**Databricks AWS Gen AI workshop**

Gen AI Requirement

Delta table : Document

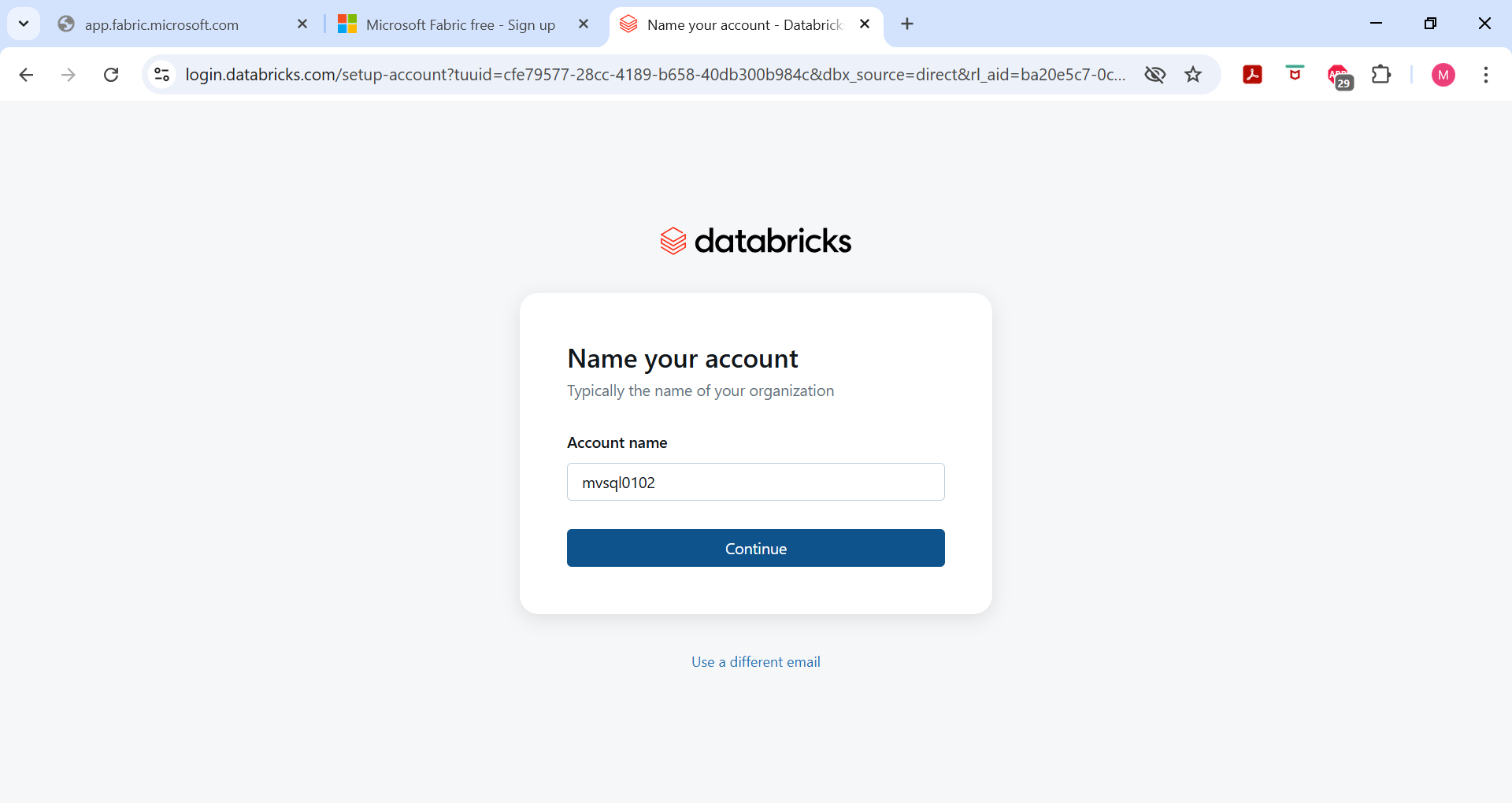
Vector search

We calculate new embeddings and feed our vector database.

