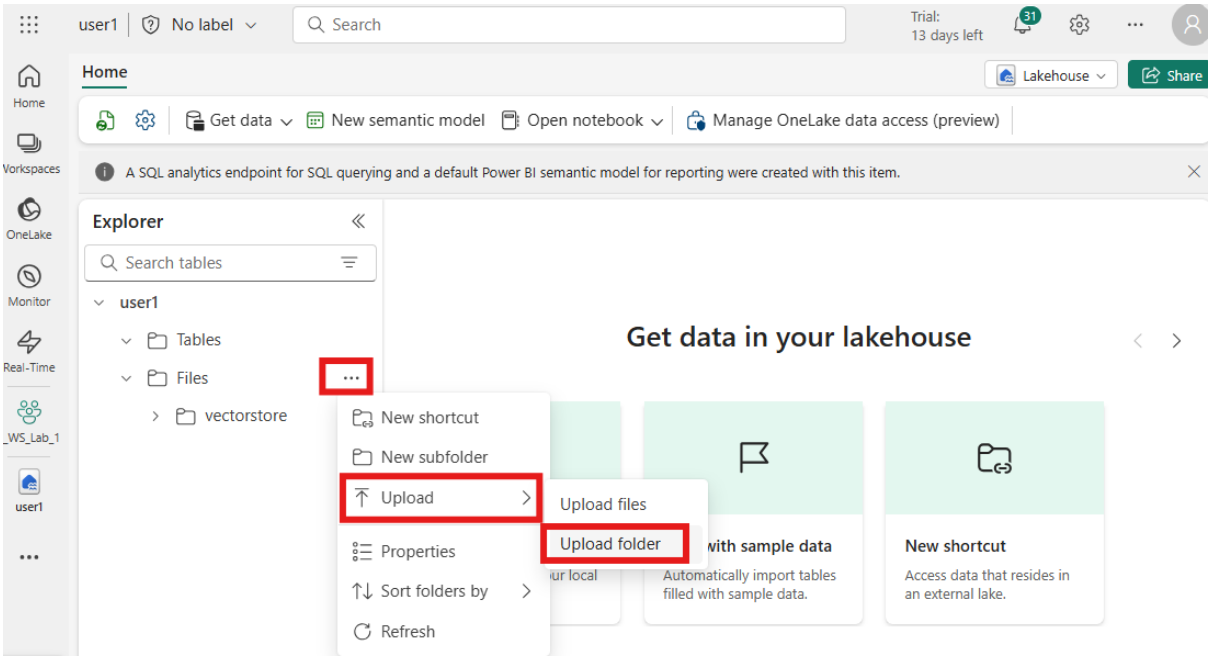


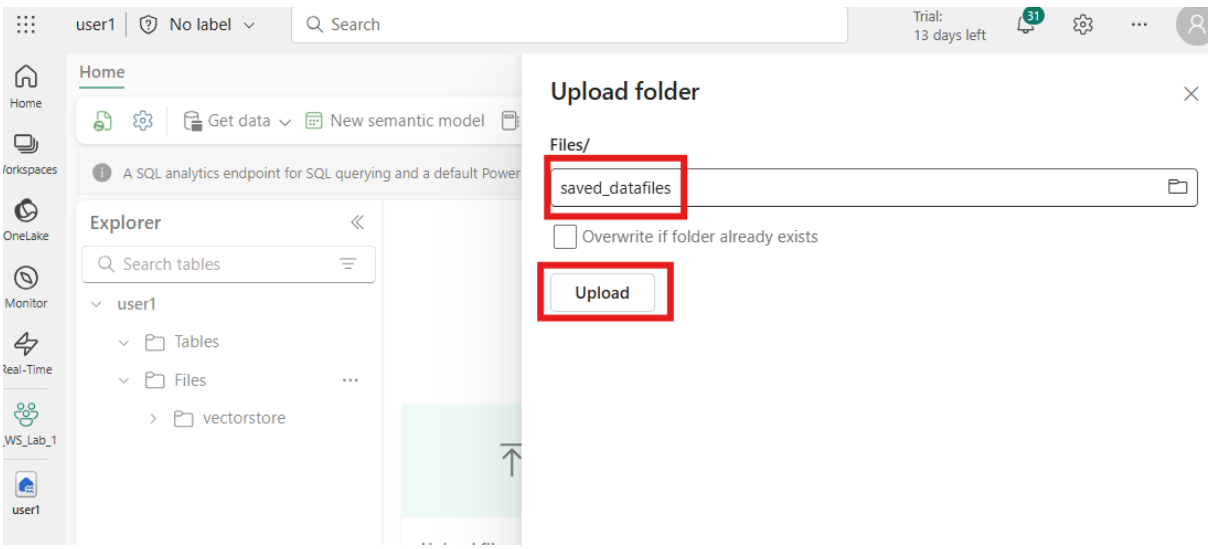
2. Lakehouse 생성이 완료되면 Lakehouse Explorer로 이동합니다.

3. Files에 “vectorstore” 이름의 New Subfolder를 추가합니다.

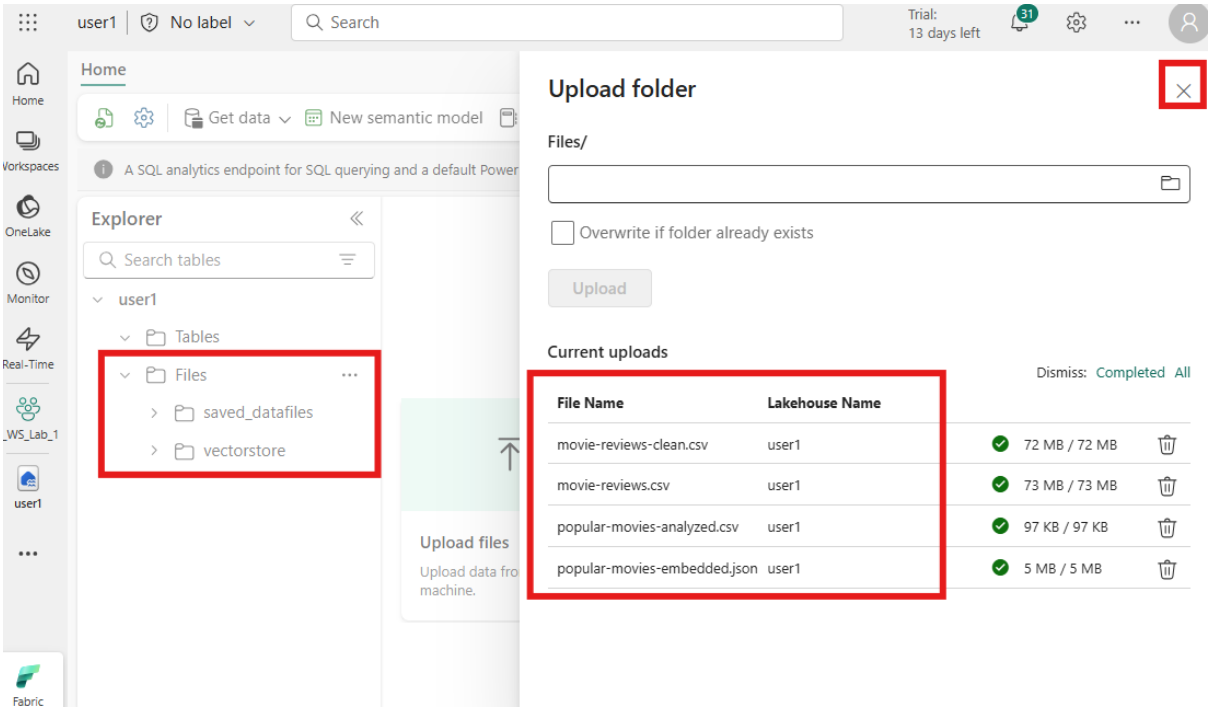


4. Files에 대해 Upload Folder를 선택해, 앞서 클론해둔 레포지터리의 Fabric\Files\saved_datafiles 폴더를 업로드합니다.



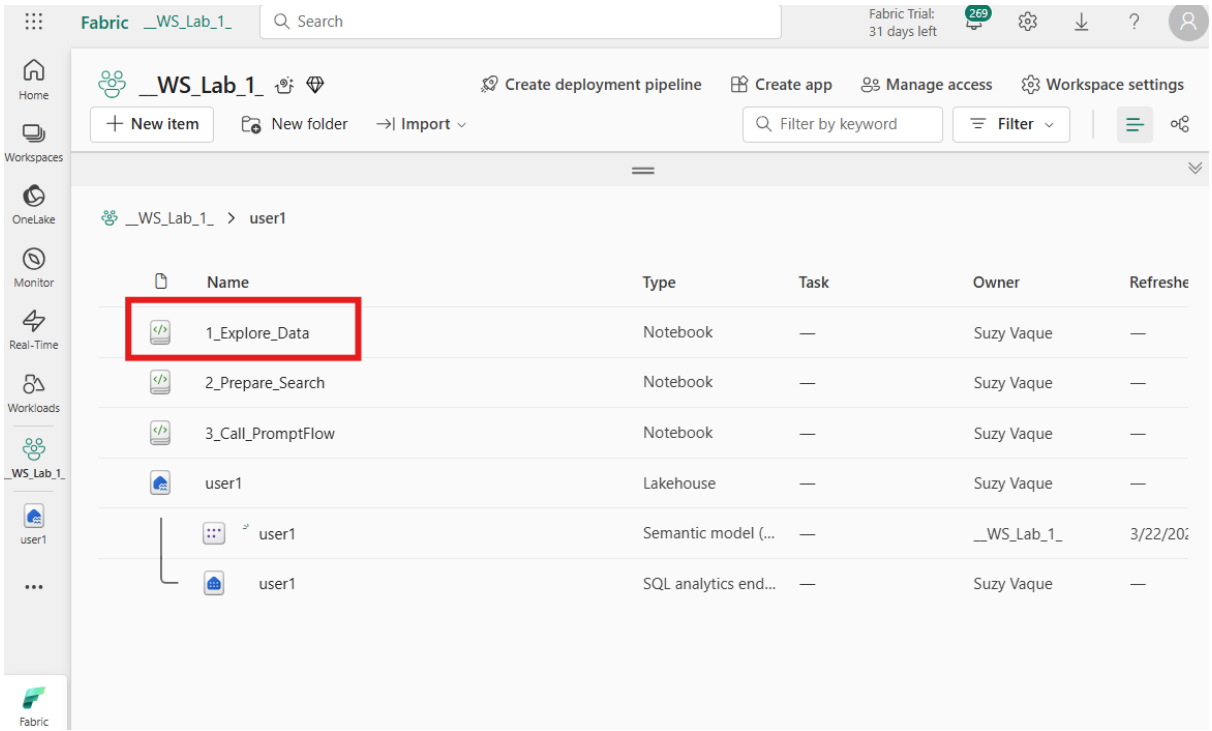


5. 업로드가 완료됐다면 Lakehouse Explorer를 종료합니다.

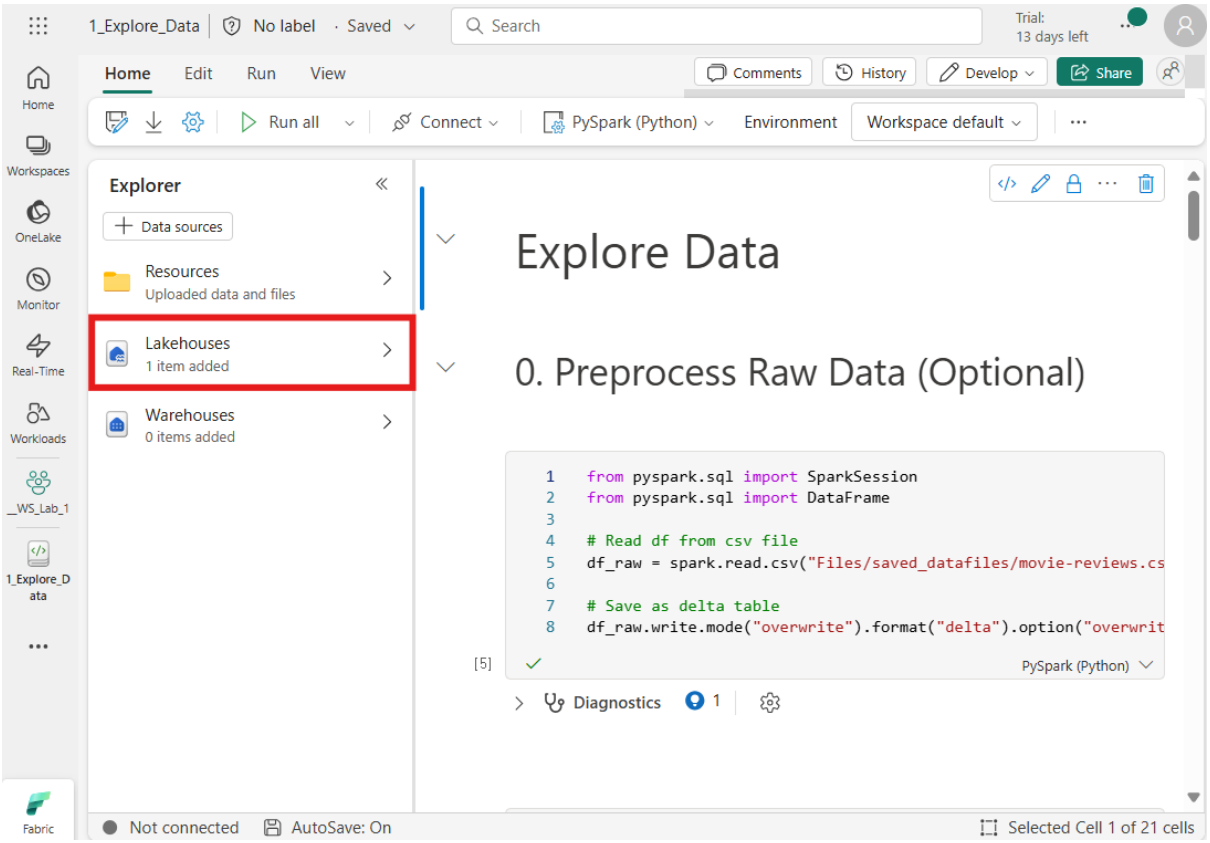


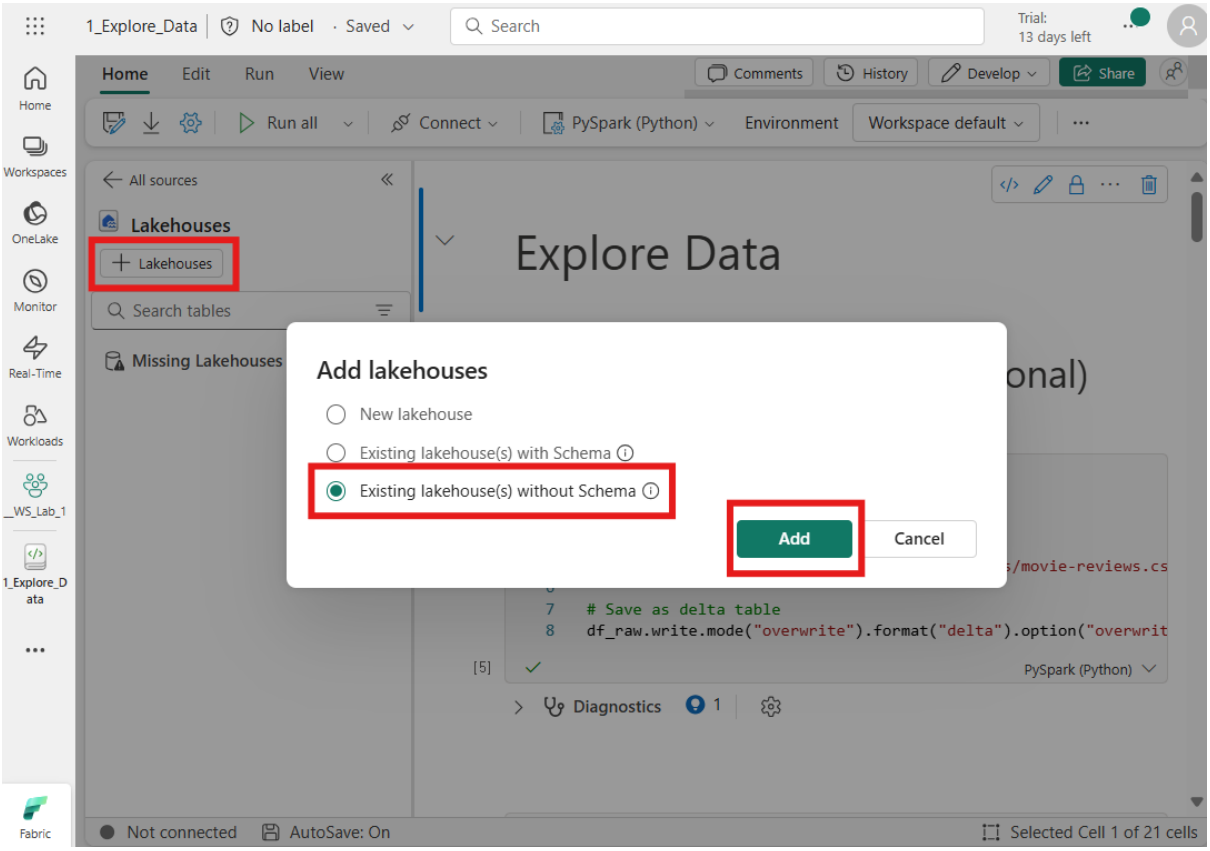
Add Lakehouse to Notebook

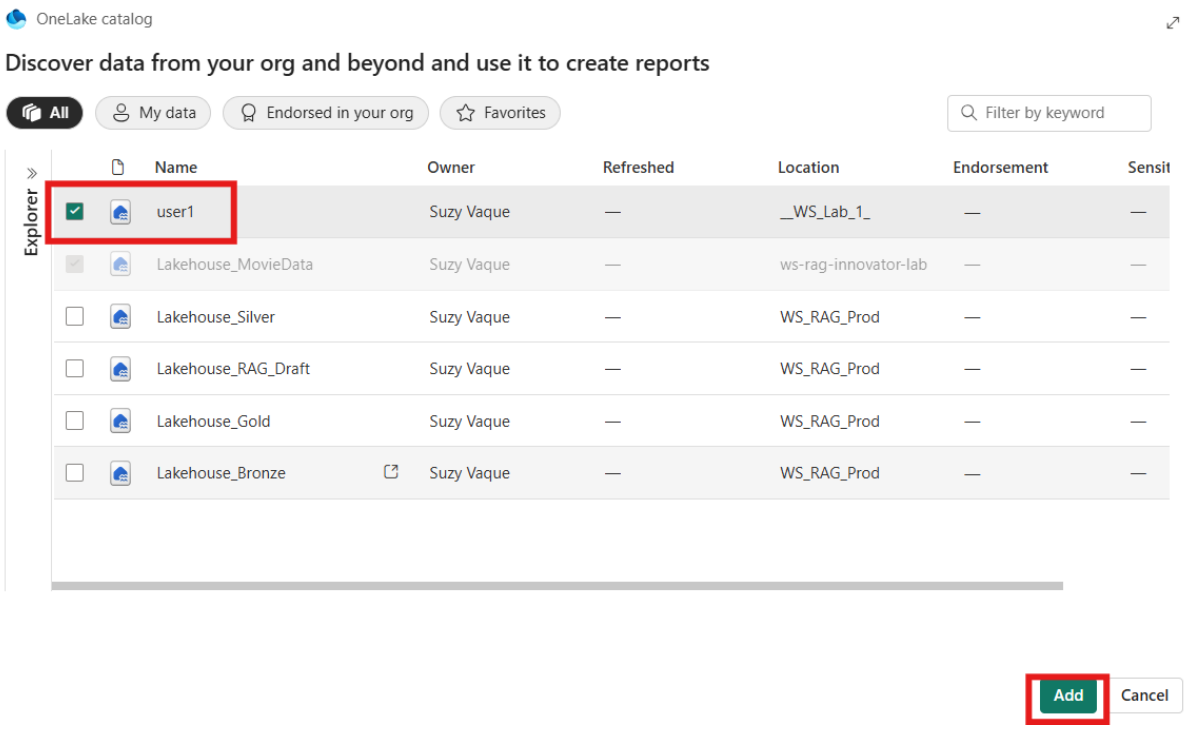
1. Workspace 폴더 내의 Fabric Notebook을 선택합니다.



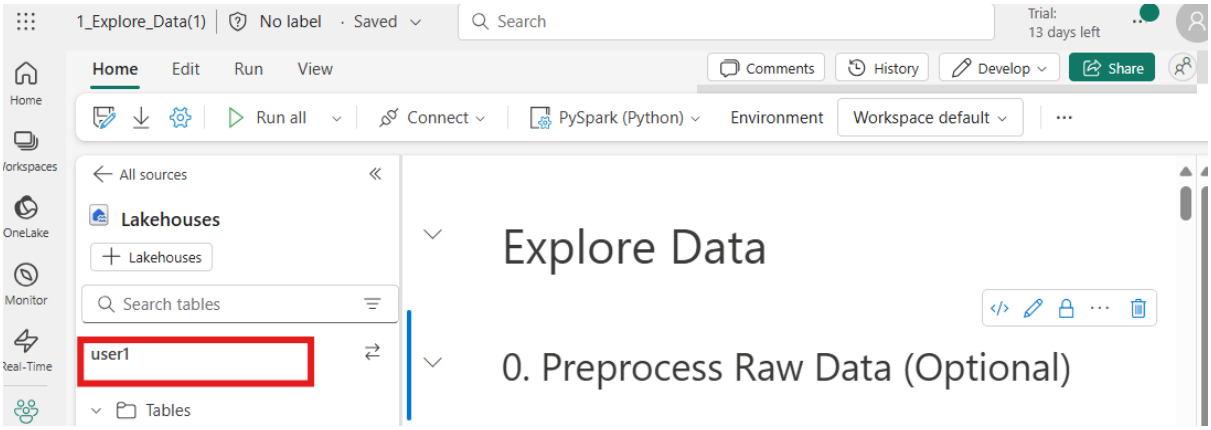
2. Lakehouses 항목을 선택해 앞서 생성한 Lakehouse를 추가합니다.

