



# Microsoft Federal Developer Summit

## Building AI Solutions

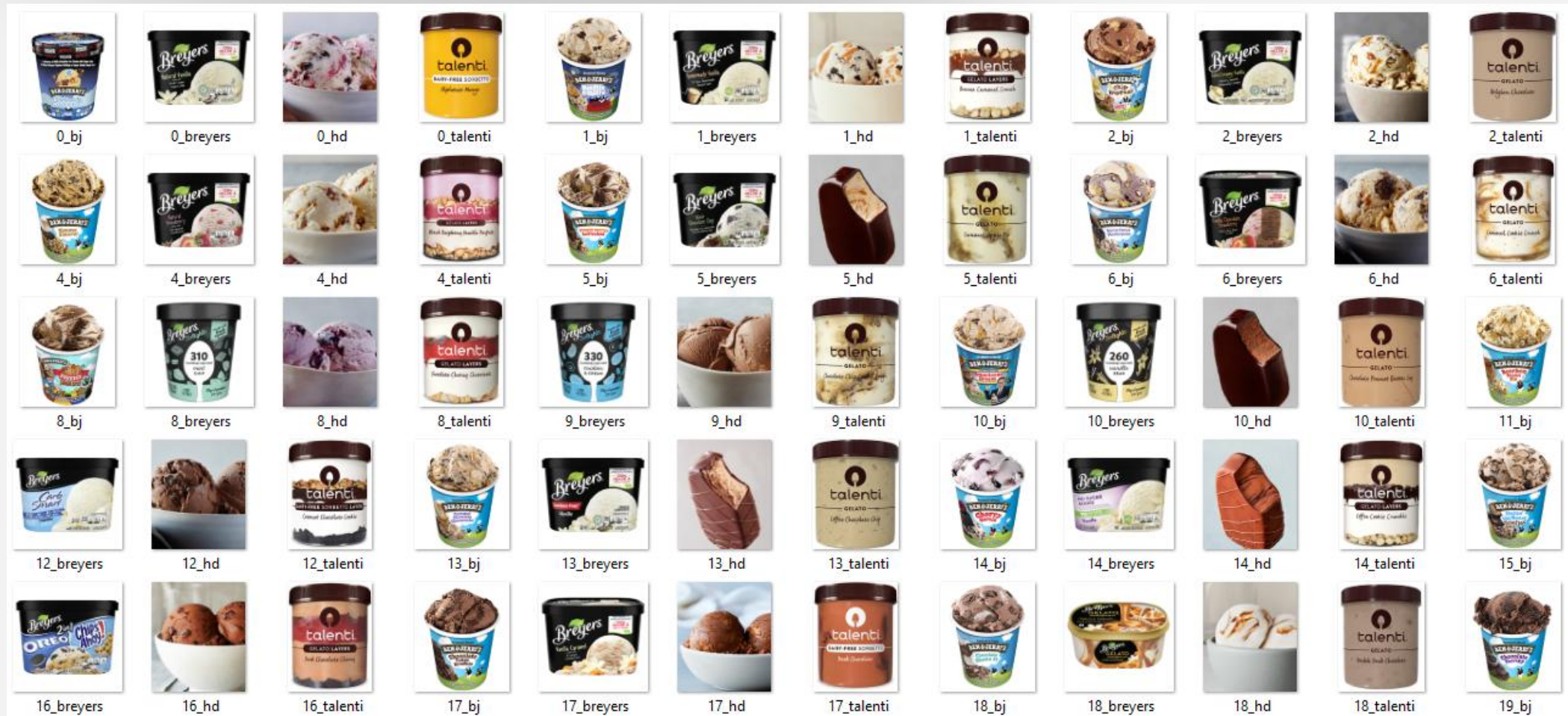
### The Semantics of AI

*Demystifying Concepts and Terminology*

Dan Biscup – Cloud Solution Architect  
Rich Posada – Cloud Solution Architect



# Ice Cream Dataset – 239 Images



# Ice Cream Bot



Microsoft Federal Developer Summit  
Building AI Solutions

## Ice Cream Bot

Settings



What flavors of ice cream are there?



Based on the sources provided, the available flavors of ice cream are:

- BEN & JERRY'S Ice Cream Sammie (no specific flavor mentioned)
- Häagen-Dazs CRISPY TR BELGIAN CHOCOLATE COCONUT & ICE
- Häagen-Dazs IS PY TR STECHOCOLATE & LEMON & R. ICE CR
- Häagen-Dazs CRISPY TR BELGIAN C COFFEE & ICE C