Langchain

Databricks account

AWS account



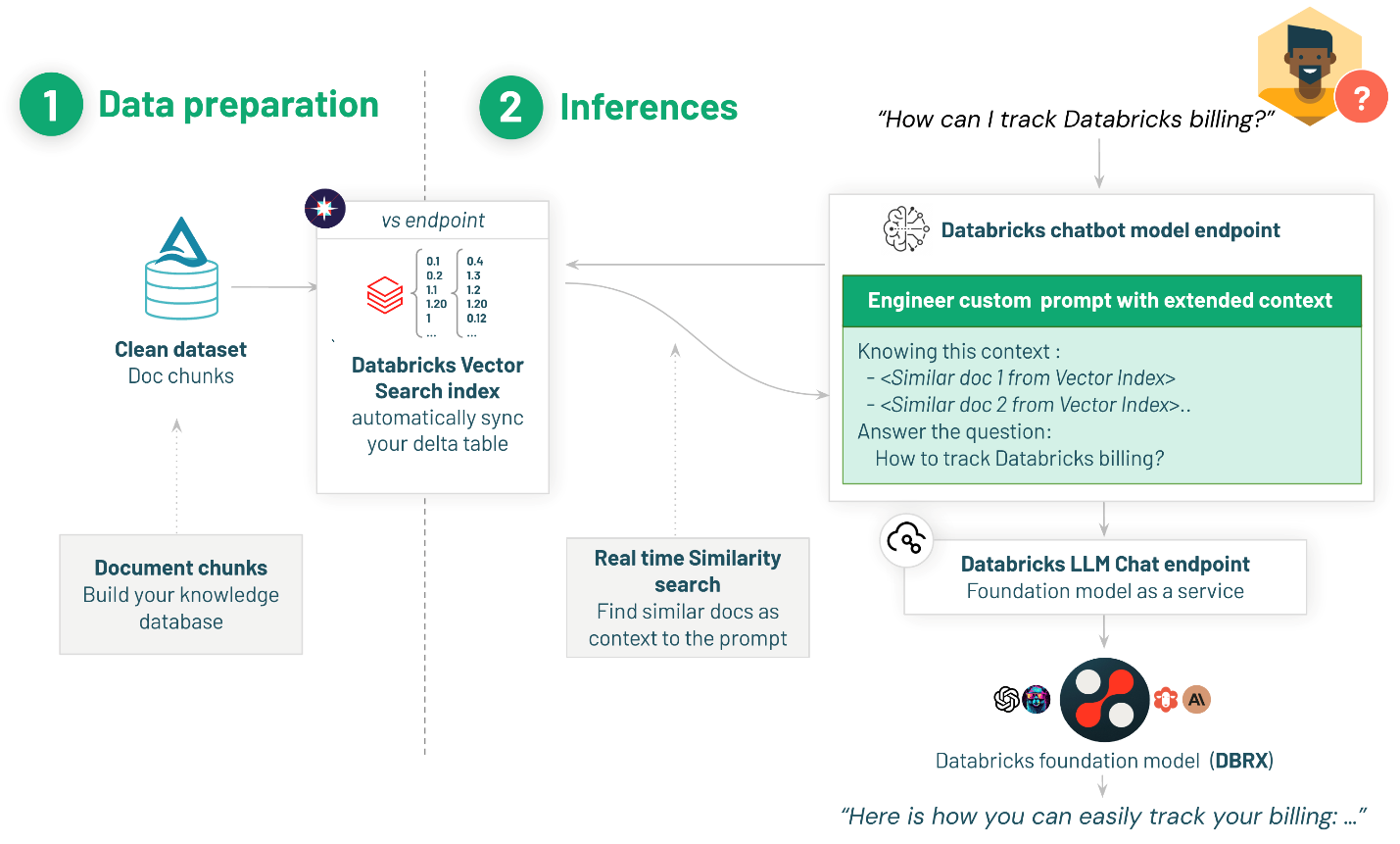
AccessKeyId