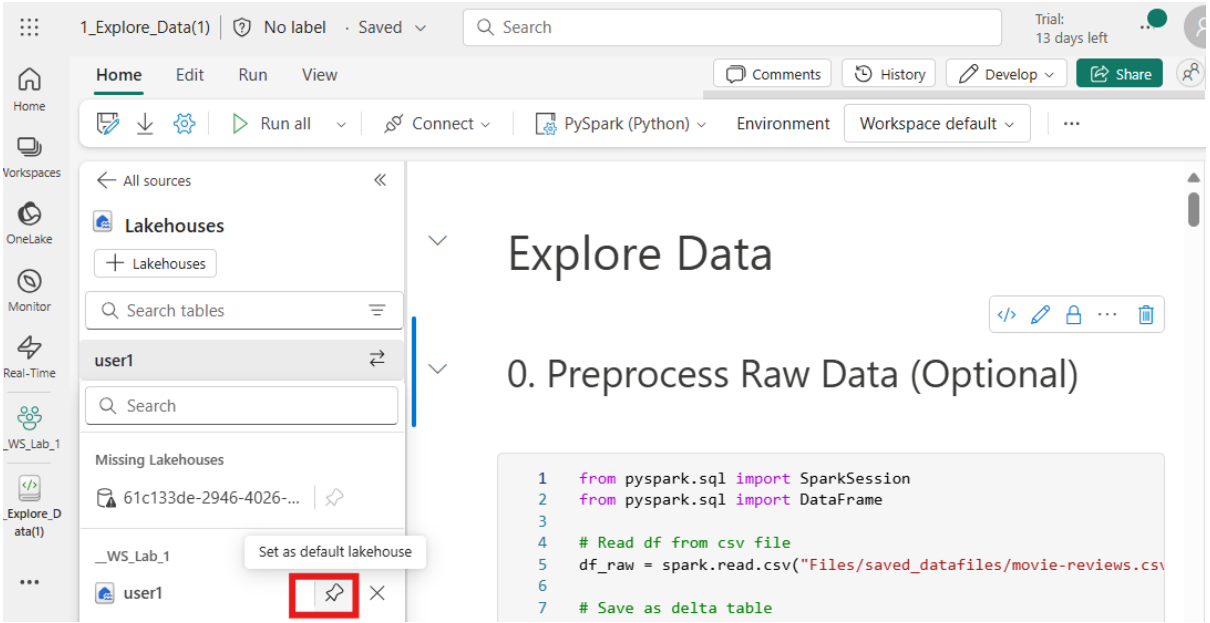




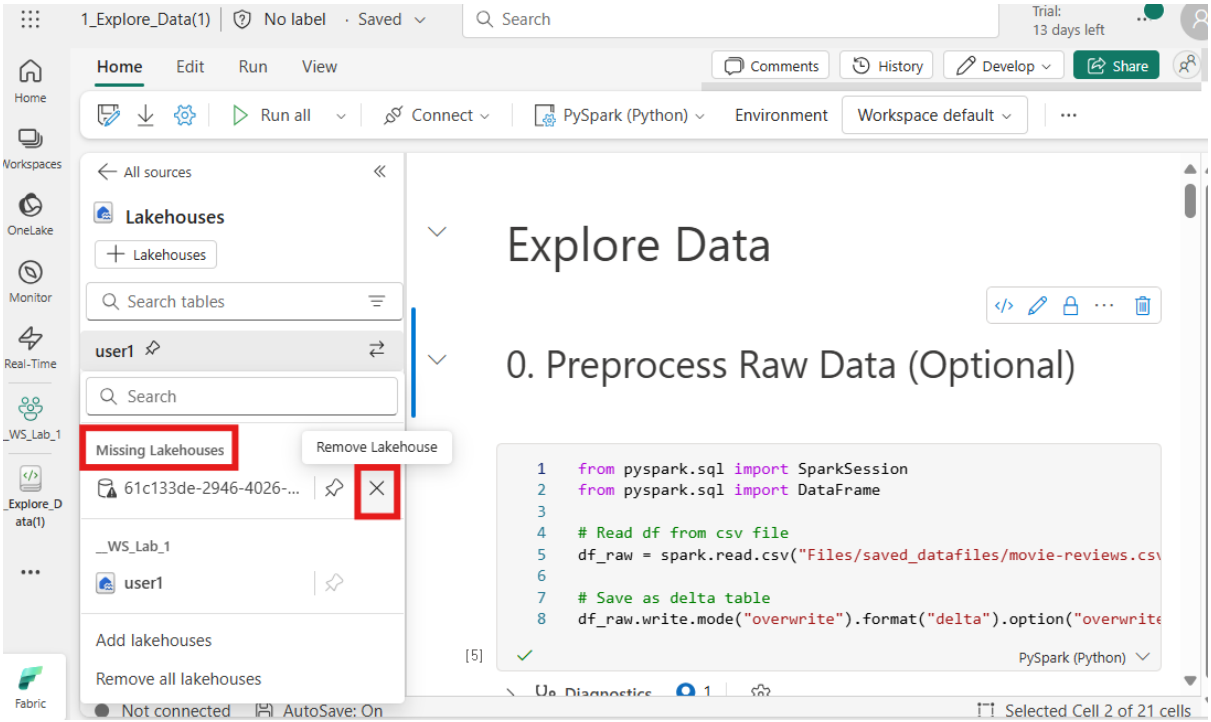


3. 추가한 Lakehouse를 Default Lakehouse로 변경합니다.





4. Missing Lakehouse는 삭제합니다.



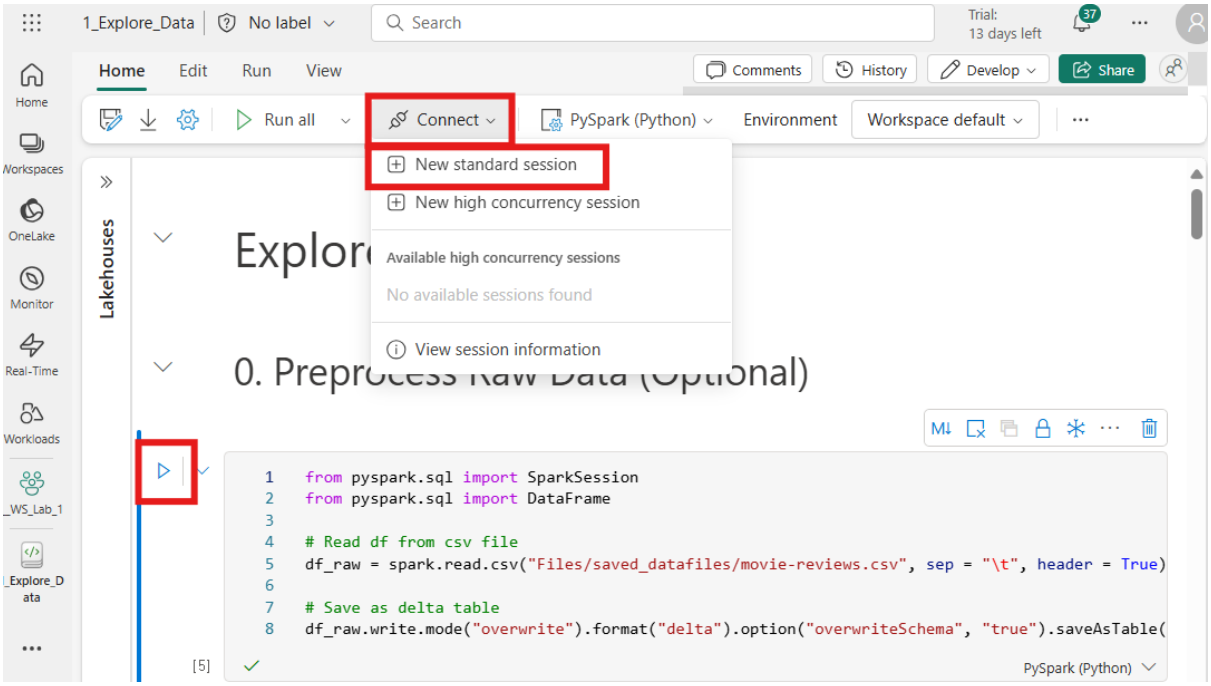
5. 다른 Fabric Notebook에서도 Step 1-4를 반복합니다.

Prepare OneLake Data for Connection

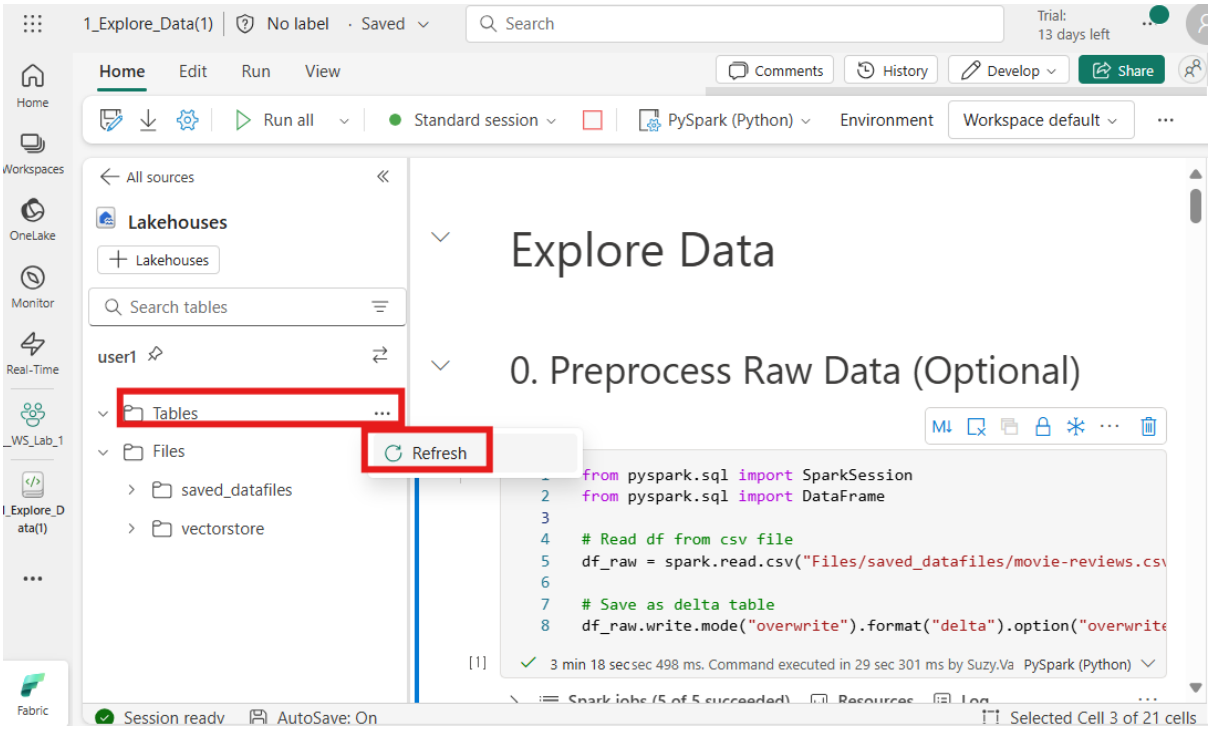
Explore Data

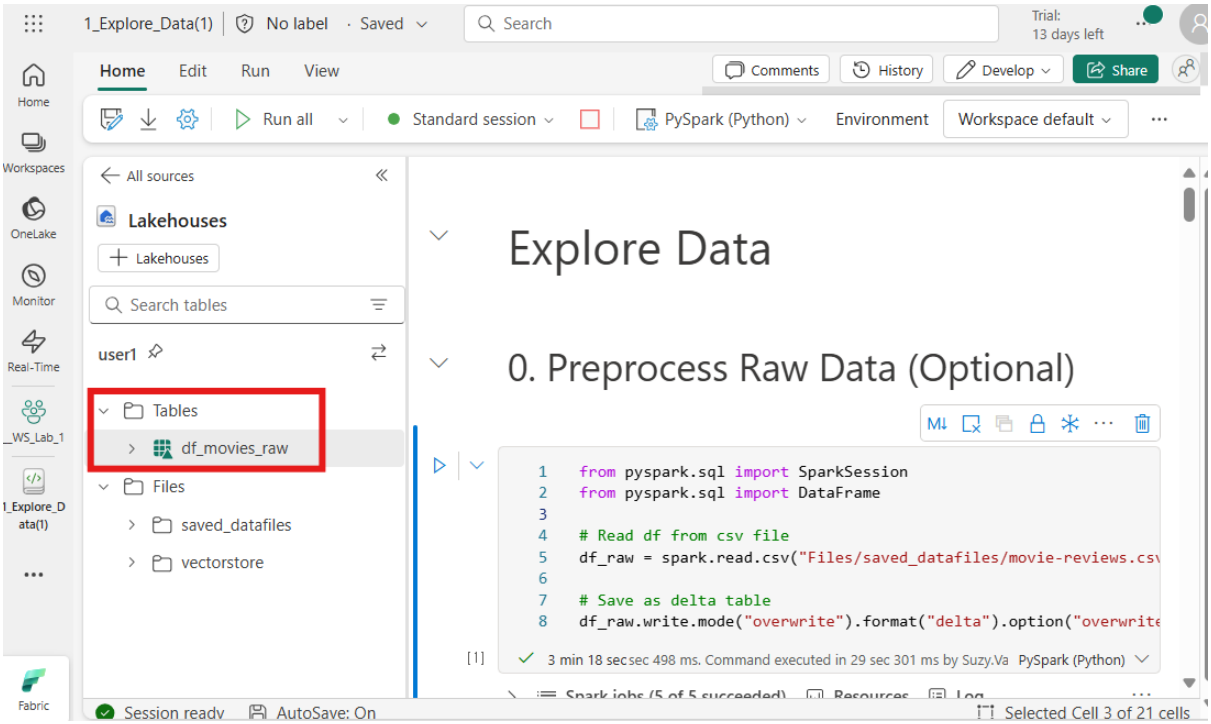
Fabric Notebook 1_Explore_Data를 실행해 데이터를 탐색하고 전처리합니다.

1. 새로운 Spark Session을 연결합니다.



2. 세션이 시작되고 첫번째 셀을 실행하면 Delta Table이 생성된 것을 확인할 수 있습니다.

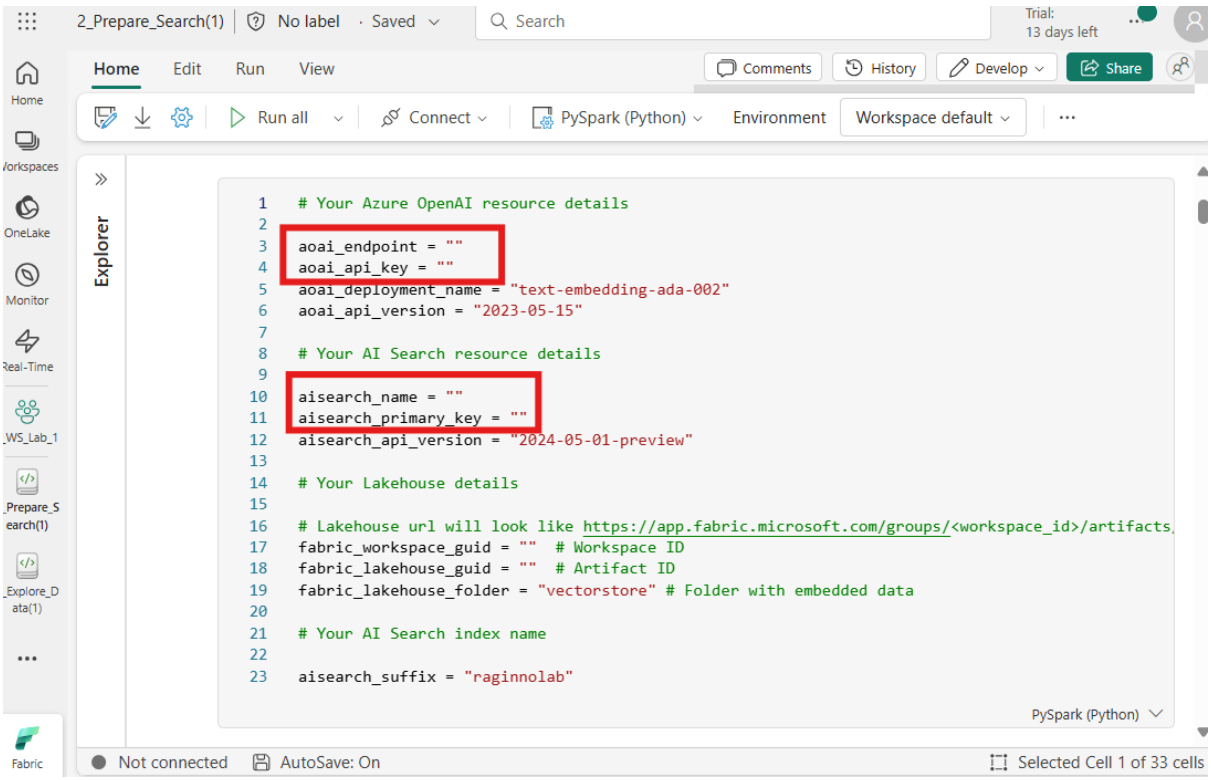




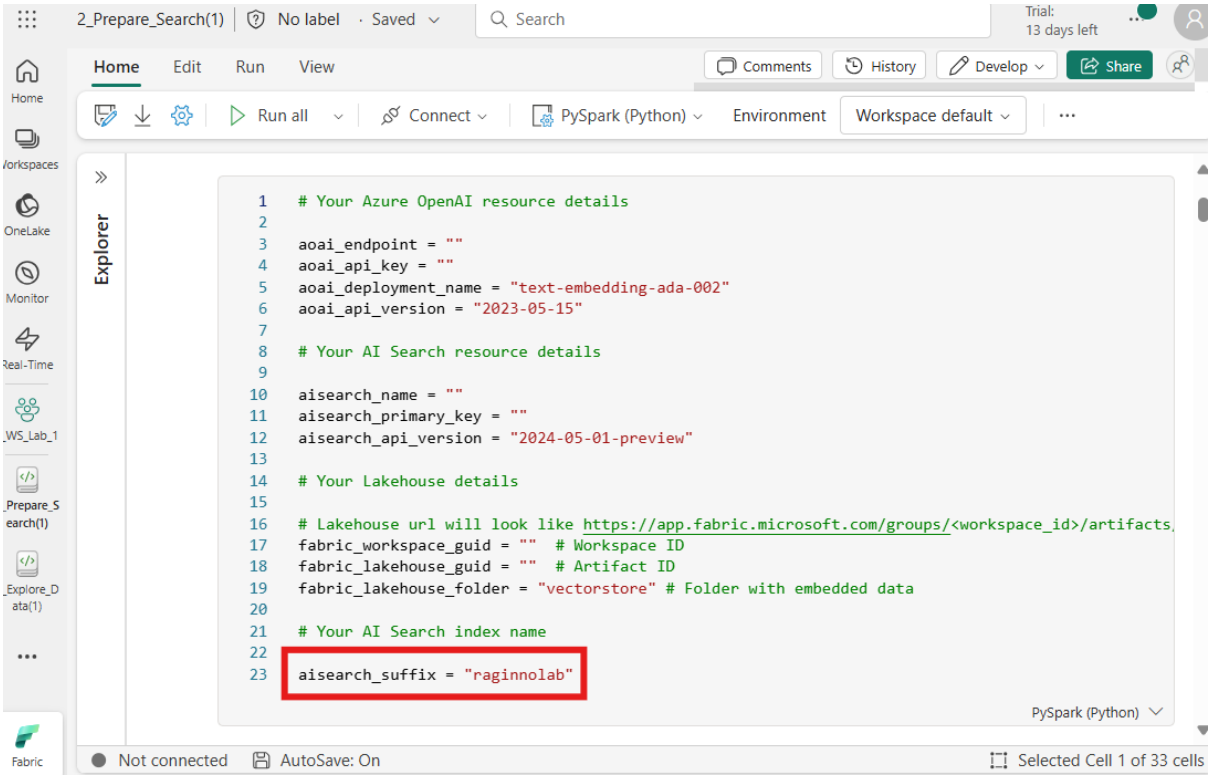
Prepare Search

Fabric Notebook 2_Prepare_Search를 실행해 데이터를 임베딩하고 AI Search 인덱스를 생성합니다.