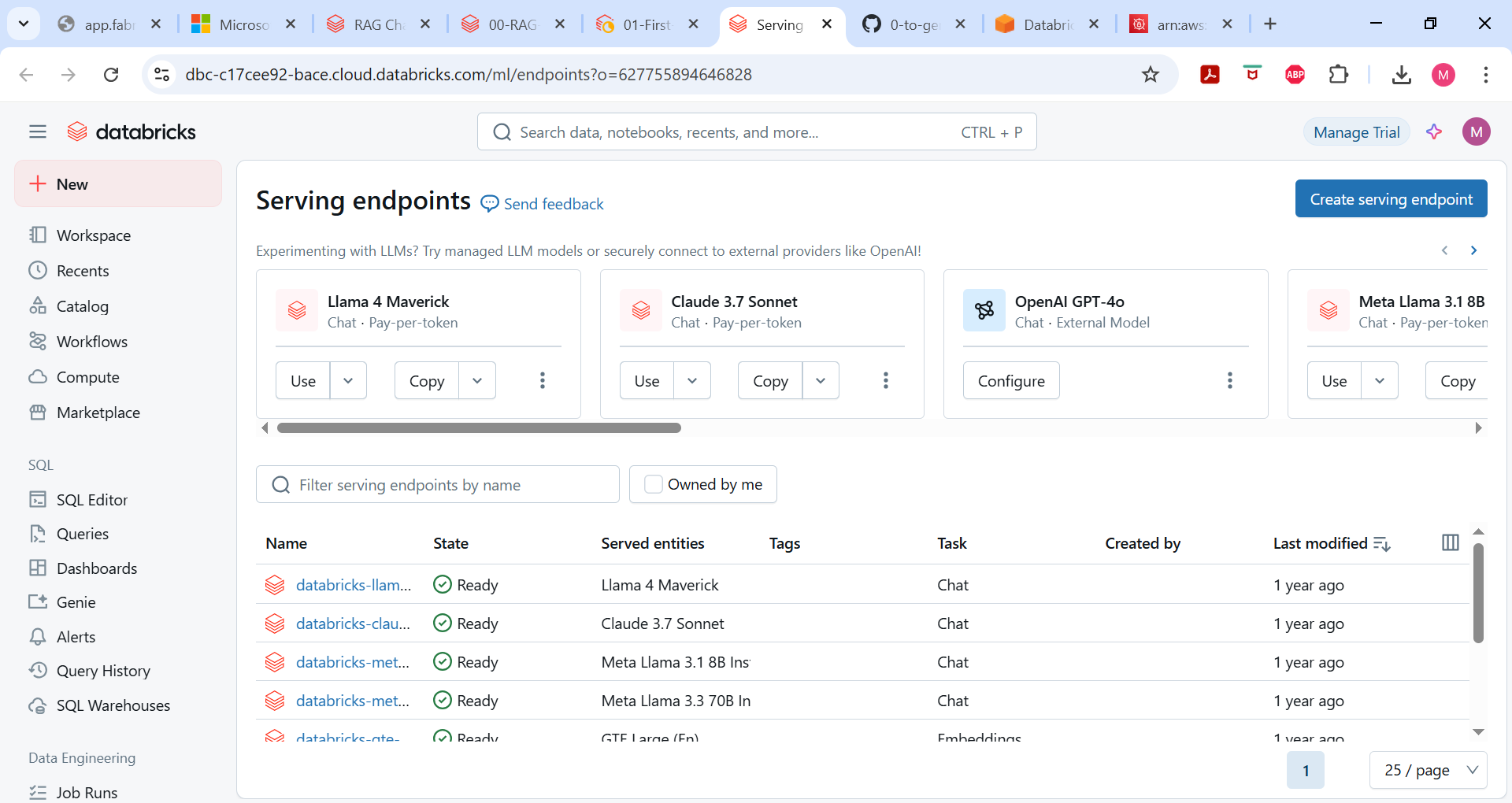
AKIAUEJCDRKZIZ6NC3V3

SecretAccessKey