1. User Setting 정보를 바탕으로 Azure OpenAI와 AI Search의 리소스 정보를 추가합니다.

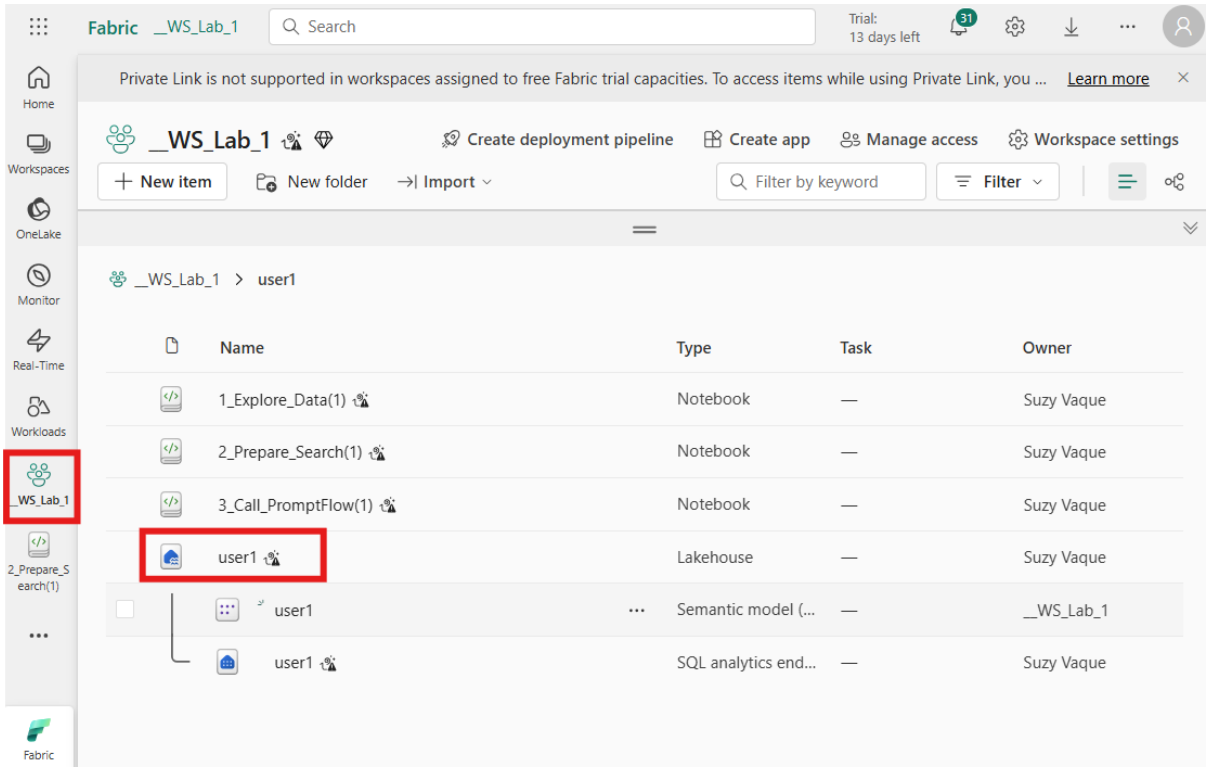


2. AI Search의 Suffix로는 Custom Suffix를 작성합니다.

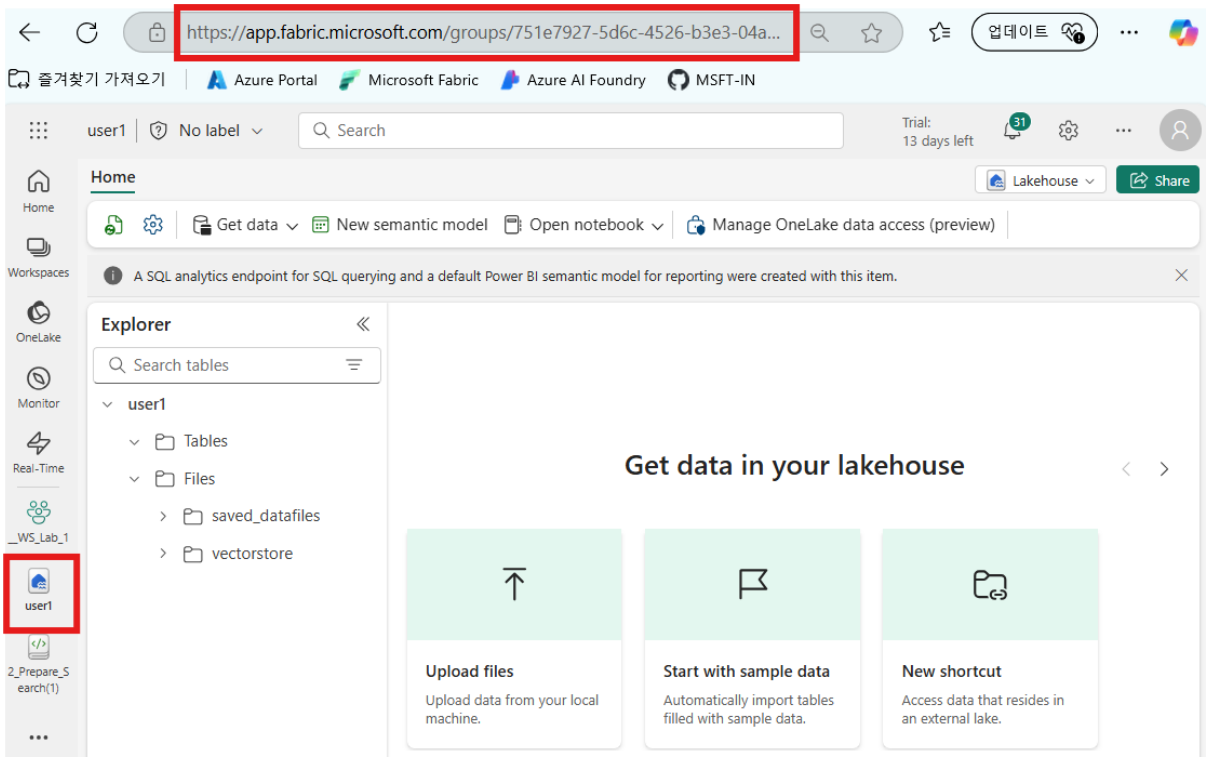


3. Lakehouse 정보를 업데이트합니다.

   A. Workspace 폴더의 Lakehouse를 선택합니다.



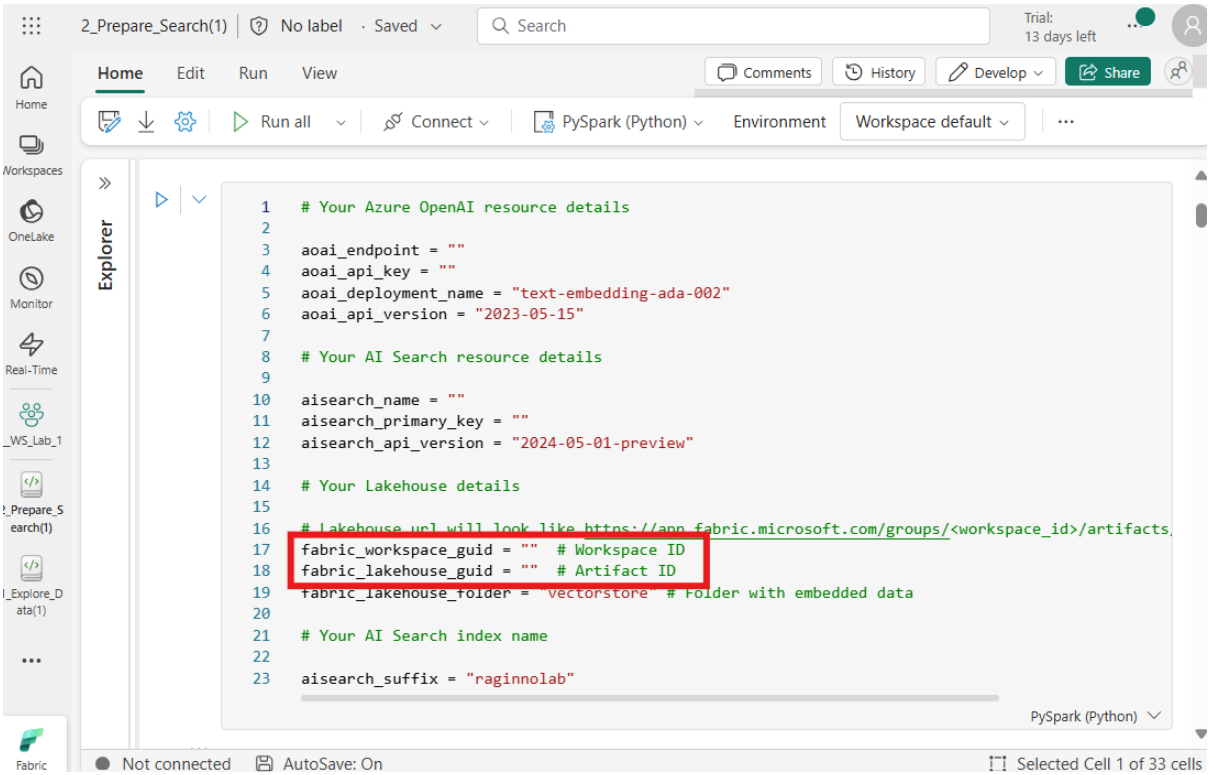
B. Lakehouse에 접속한 상태에서 url을 확인합니다.



C. url의 구조는 다음과 같습니다.

```
https://app.fabric.microsoft.com/groups/{WORKSPACE_GUID}/lakehouses/{ARTIFACT_GUID}
?experience=fabric-developer
```

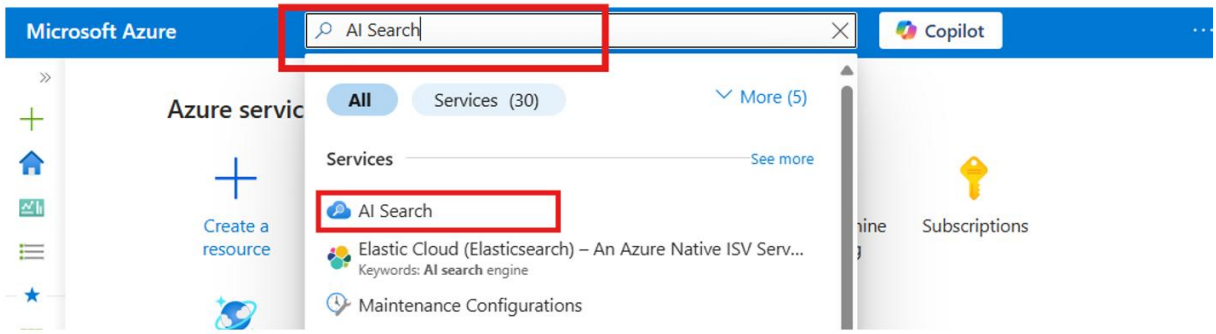
D. Workspace GUID와 Artifact GUID 값으로 노트북의 정보를 업데이트합니다.

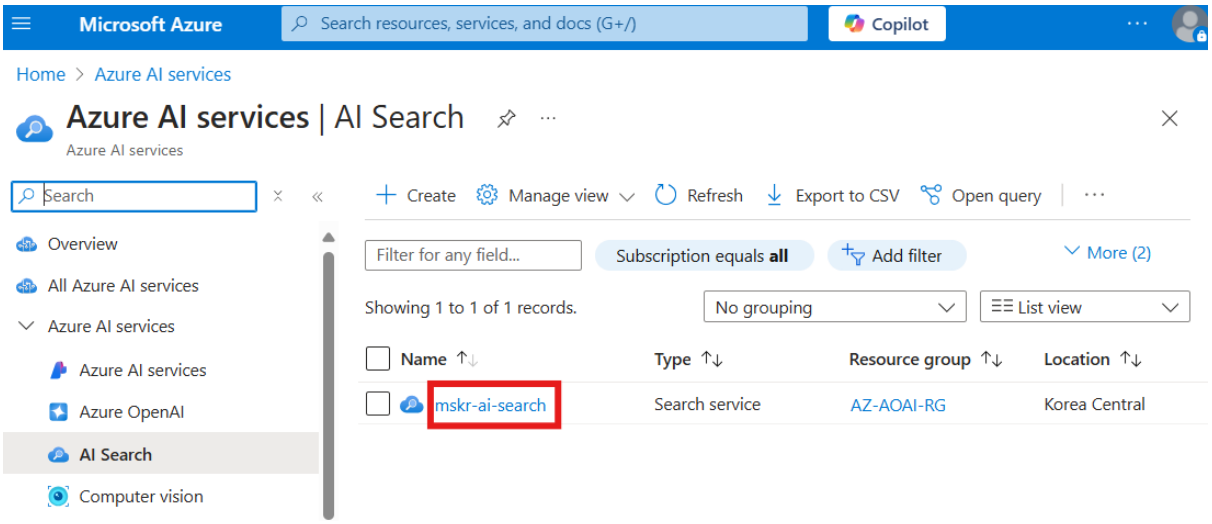


Check AI Search Resource

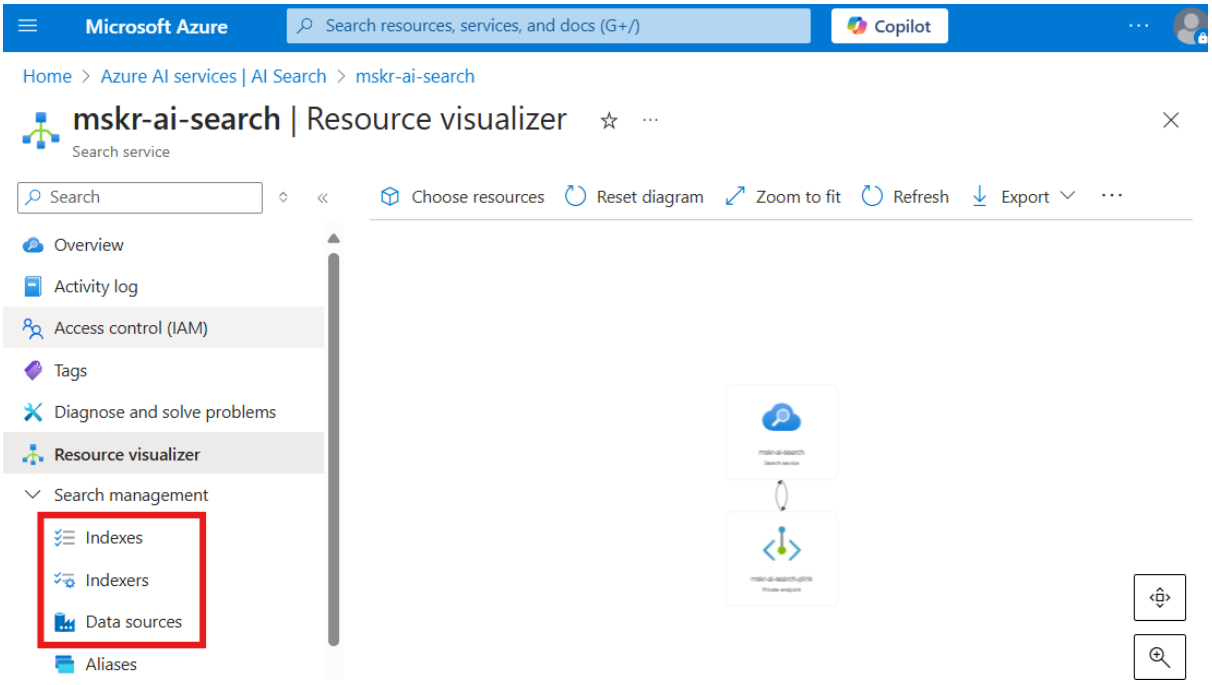
Fabric Notebook 2_Prepare_Search까지 완료했다면, AI Search 리소스에 새로운 Index가 생성됐을 것입니다.

1. Azure Portal에서 AI Search를 검색하고 자신의 리소스를 선택합니다.

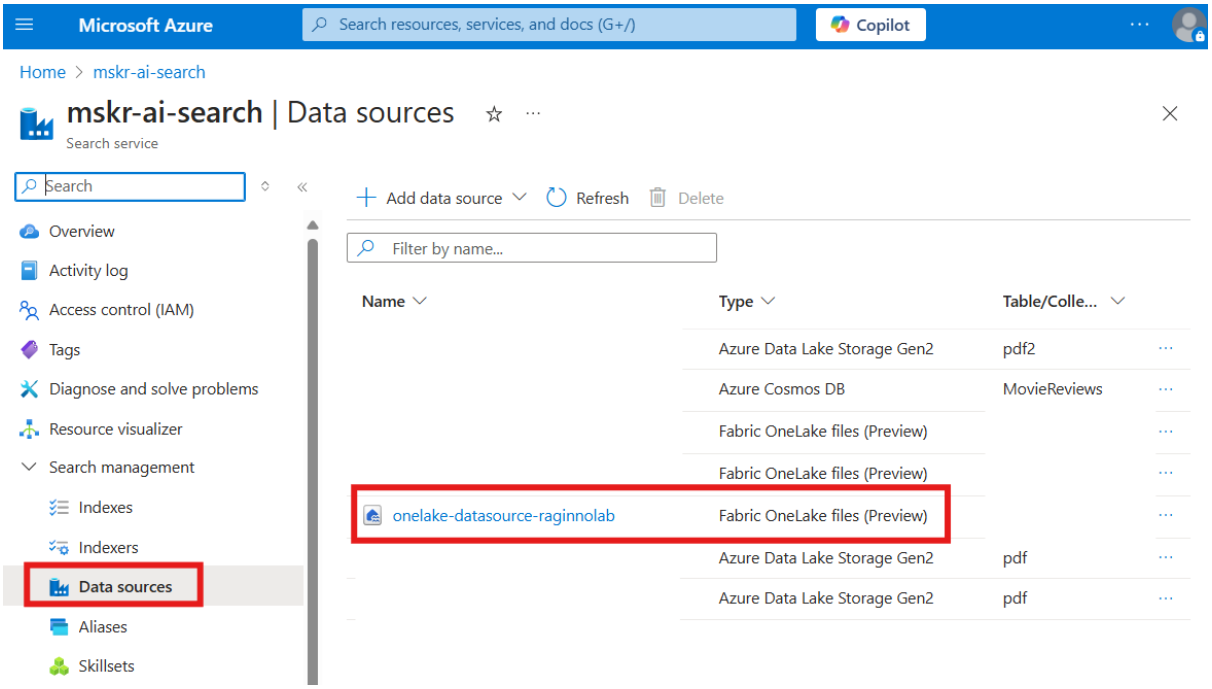




2. 노트북에서 SDK를 통해 생성한 Index, Indexer, Data Source를 확인할 수 있습니다.



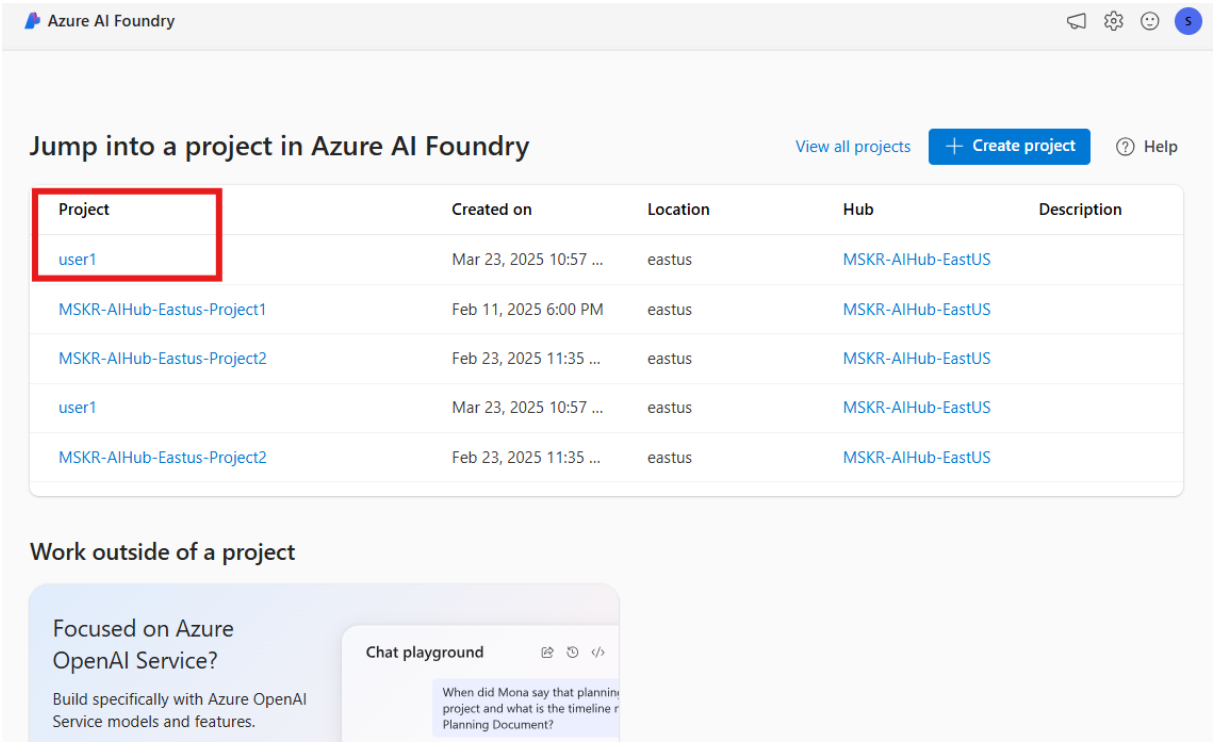
3. Data Source를 확인해보면, Indexer가 One Lake를 바라보고 있음을 확인할 수 있습니다.



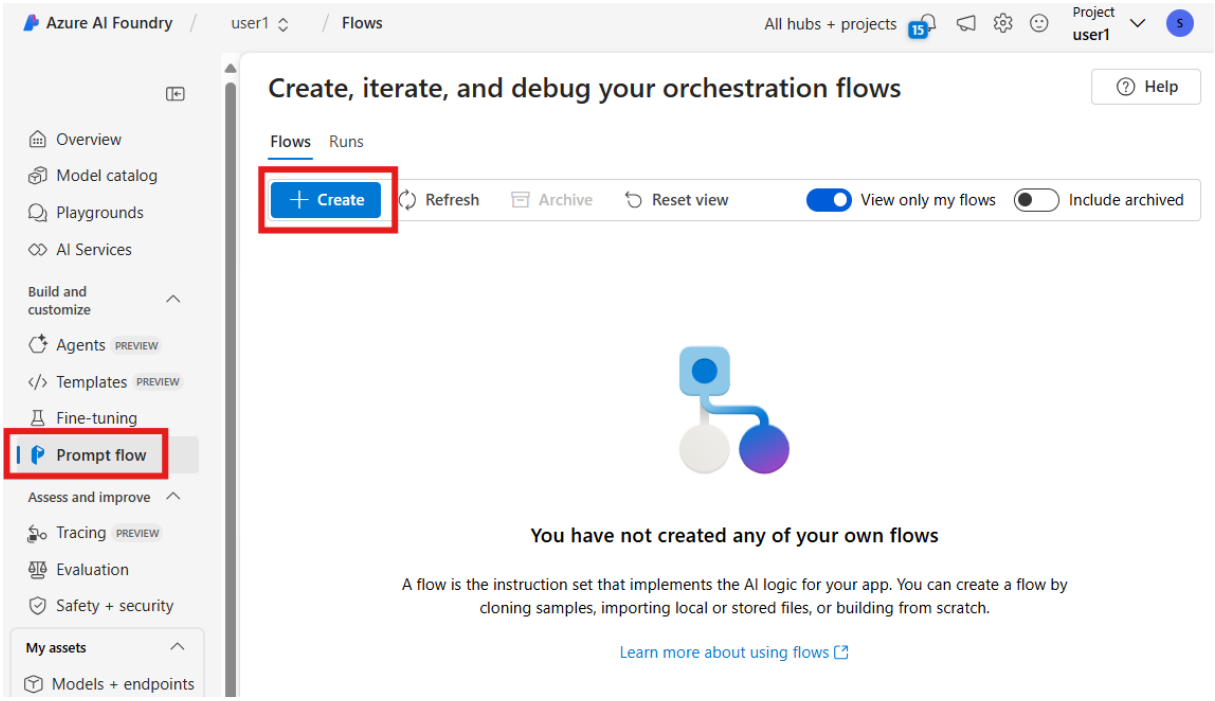
Create AI Applications on OneLake Data

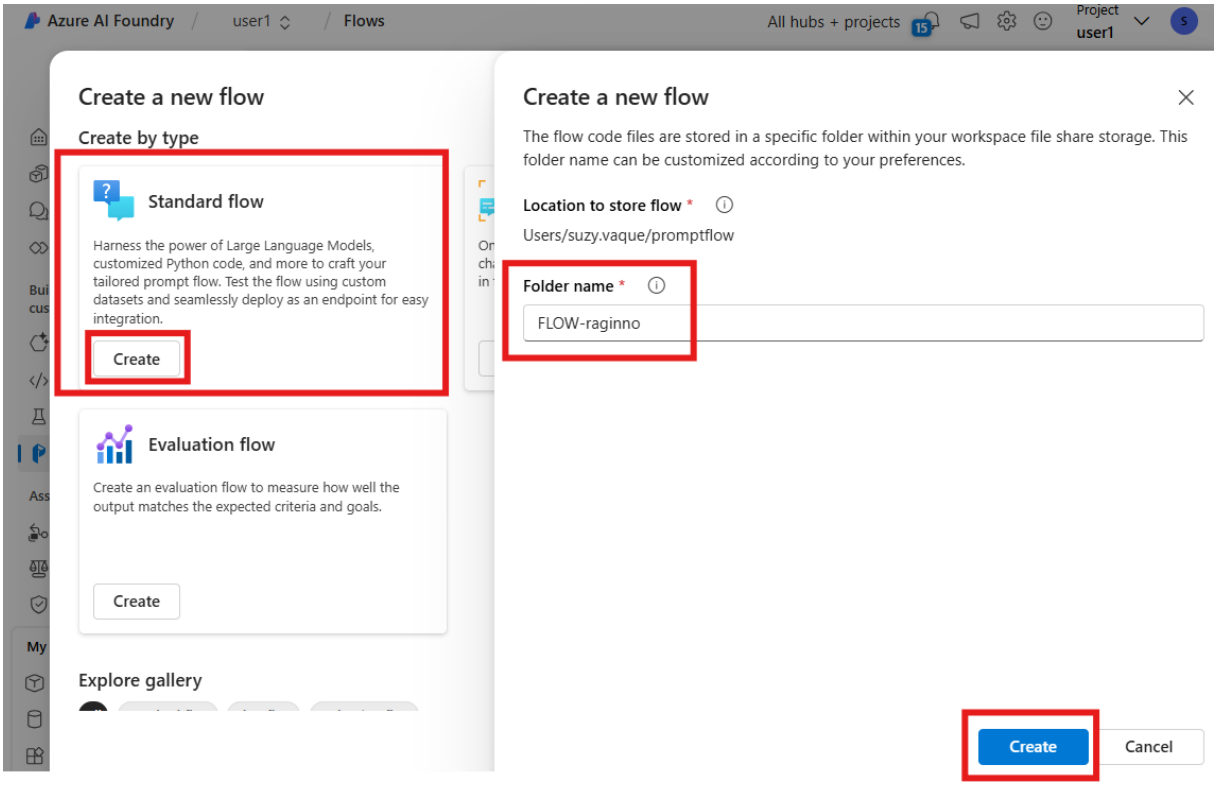
Get Started with PromptFlow

1. Foundry Portal의 AI Project로 이동합니다.



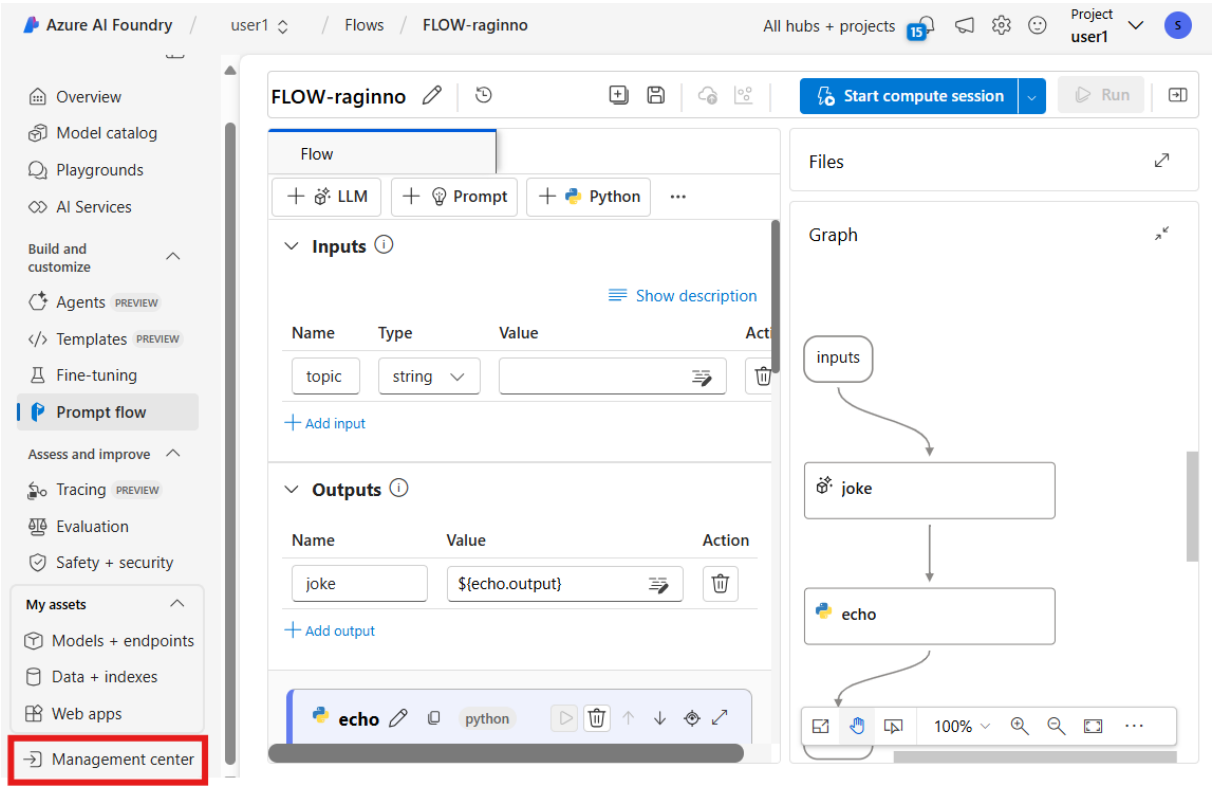
2. Prompt Flow 서비스로 이동해 새로운 Standard Flow를 생성합니다.



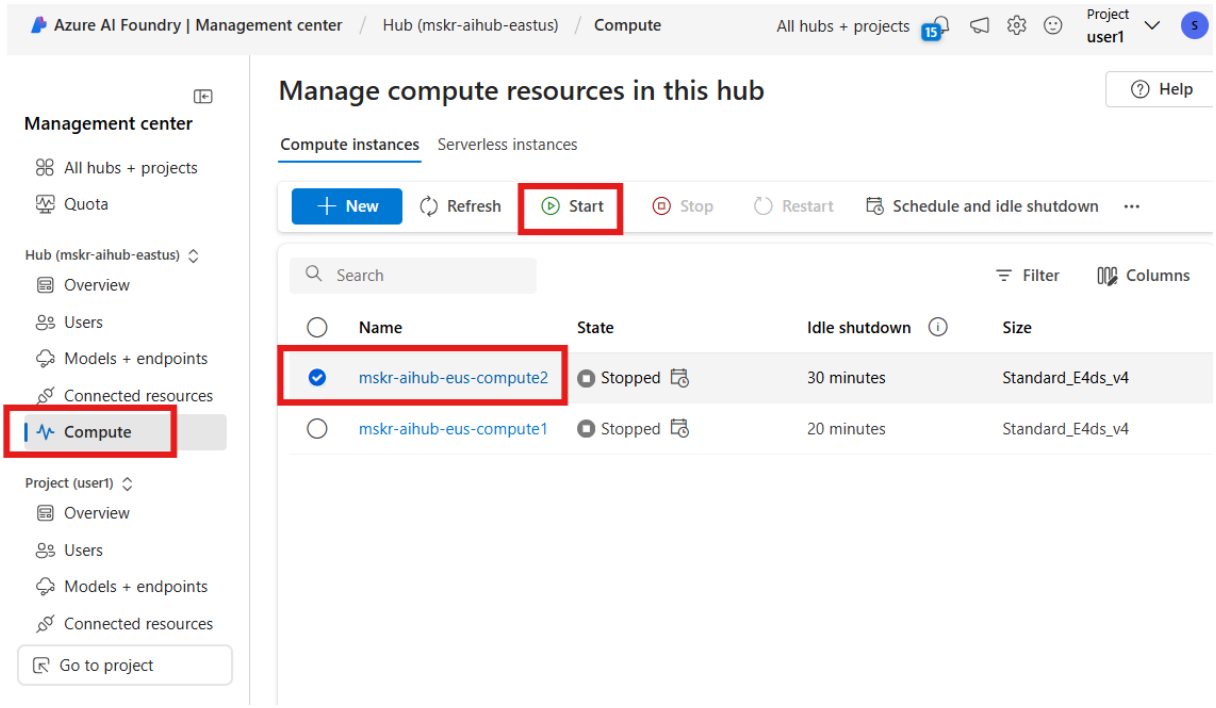


3. Flow를 실행하기 위해 배정받은 Compute Instance를 연결합니다.

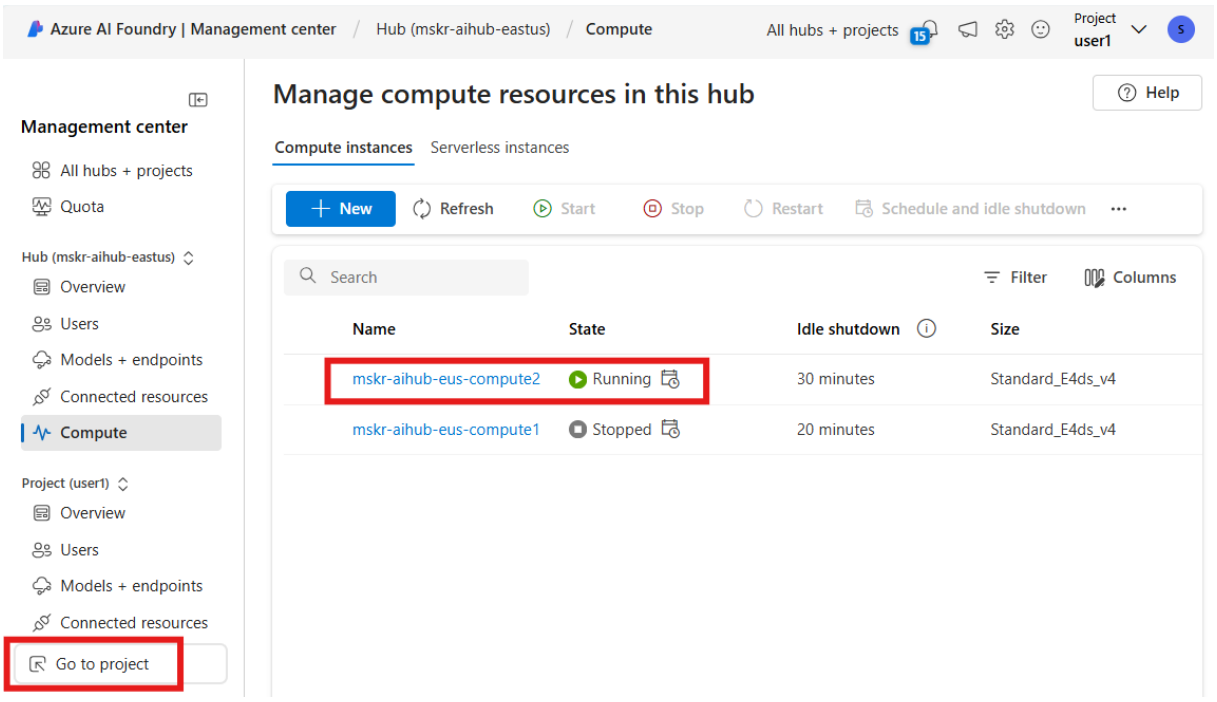
A. AI Hub의 Management Center로 이동합니다.



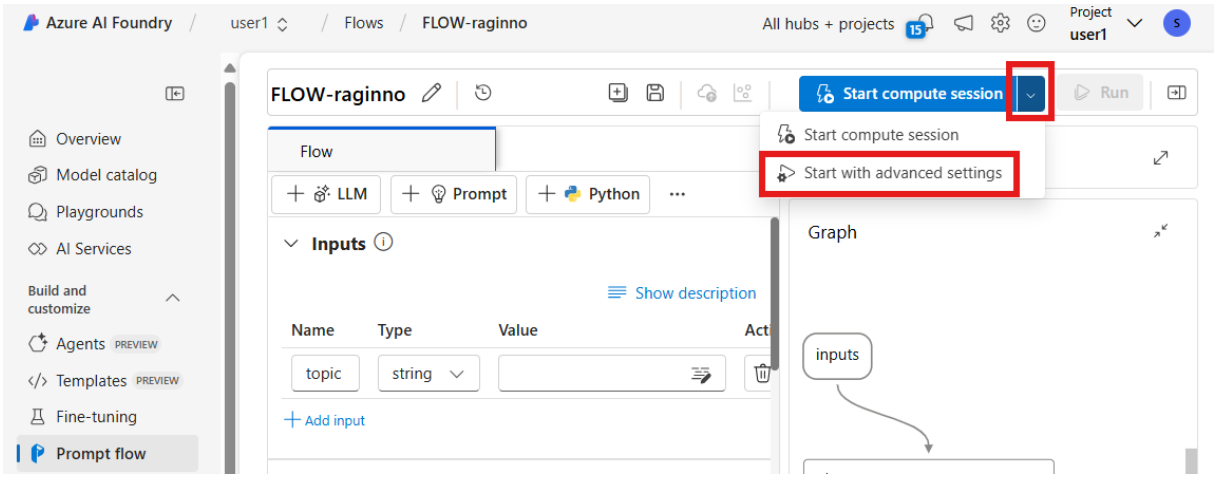
B. 배정받은 Compute를 실행합니다.

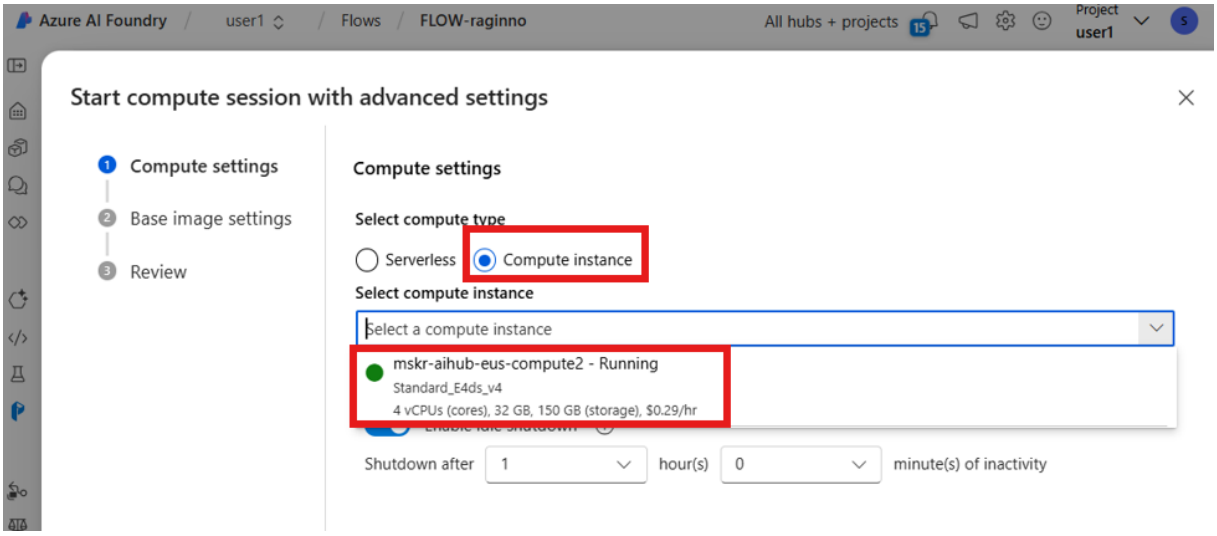


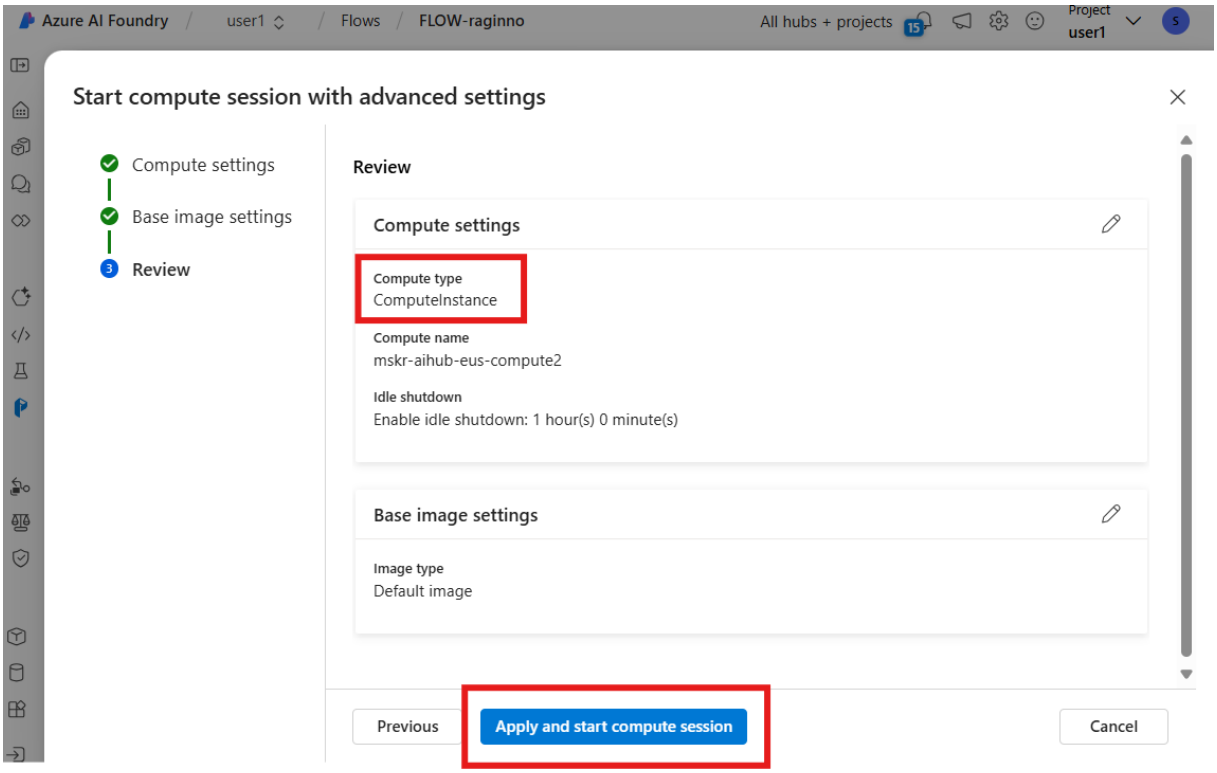
C. Compute 실행이 완료되면 다시 AI Project의 Flow로 이동합니다.