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

Retrieval Augmented Generation

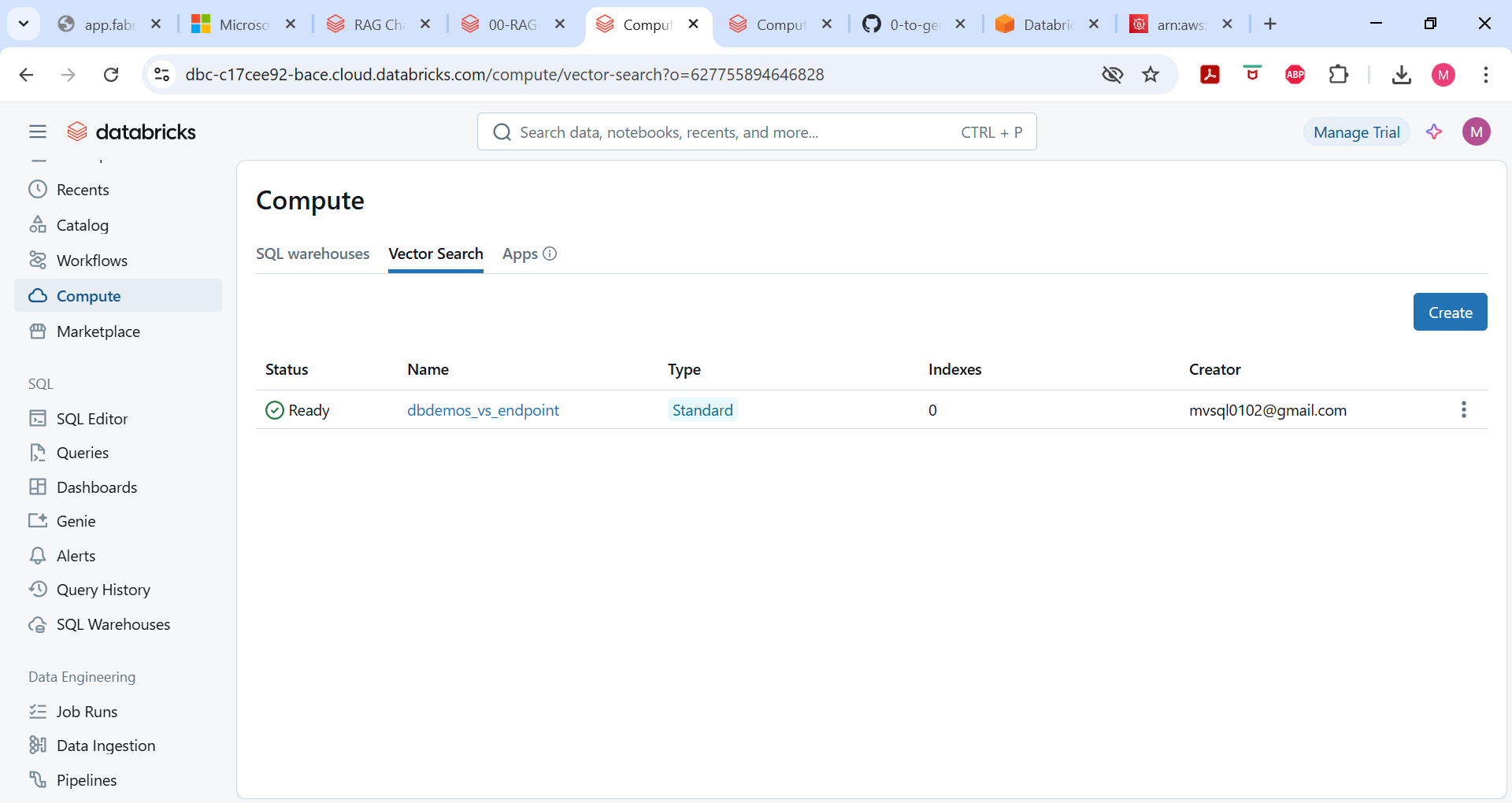