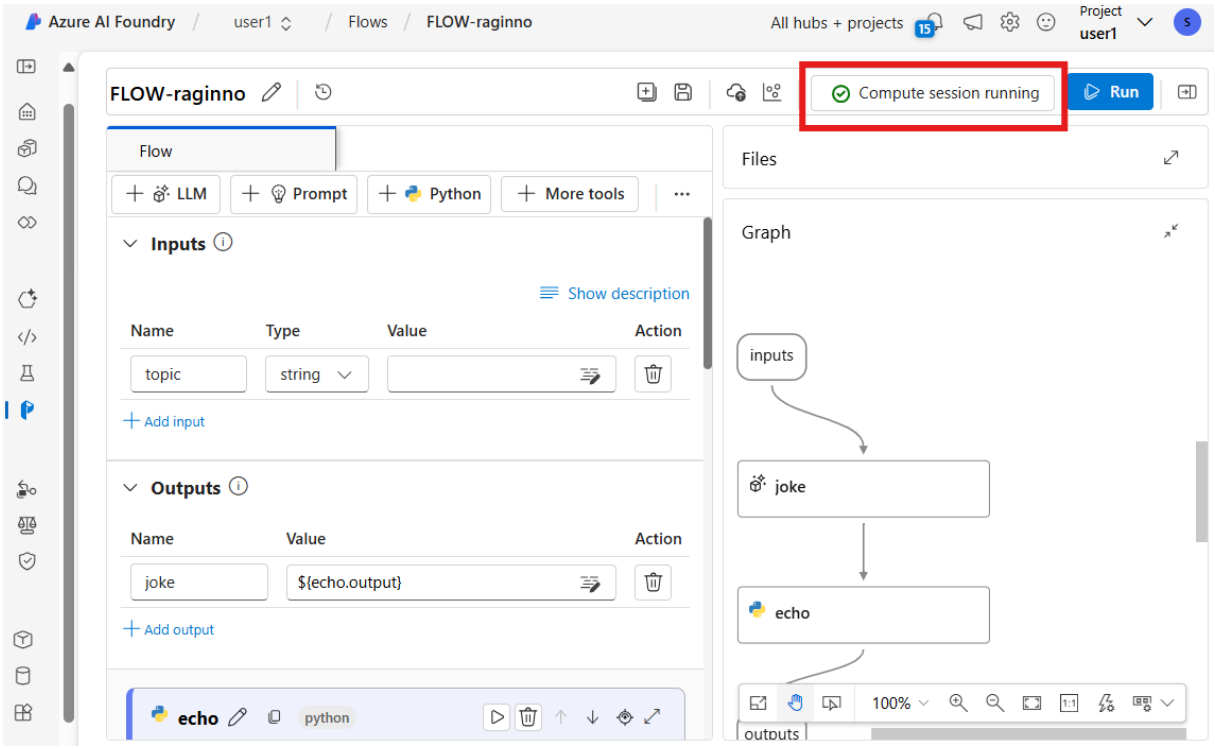


D. Compute Session의 Start with Advanced Settings 옵션을 선택해 Compute Instance를 연결합니다.



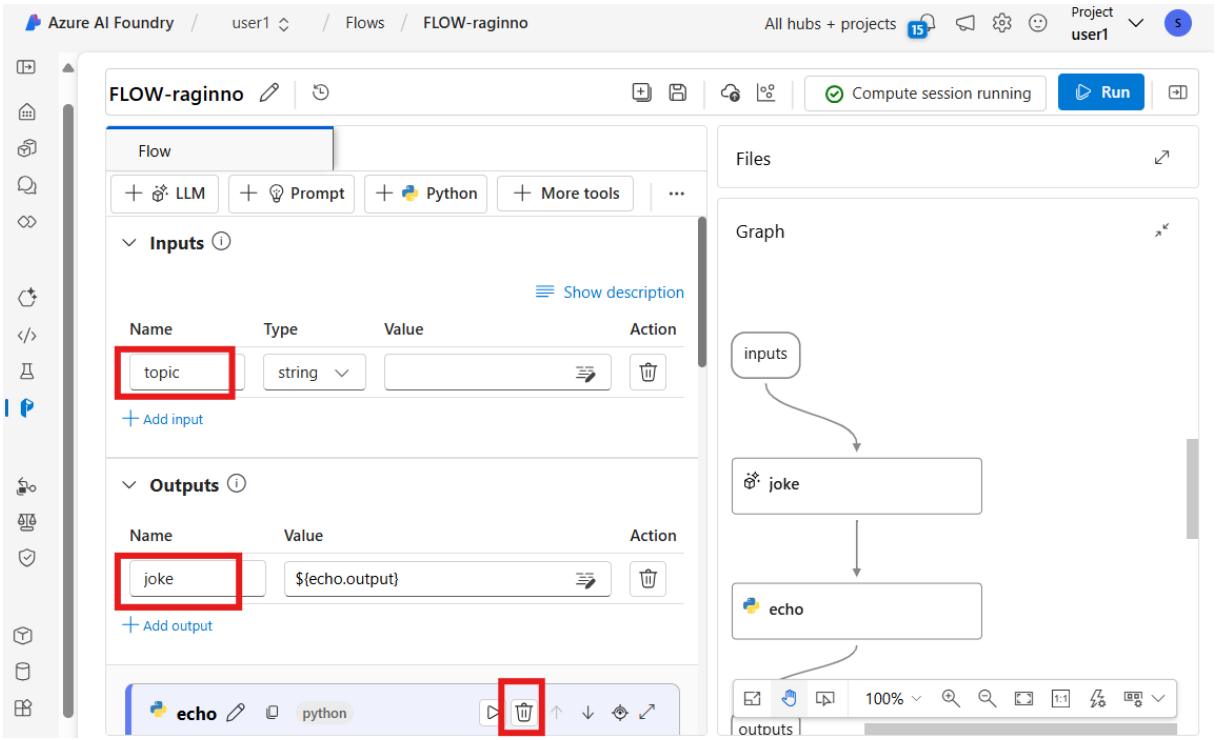


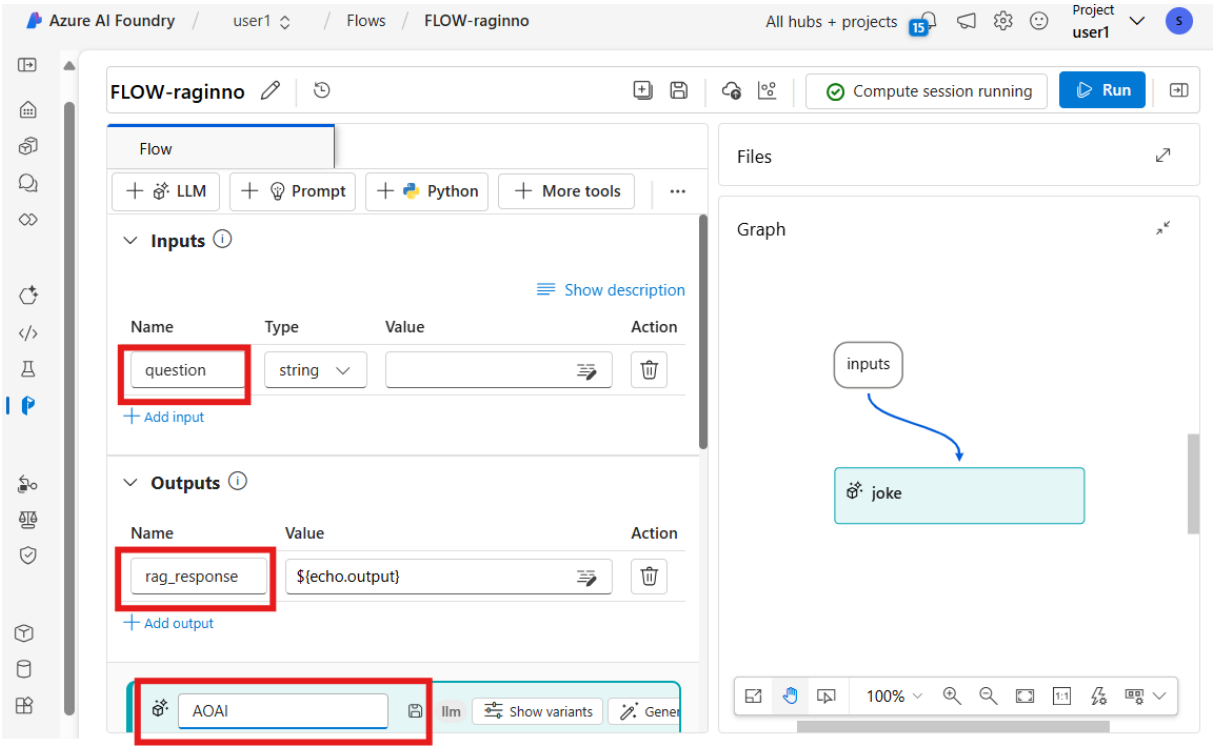




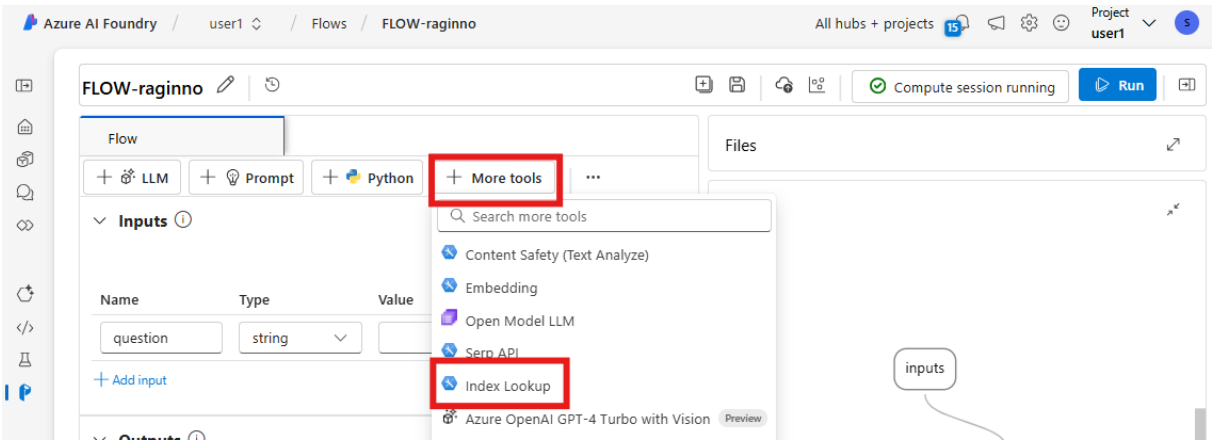
Create RAG Flow

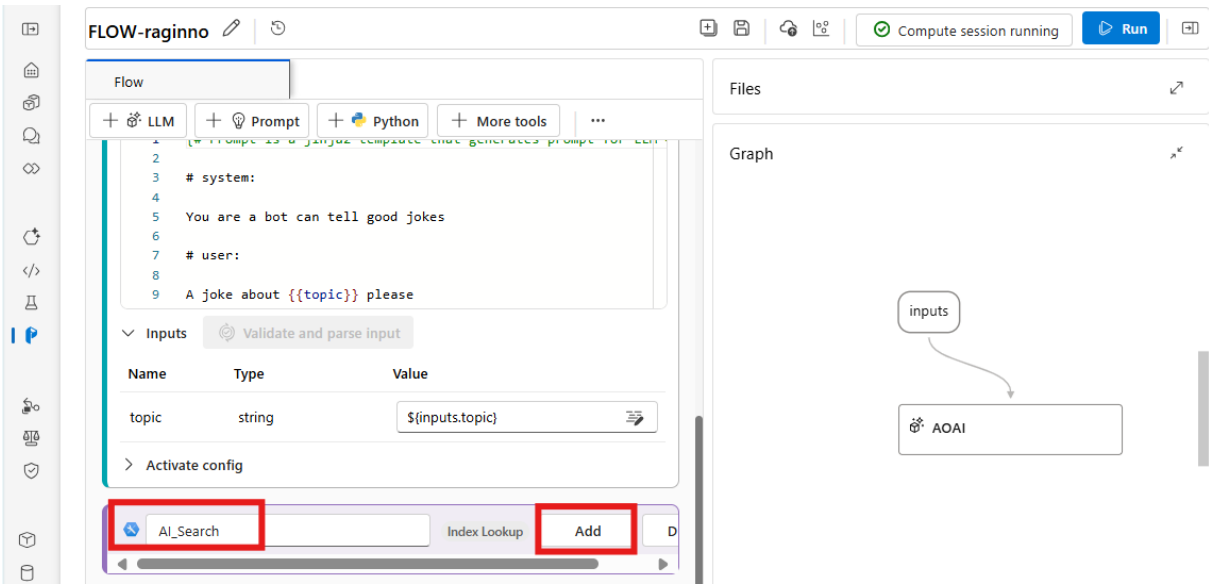
1. Flow에 기본적으로 생성된 Input, Output, LLM Node의 이름을 적절하게 변경하고, Python Node를 삭제합니다.





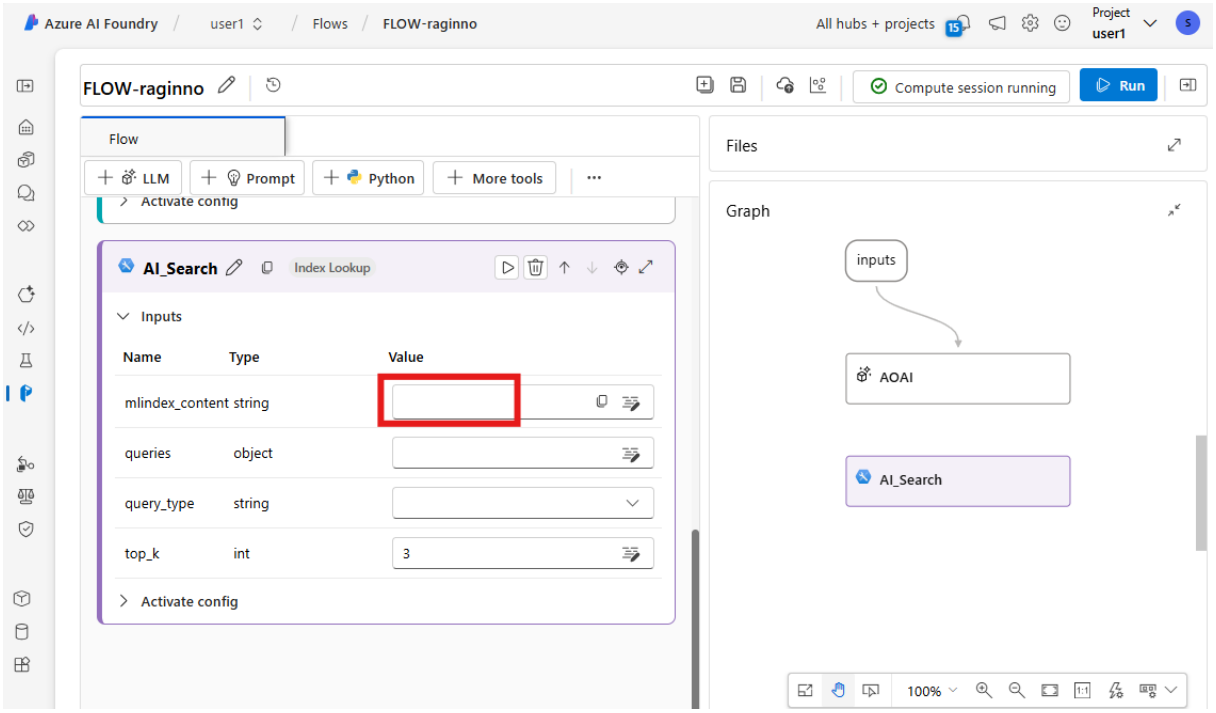
2. Index Lookup Tool을 추가합니다.

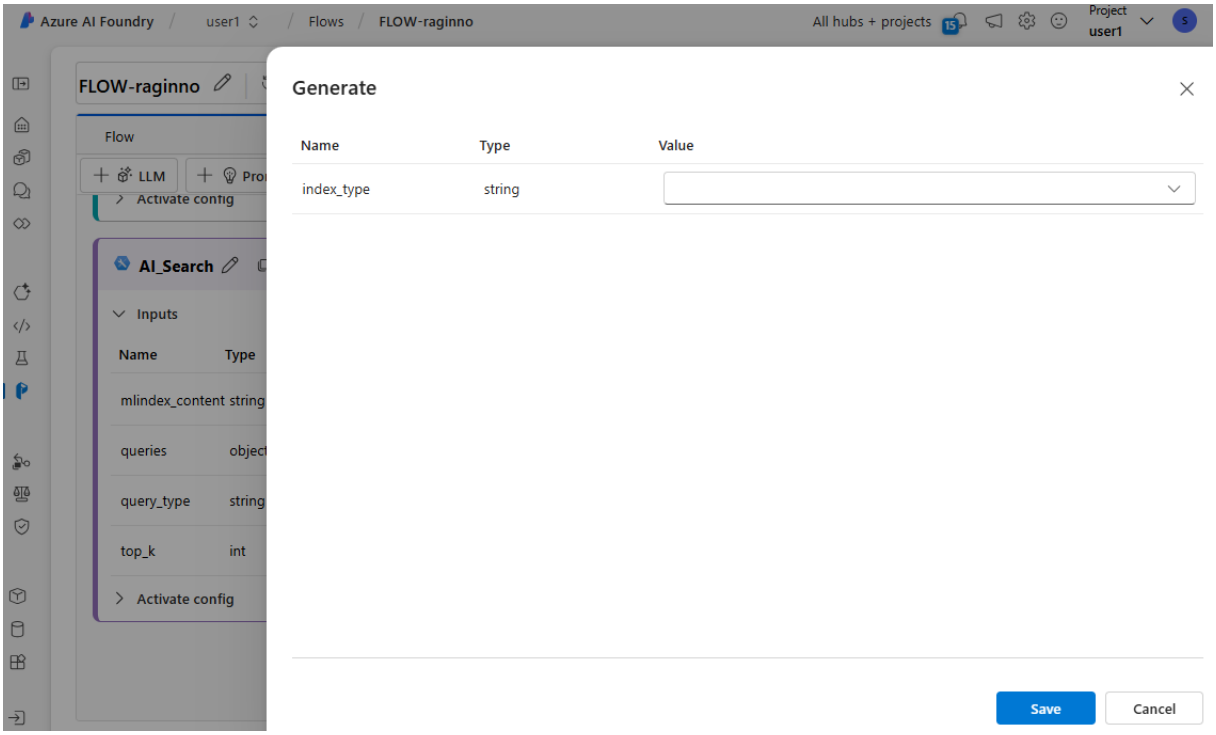




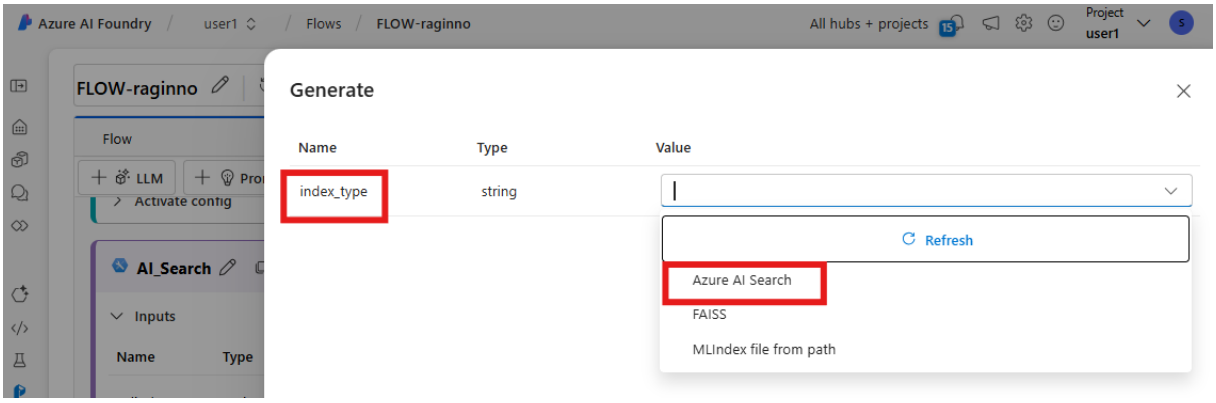
3. Index Lookup에 OneLake 기반 AI Search Index를 연결합니다.

   A. Lookup Generate으로 이동합니다.

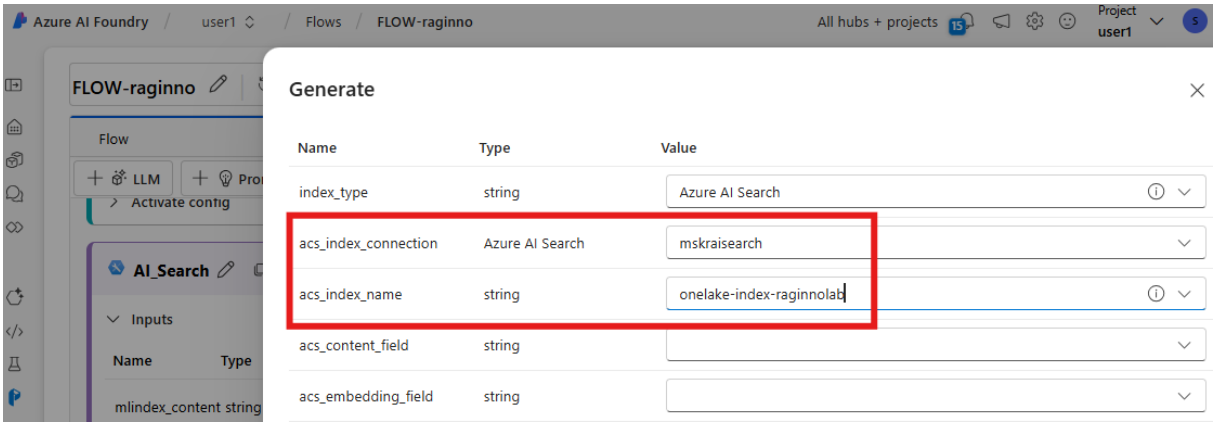




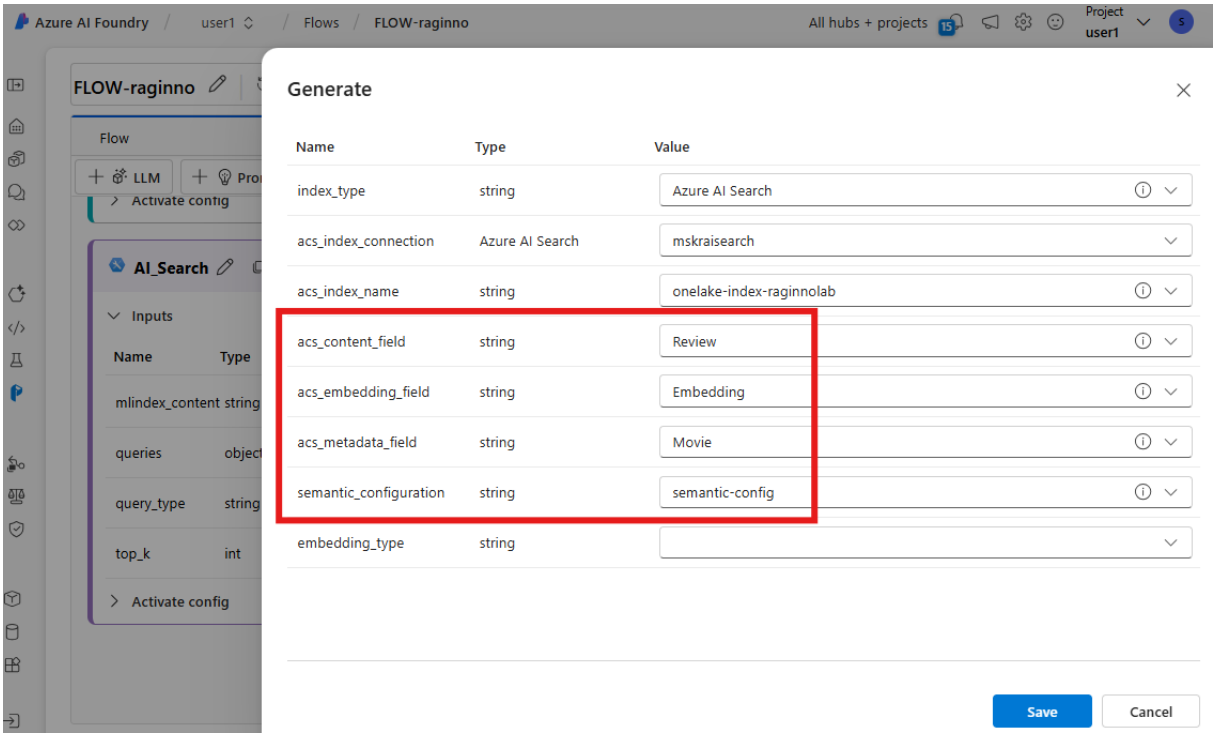
B. Index Type으로 Azure AI Search를 선택합니다.



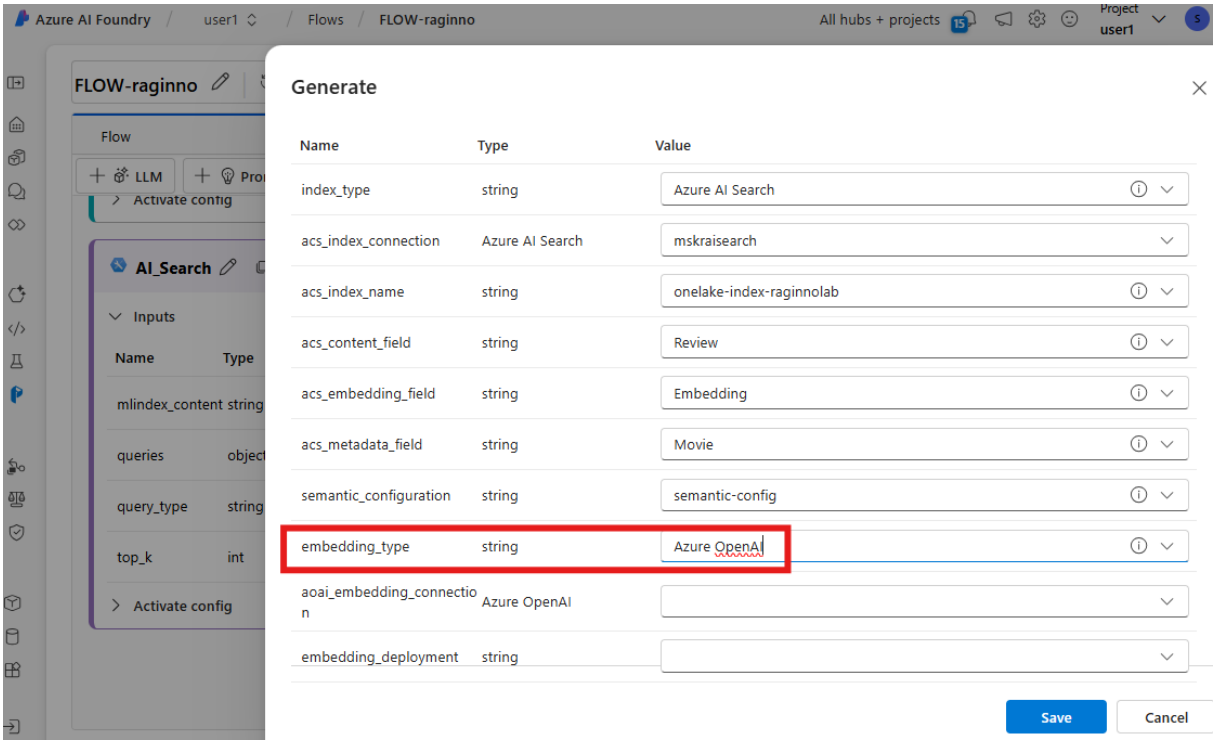
C. 배정받은 AI Search 리소스와 앞서 생성한 Index를 선택합니다.



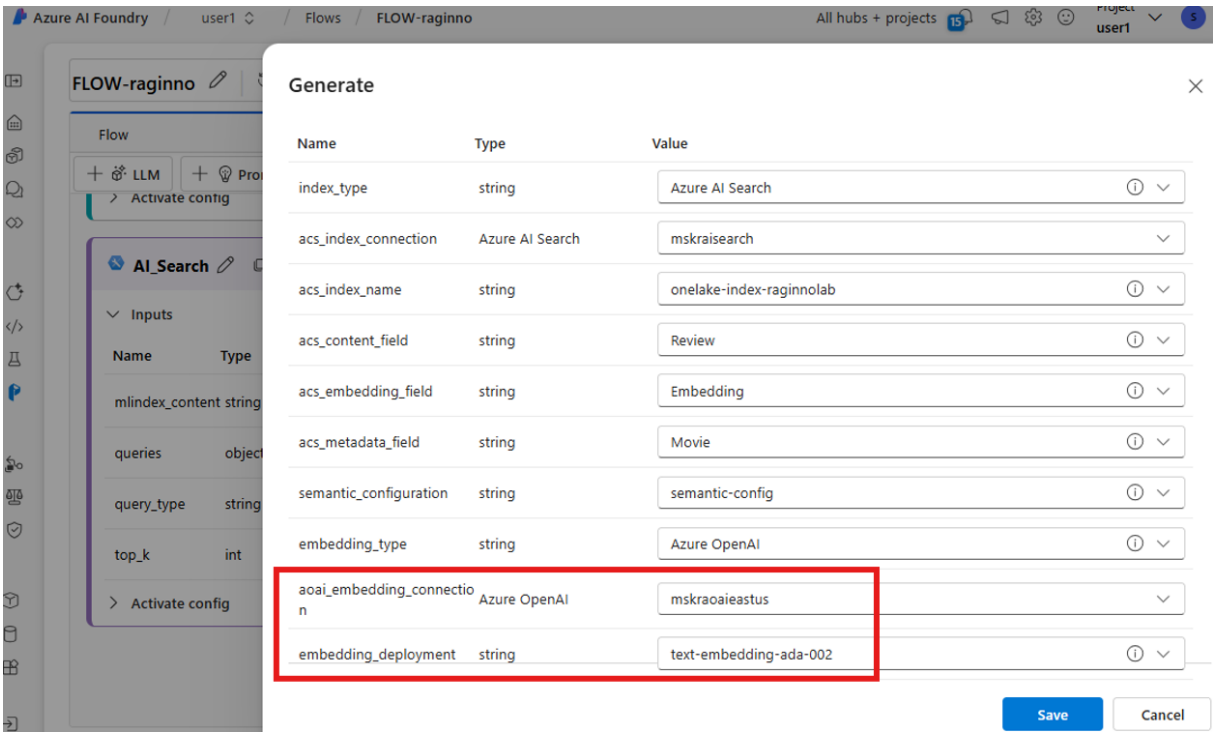
D. 필드 정보를 선택합니다.



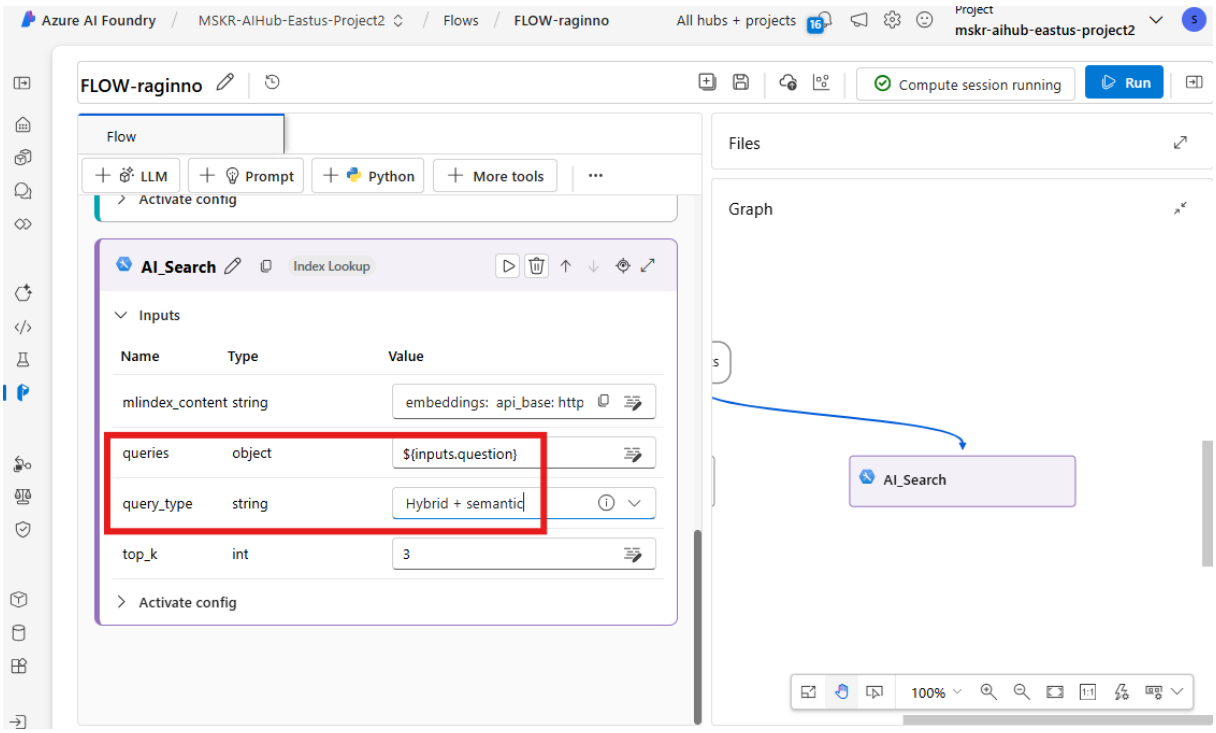
E. 쿼리를 임베딩해 Vector Search할 수 있도록 Embedding Type으로 Azure OpenAI를 선택합니다.



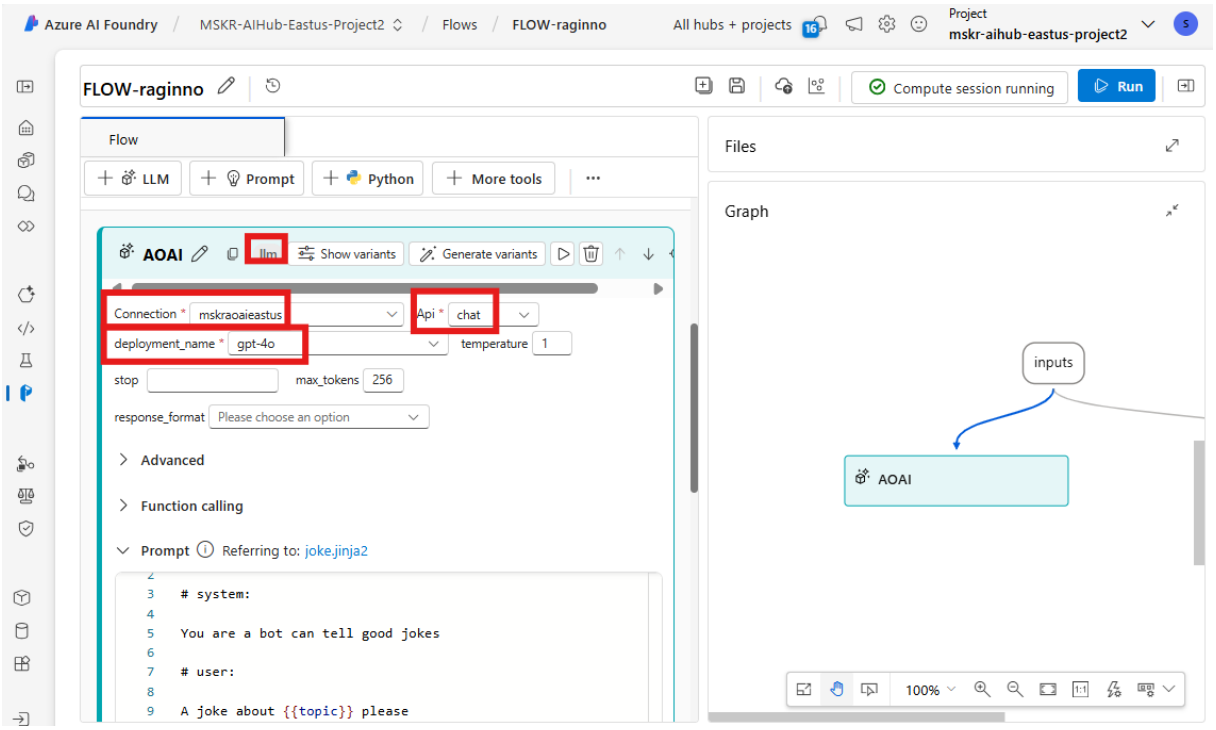
F. 배정받은 Azure OpenAI 리소스를 연결합니다.



G. 생성된 Index Lookup Node의 Queries와 Query Type을 지정합니다.

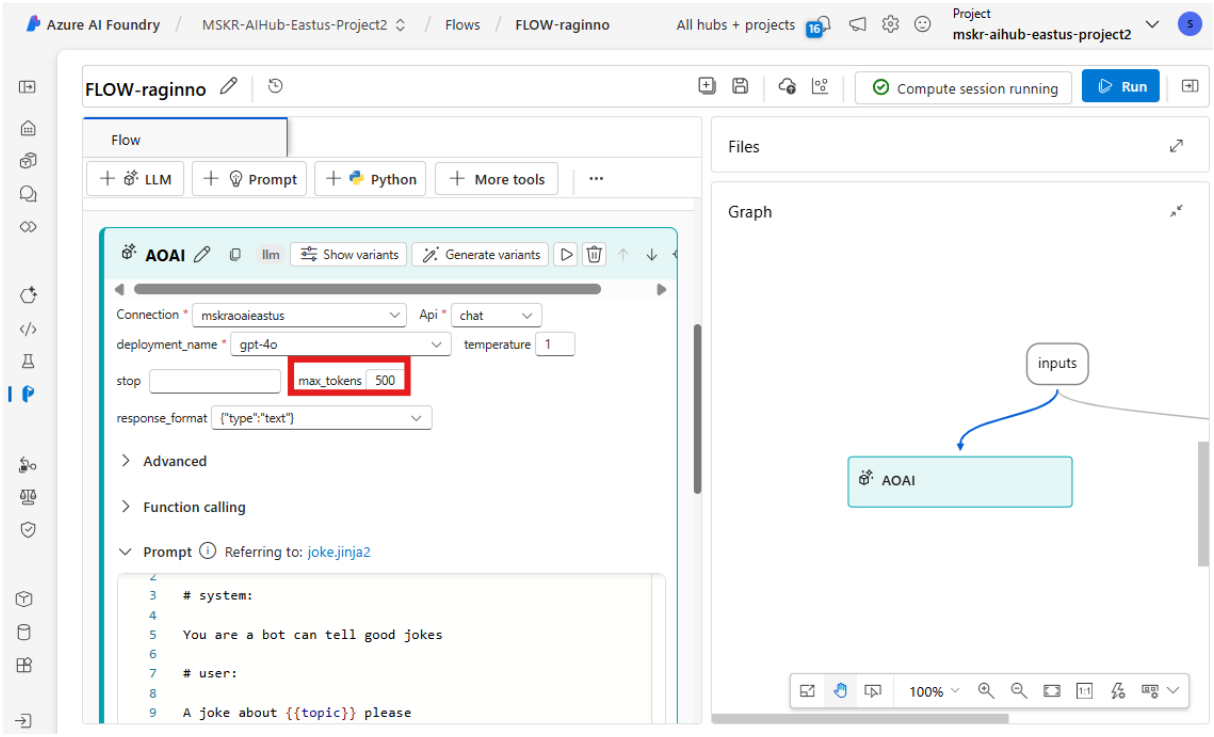


4. LLM Node에 배정받은 Azure OpenAI 리소스를 연결합니다.

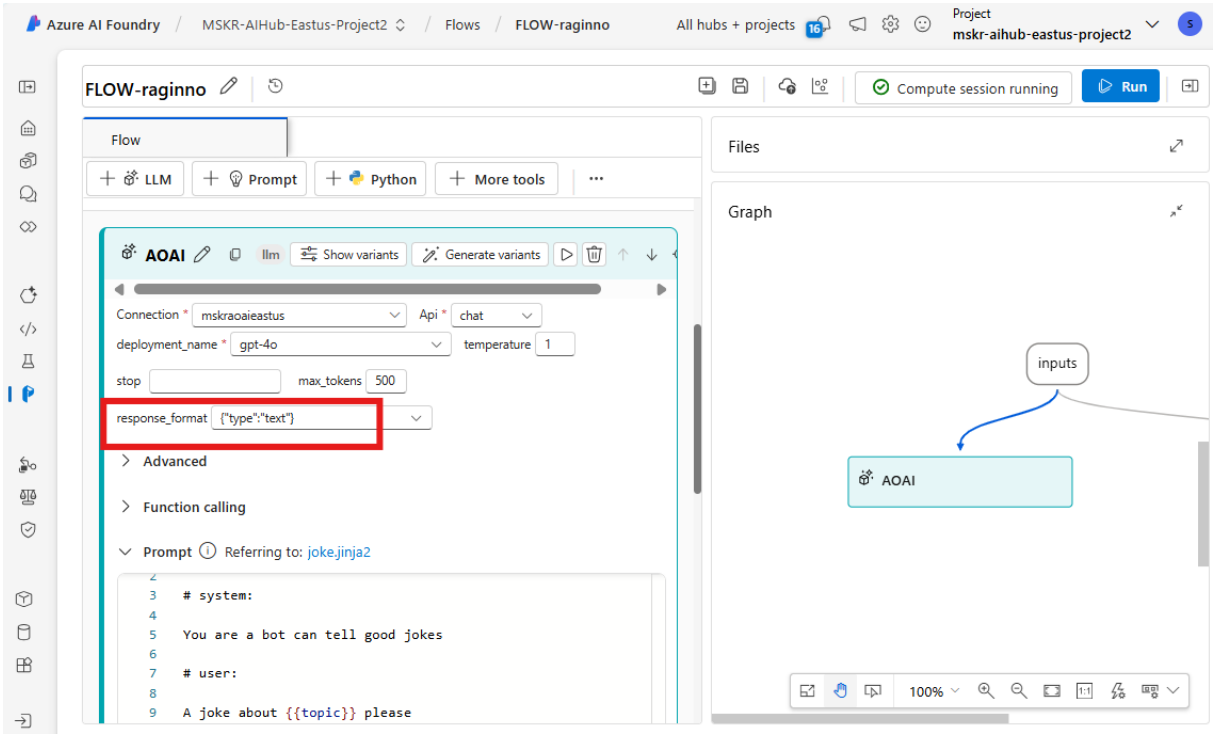


5. LLM Node를 설정합니다.

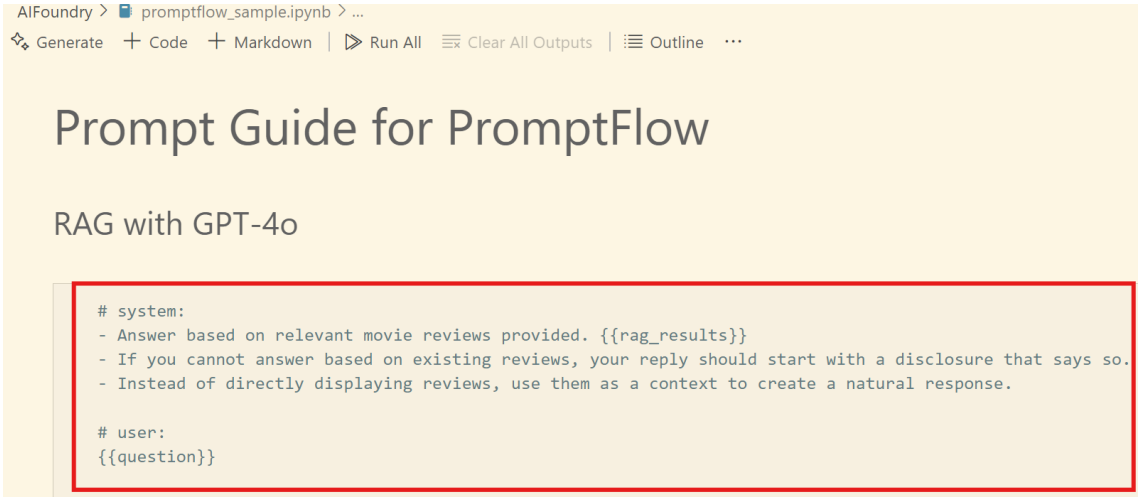
A. Max Tokens를 수정합니다.