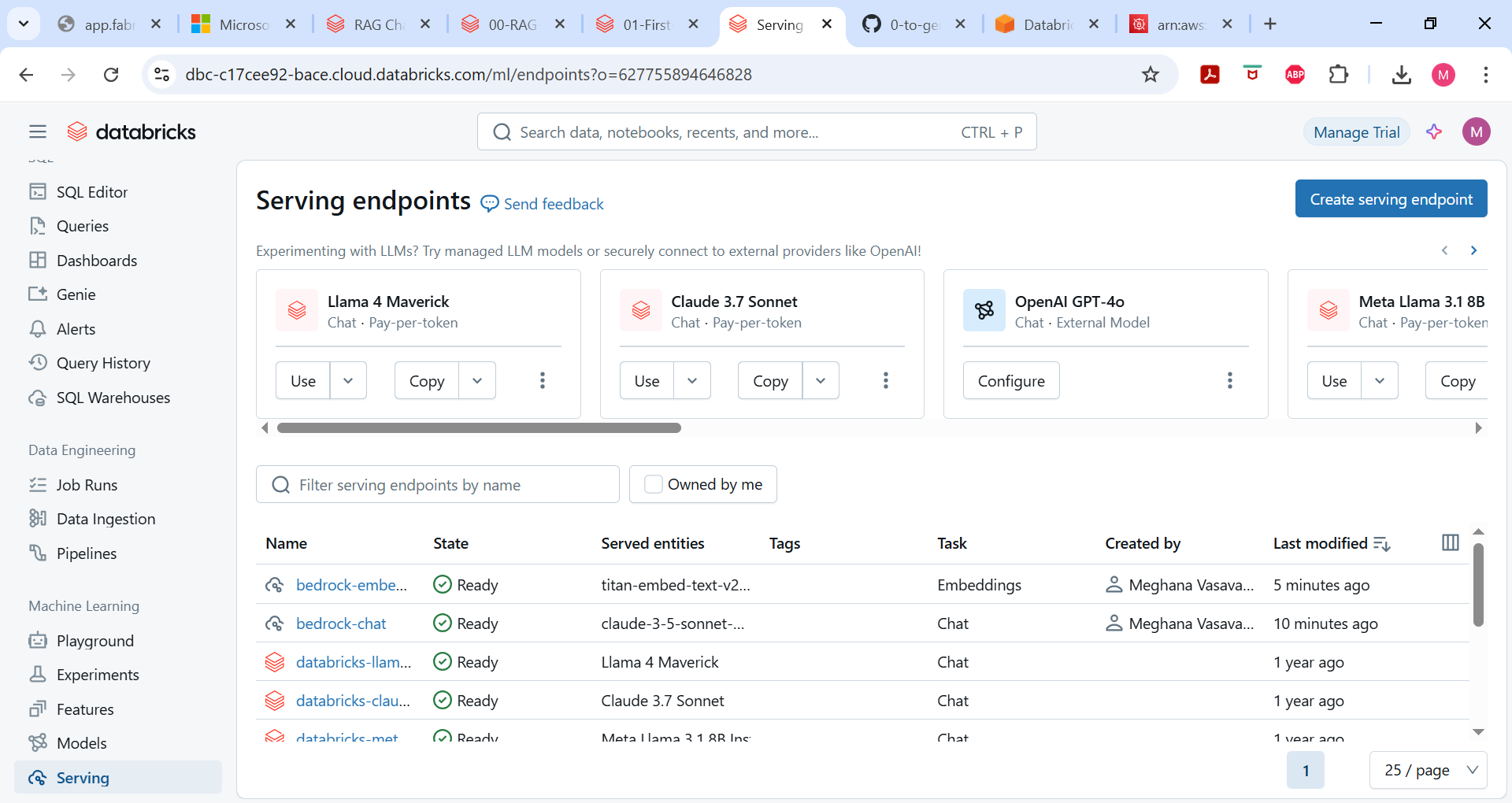


Embedding models are used whenever you ask a question:

Create 2 serving endpoints

1. Bedrock-chat using Anthropic Claude
2. Bedrock-embeddings: using Amazon Titan





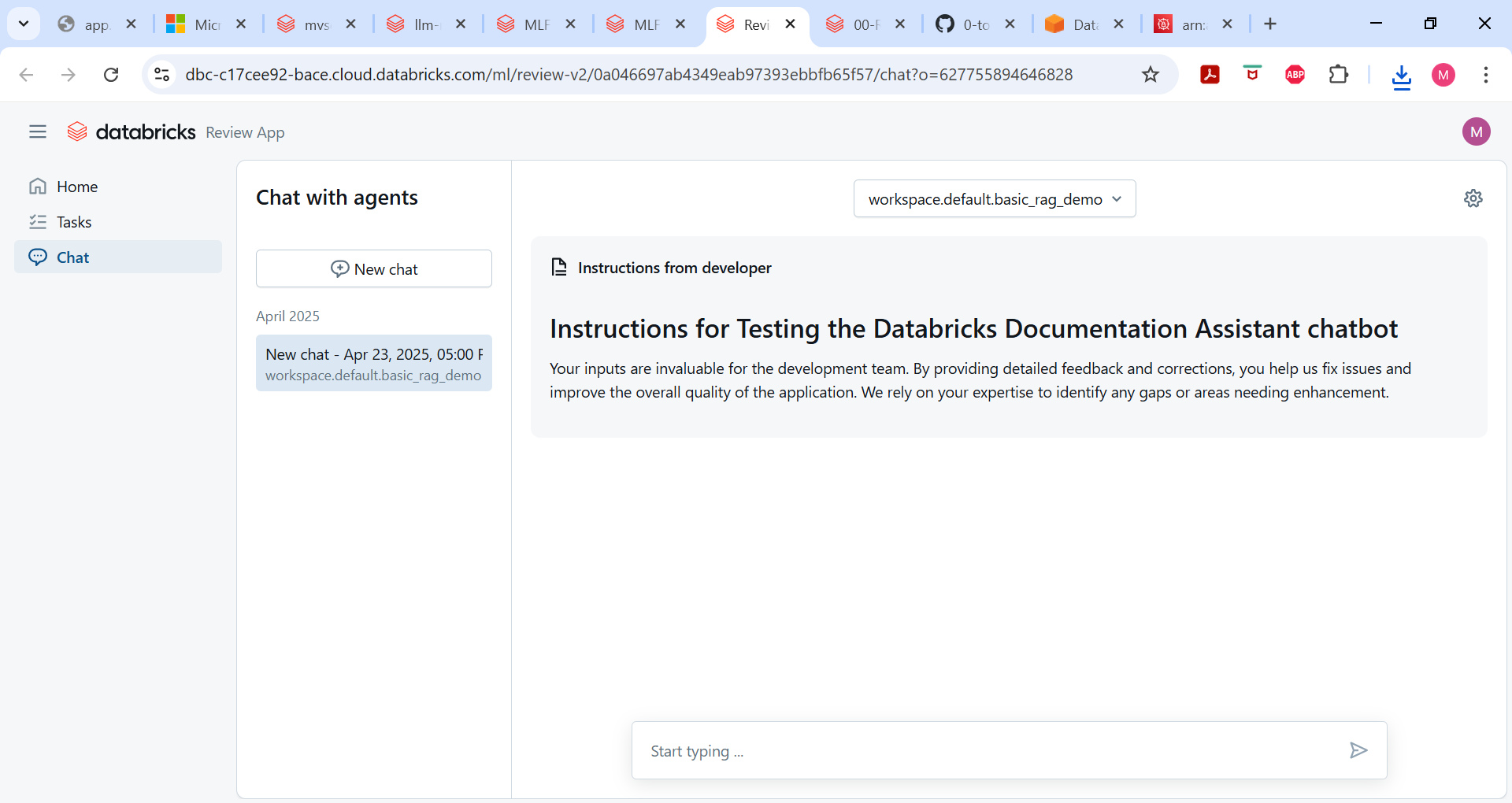
**databricks-gte-large-en : for embeddings**

**databricks-claude-3-7-sonnet for chat**

[**https://catalog.us-east-1.prod.workshops.aws/event/dashboard/en-US**](https://catalog.us-east-1.prod.workshops.aws/event/dashboard/en-US)

[**https://github.com/ioannisatdatabricks/0-to-genai-with-databricks-on-aws/blob/main/Setup%20Instructions%20for%20Databricks%20Express.md**](https://github.com/ioannisatdatabricks/0-to-genai-with-databricks-on-aws/blob/main/Setup%20Instructions%20for%20Databricks%20Express.md)

[**https://us-west-2.console.aws.amazon.com/secretsmanager/secret?region=us-west-2&name=arn:aws:secretsmanager:us-west-2:284076247730:secret:MyUserAccessKey-lmUnqx**](https://us-west-2.console.aws.amazon.com/secretsmanager/secret?region=us-west-2&name=arn:aws:secretsmanager:us-west-2:284076247730:secret:MyUserAccessKey-lmUnqx)



**https://github.com/ioannisatdatabricks/0-to-genai-with-databricks-on-aws/blob/main/Setup%20Instructions%20for%20Databricks%20Express.md**