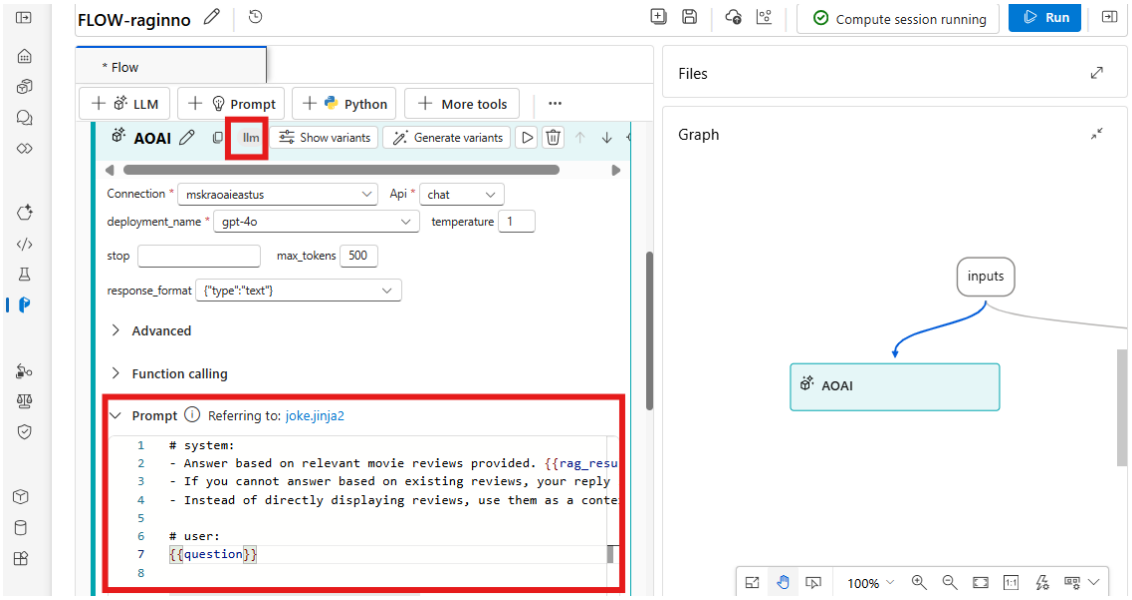


B. Response Type으로 Text를 선택합니다.

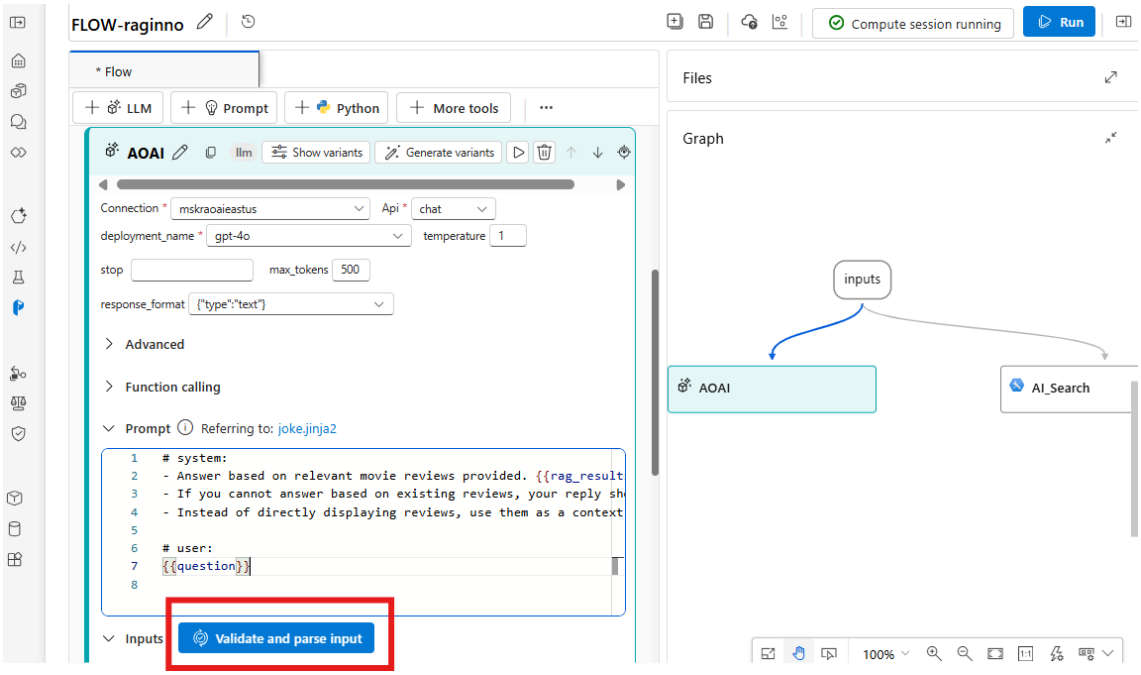


C. 클론된 폴더의 AIFoundry\promptflow_sample.ipynb에 있는 RAG 프롬프트 샘플을 LLM 프롬프트로 사용합니다.

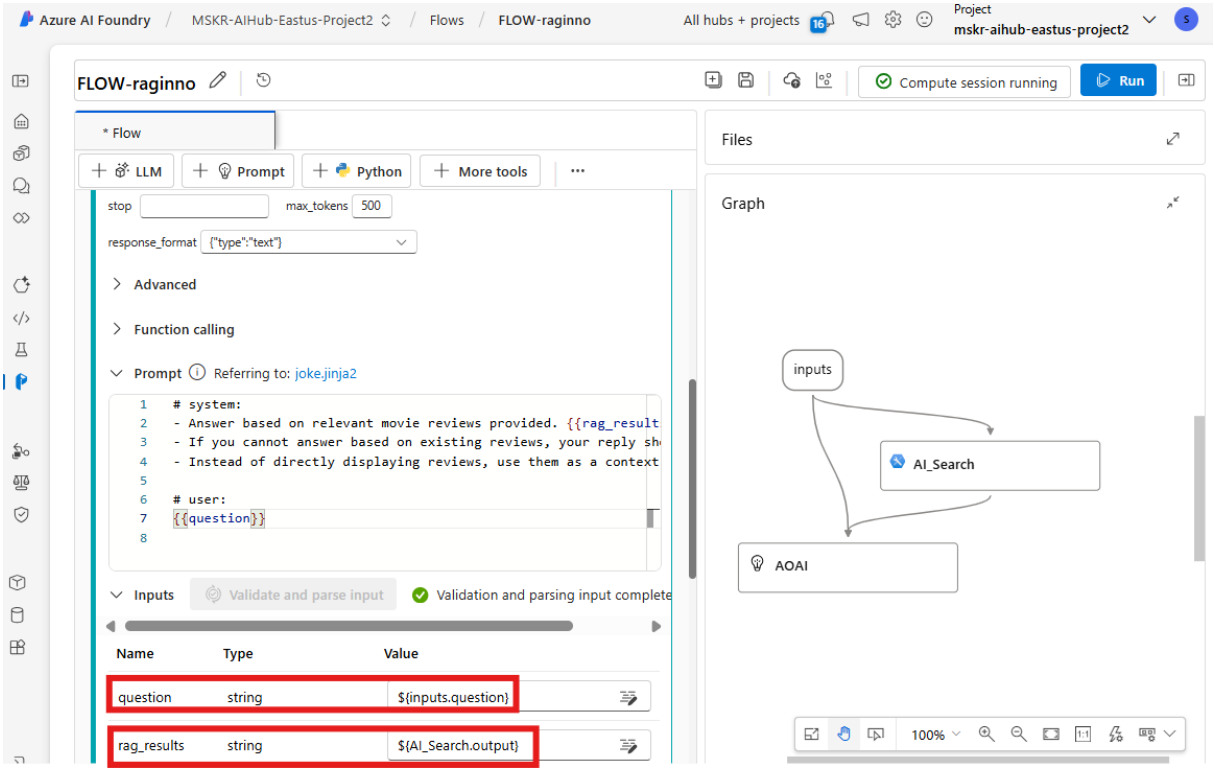




D. Validate and Parse Input을 선택합니다.

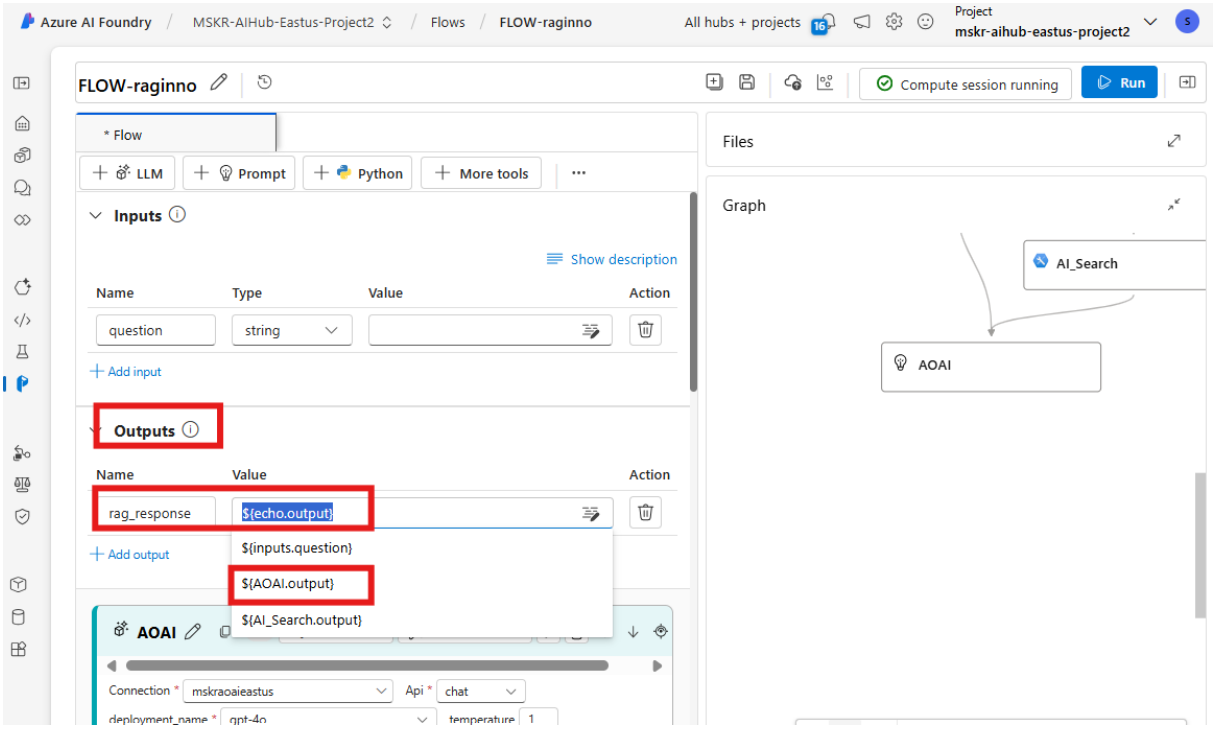


E. Validate가 완료된 후 Input을 적절한 값으로 연결합니다.

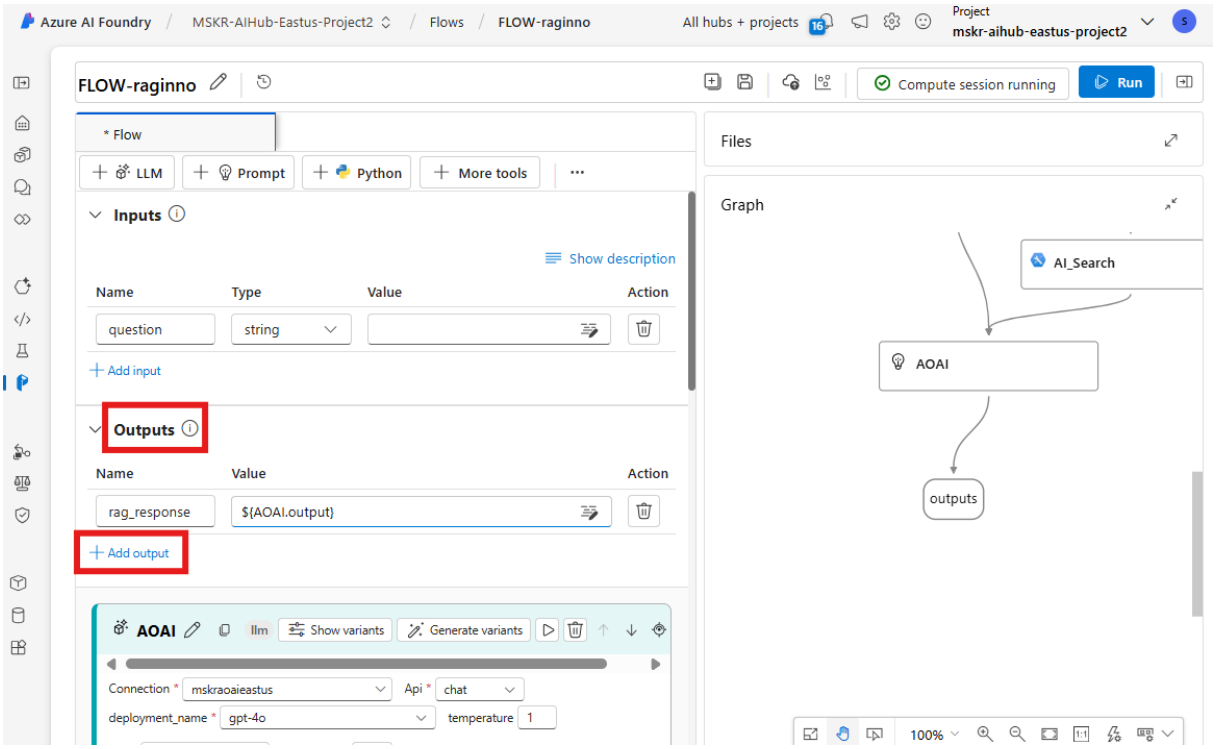


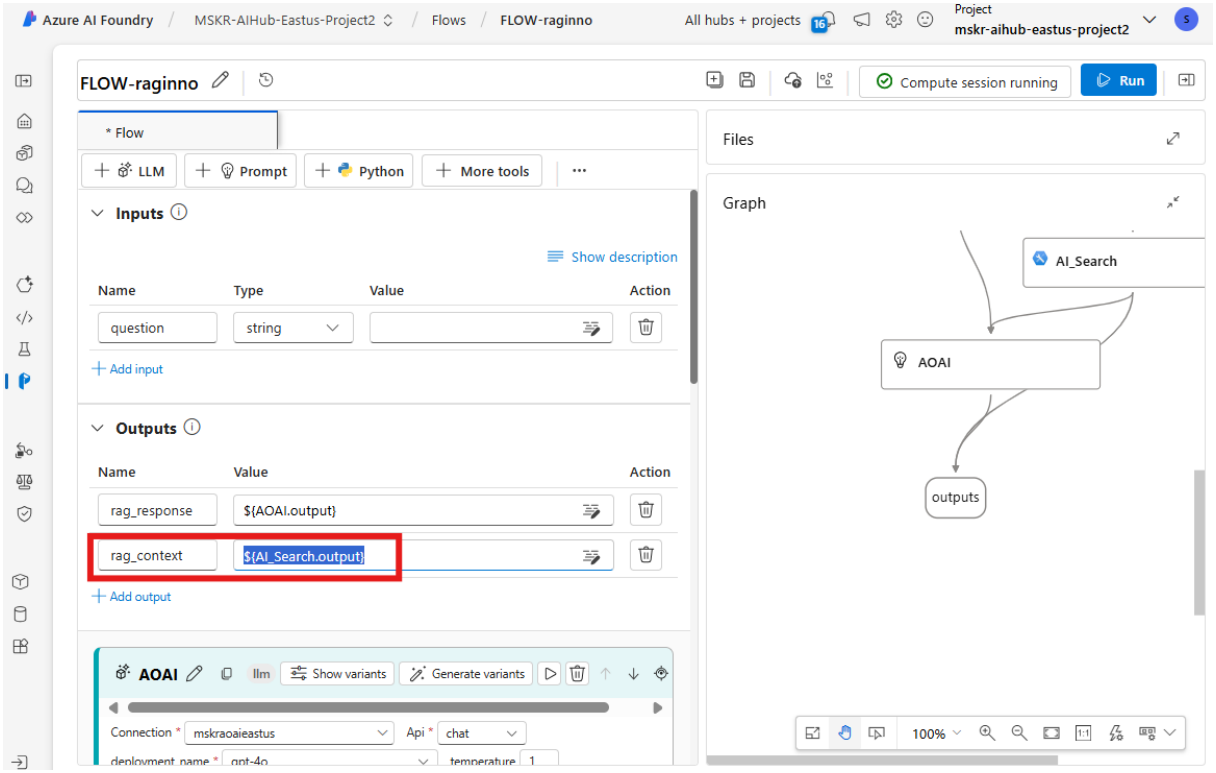
6. Flow의 Output을 수정합니다.

A. LLM Node의 Output을 RAG Response에 연결합니다.



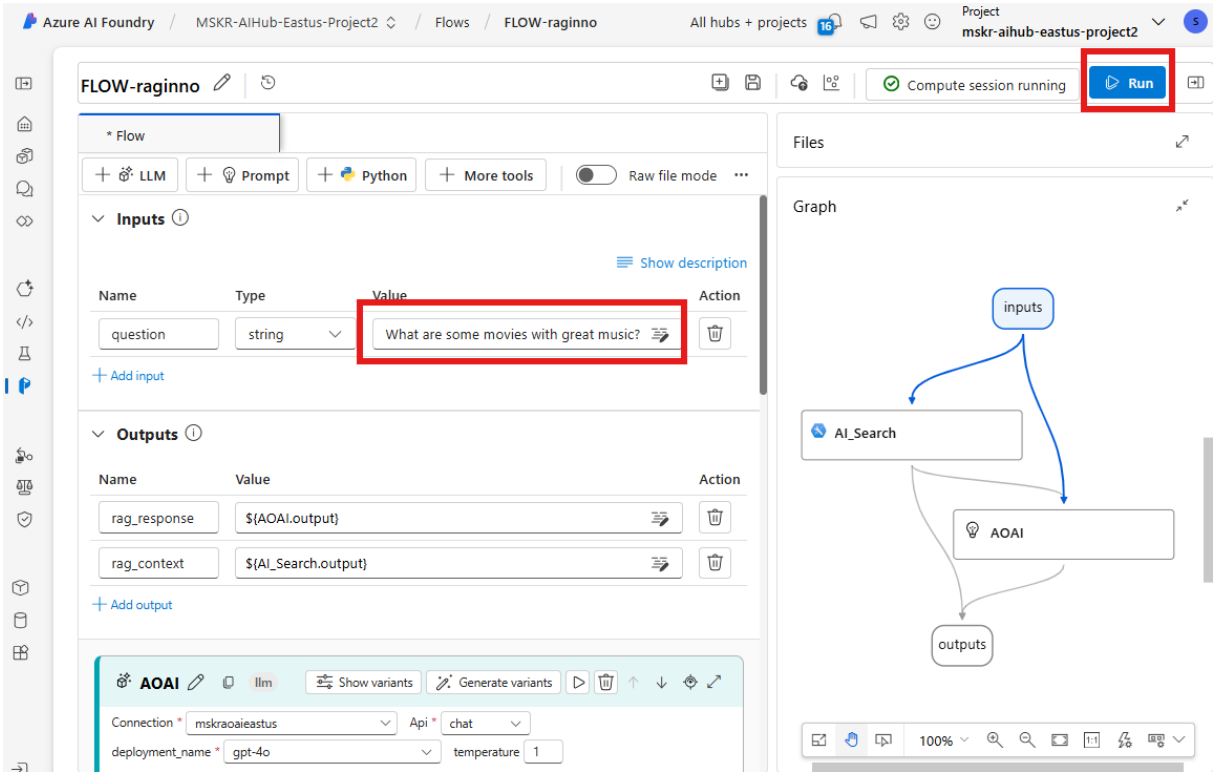
B. Add Output을 선택해 RAG Context를 추가하고, Index Lookup Node의 Output을 연결합니다.



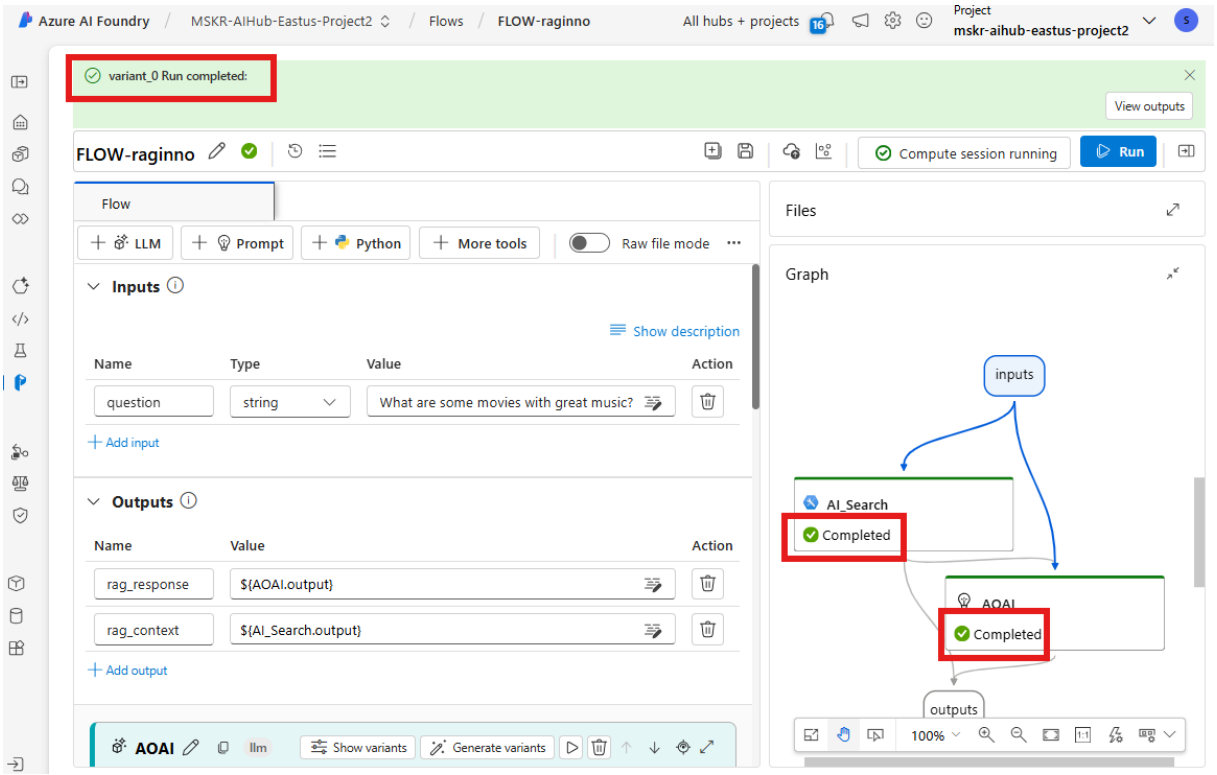


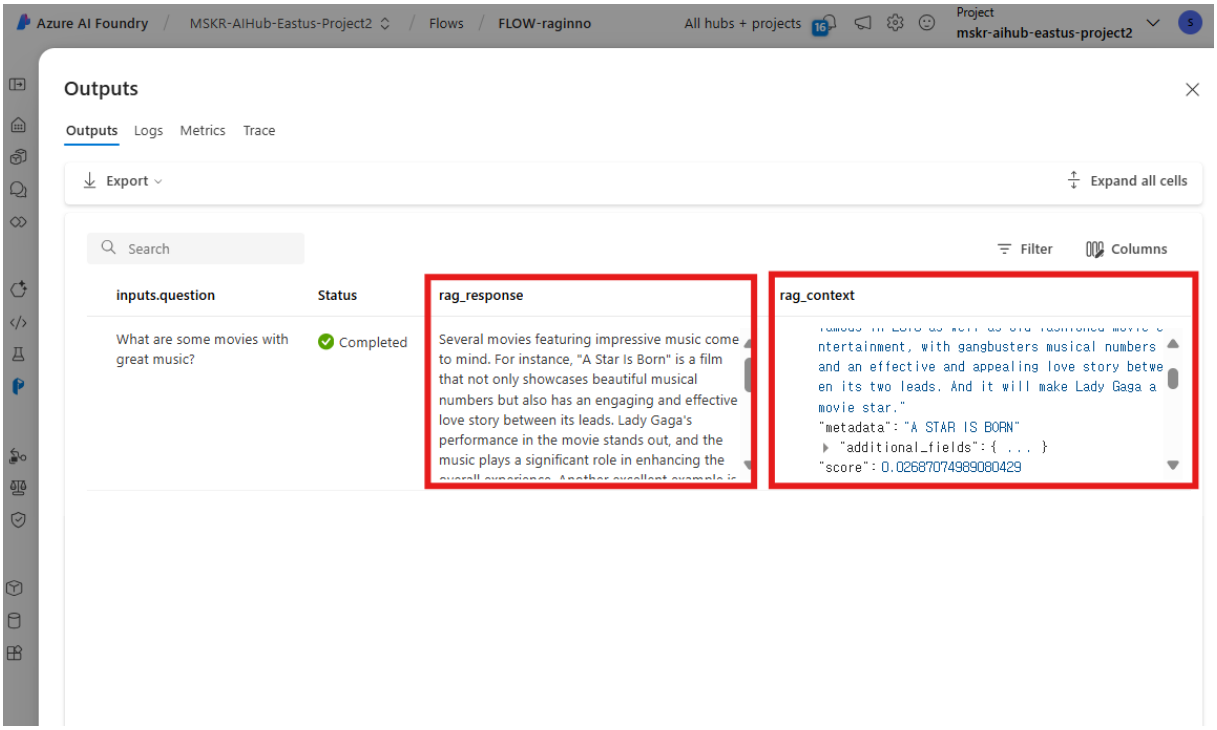
7. Flow 수정이 완료됐으면, Input 질문을 작성하고 Flow를 실행합니다.

   A. Clone한 폴더의 AIFoundry\promptflow_sample.ipynb에서 Input으로 사용 가능한 Sample Questions를 확인할 수 있습니다.



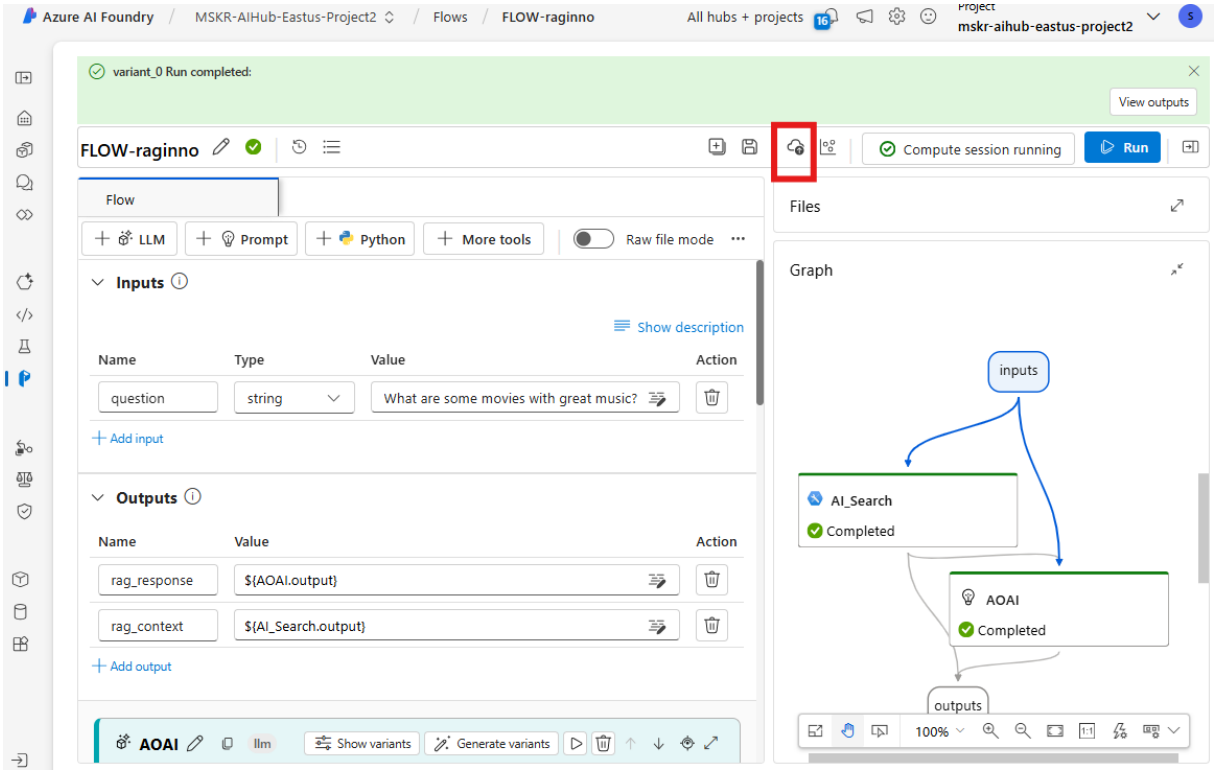
8. Flow 실행 결과를 확인합니다.



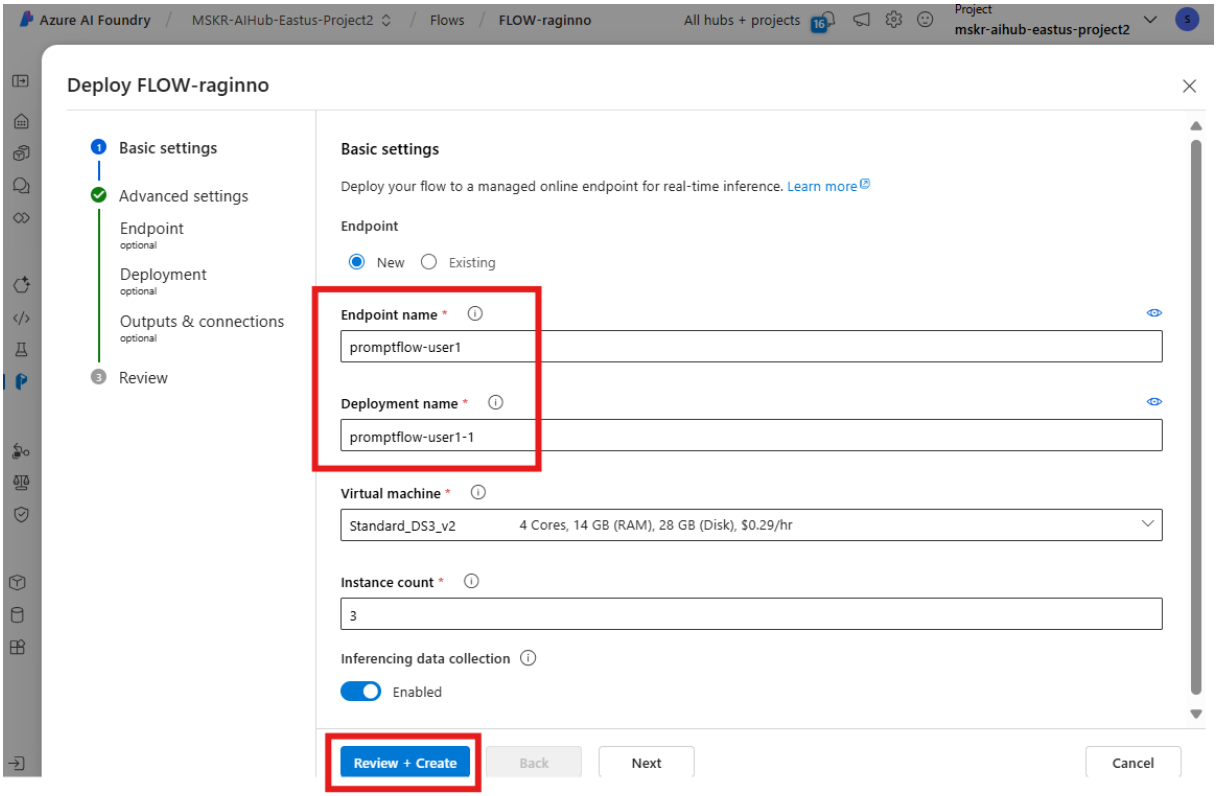


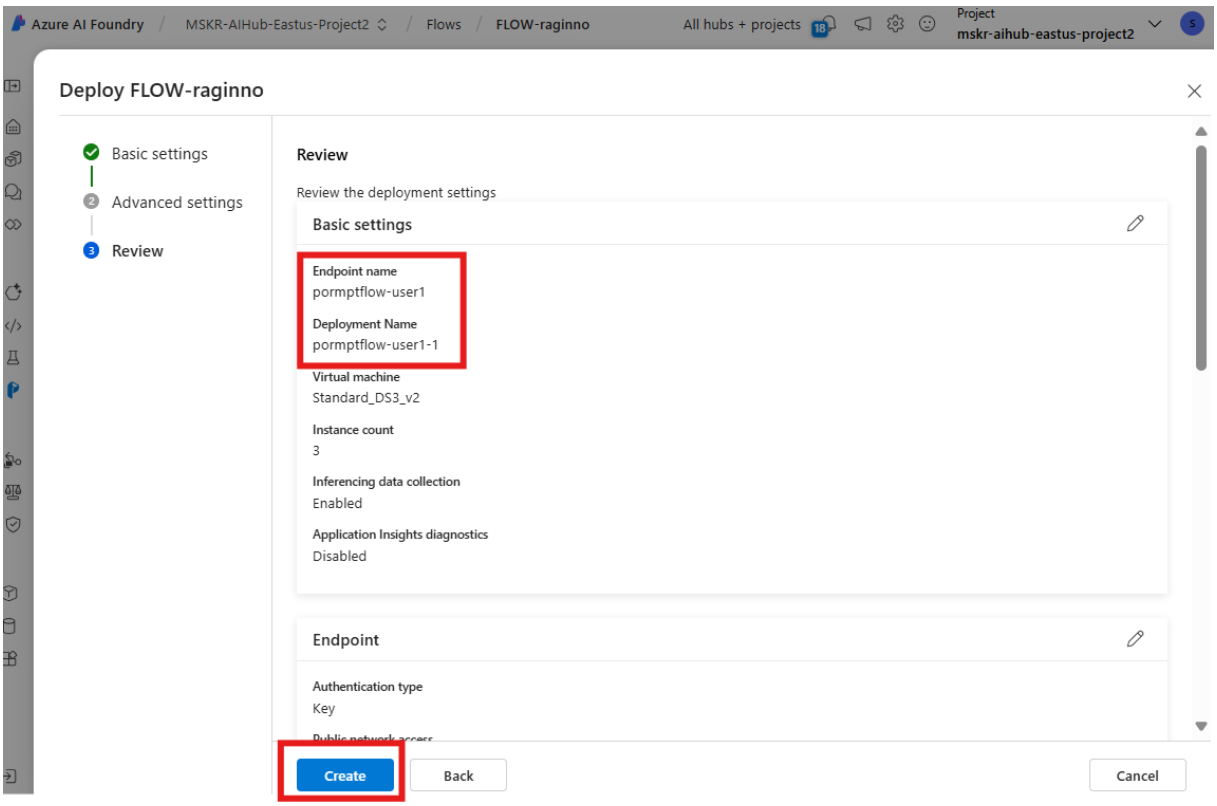
Call Deployed Endpoint

1. Flow 화면에서 Deploy를 선택합니다.

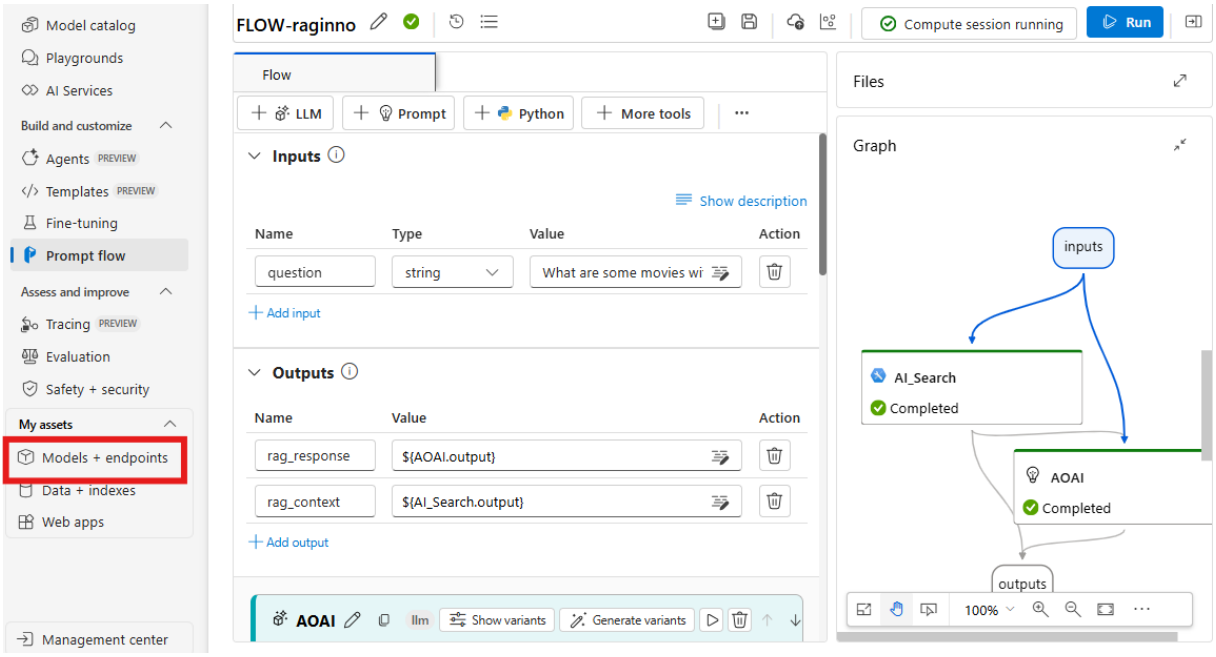


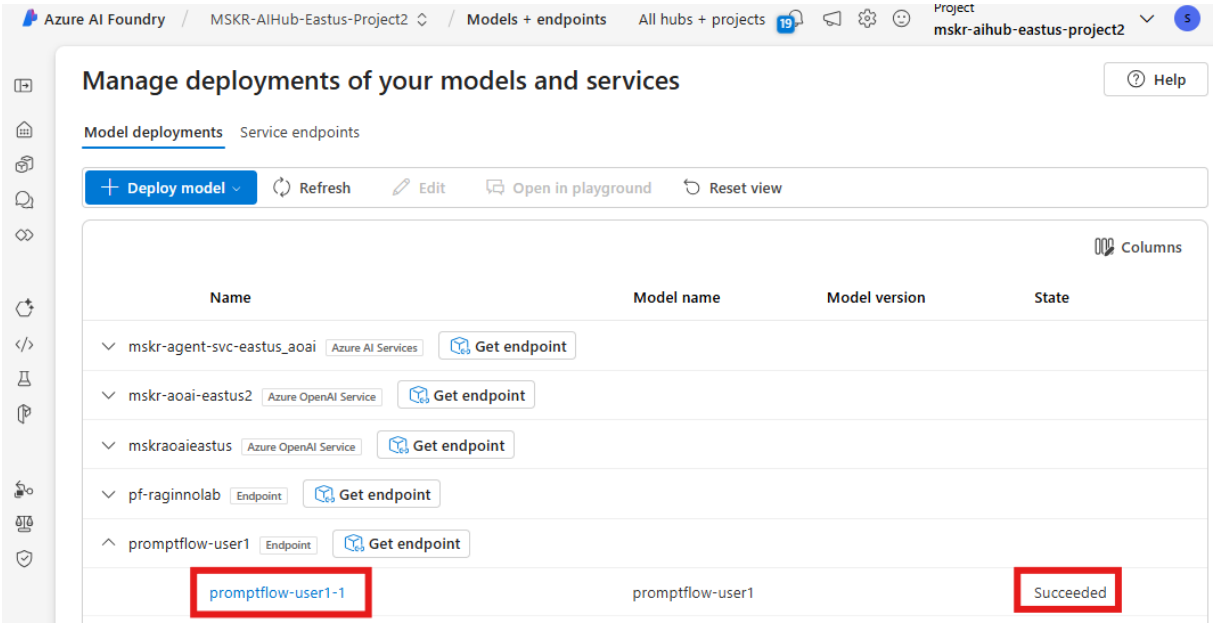
2. Custom Suffix를 포함한 이름으로 Endpoint와 Deployment명을 정하고 Create합니다.

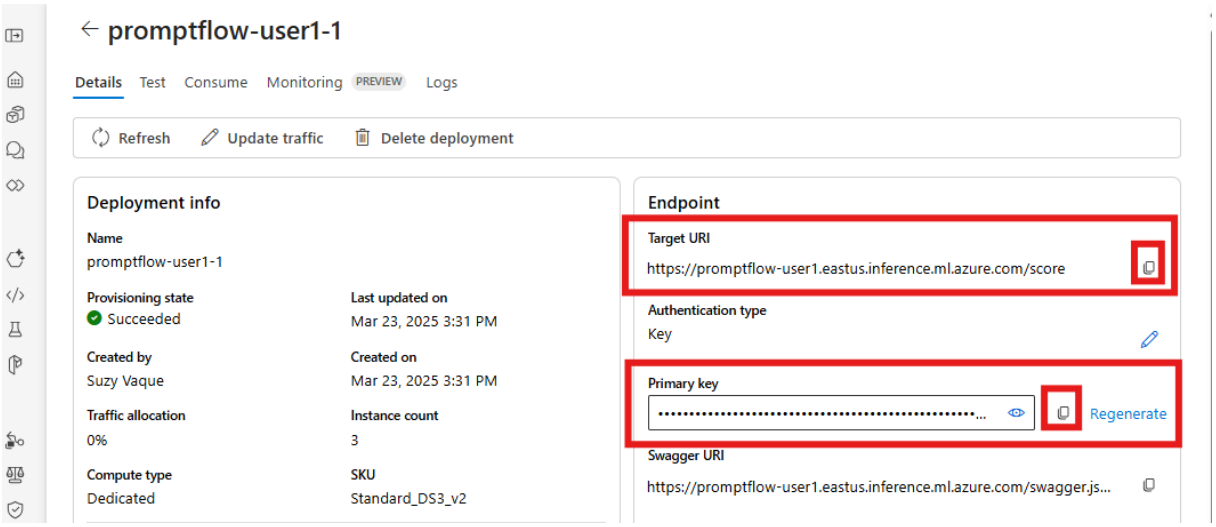




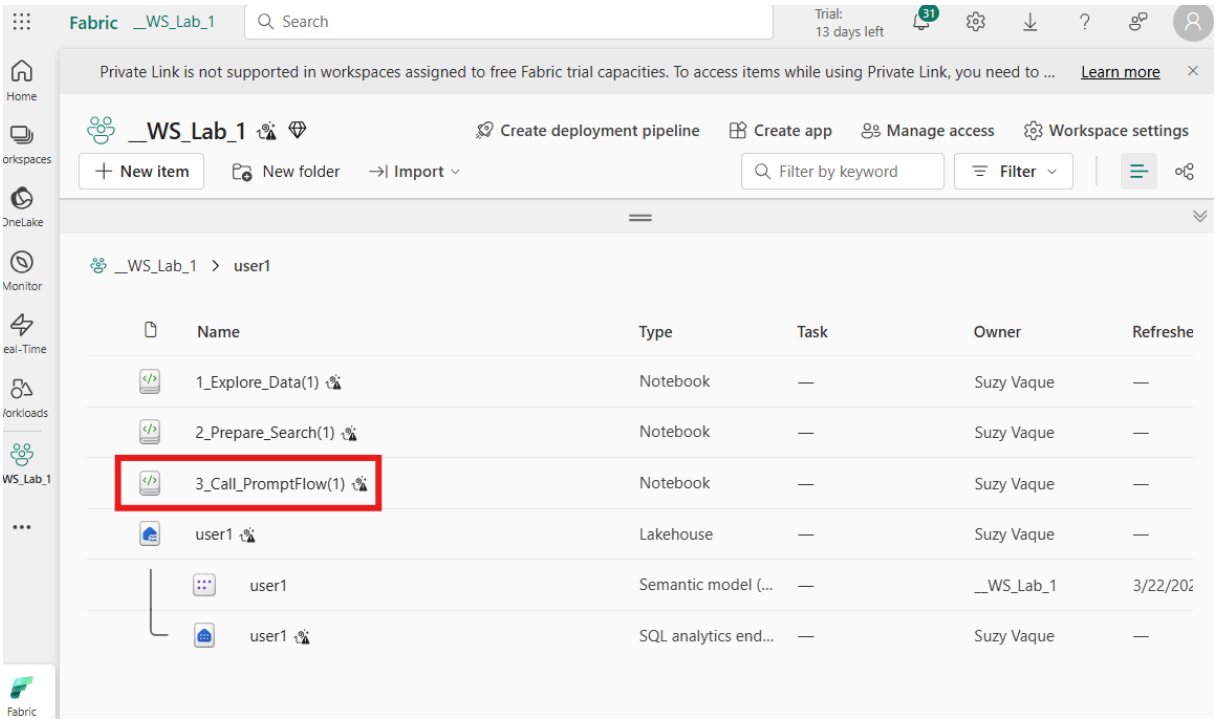
3. 배포가 완료되면 Endpoint와 Key 값을 확인합니다.



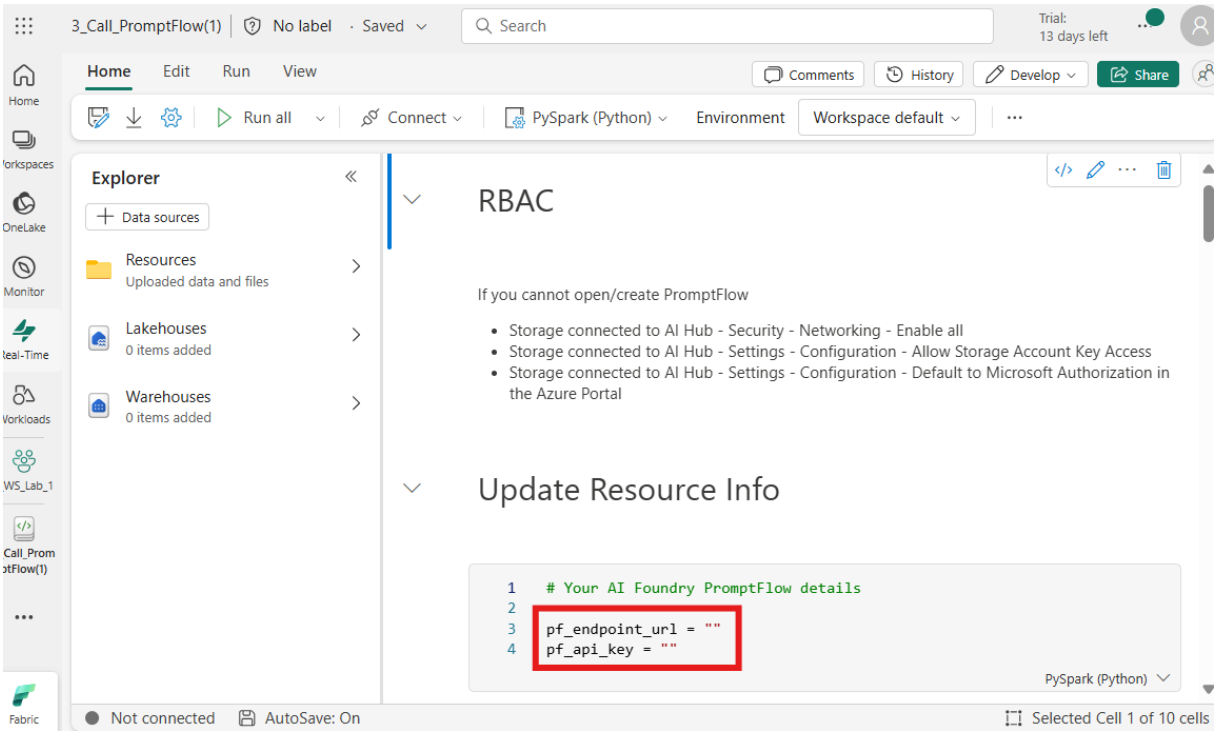




4. Fabric Portal의 3_Call_PromptFlow 노트북으로 이동합니다.



5. Flow의 Endpoint와 Key 값을 업데이트합니다.



6. Run PromptFlow 항목의 셀들을 실행하면, Fabric Notebook 상에서 REST API 방식으로 Flow를 호출하고 Lakehouse의 데이터를 기반으로 한 RAG 답변을 얻을 수 있습니다.