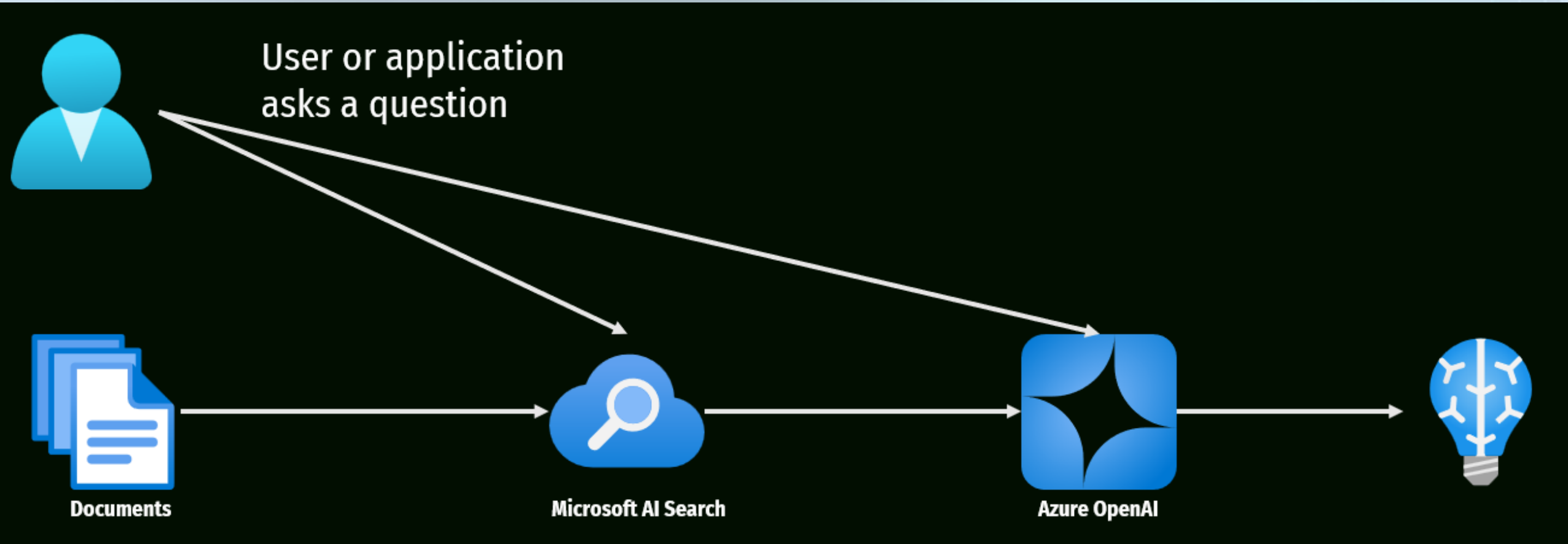
However, this is not an exhaustive list of all ice cream flavors available.

Start New Session

Ask your ice cream question here...



# Retrieval-Augmented Generation (RAG) Pattern





# Ice Cream Bot Architecture



## **Azure App Service**

Docker Container  
Python Streamlit  
Calls Azure AI Search  
Call Azure OpenAI



## **Azure AI Search**

Ice Cream Vector Index



## **Azure AI Services**

Multimodal Embedding  
Vectorize Images



## **Azure OpenAI Service**

LLM  
Azure Infrastructure  
Azure Security



Microsoft Federal Developer Summit  
Building AI Solutions

# Large Language Model (LLM) and Tokens

- LLM is type of AI model trained on massive amounts of data that can pre generate coherent
- LLMs process nu
- A token is about 4
- Mapped to a num
- Every LLM has a process for a sing
- LLM pricing is tie

TokensCharacters

1770

TokensCharacters

1770

[61443, 290, 9029, 328, 22384, 558, 1993, 421, 12432, 122861, 198, 7633, 19354, 29338, 15, 198, 168747]

TextToken IDs

TextToken IDs

# The Art and Science of Prompt Engineering

Prompt engineering is the process of designing and refining prompts to guide artificial intelligence (AI) models to produce the desired responses.

New Paradigm for Full-Stack Developers:

1. Shift from deterministic to probabilistic
2. Understanding model behavior
3. Evolving best practices
4. Ensuring ethical AI usage
5. Token and cost management



Azure AI Studio



Microsoft Federal Developer Summit  
Building AI Solutions

# Prompt Engineering Key Concepts

1. Role Assignment
2. Instruction
3. Few-Shot Prompting
4. Zero-Shot Prompting
5. Iterative Refinement
6. Chain of Thought
7. Combining Prompts

## 7. Combining Prompts

Breaks large tasks into manageable steps.

Example:

1. Generate an outline for a blog post.
2. Expand the outline into full sections.





# Azure AI Search Terminology

- Embeddings
- Vectors
  - Image 1024
  - Text 1536
- Chunking
- Skillset
- Index
- Indexer
- Semantic Ranker
- BM25
- RBAC
  - Lead with security!



# Why AI Search?

- Common use cases
  - ChatBot
  - Data pipeline
  - Model does not know your data
  - All about RAG Retrieval-Augmented Generation
- Many use open source models
  - Stich your own solution together
  - Where do you store vectors?
  - How do you embed your content?
- AI Search does it all

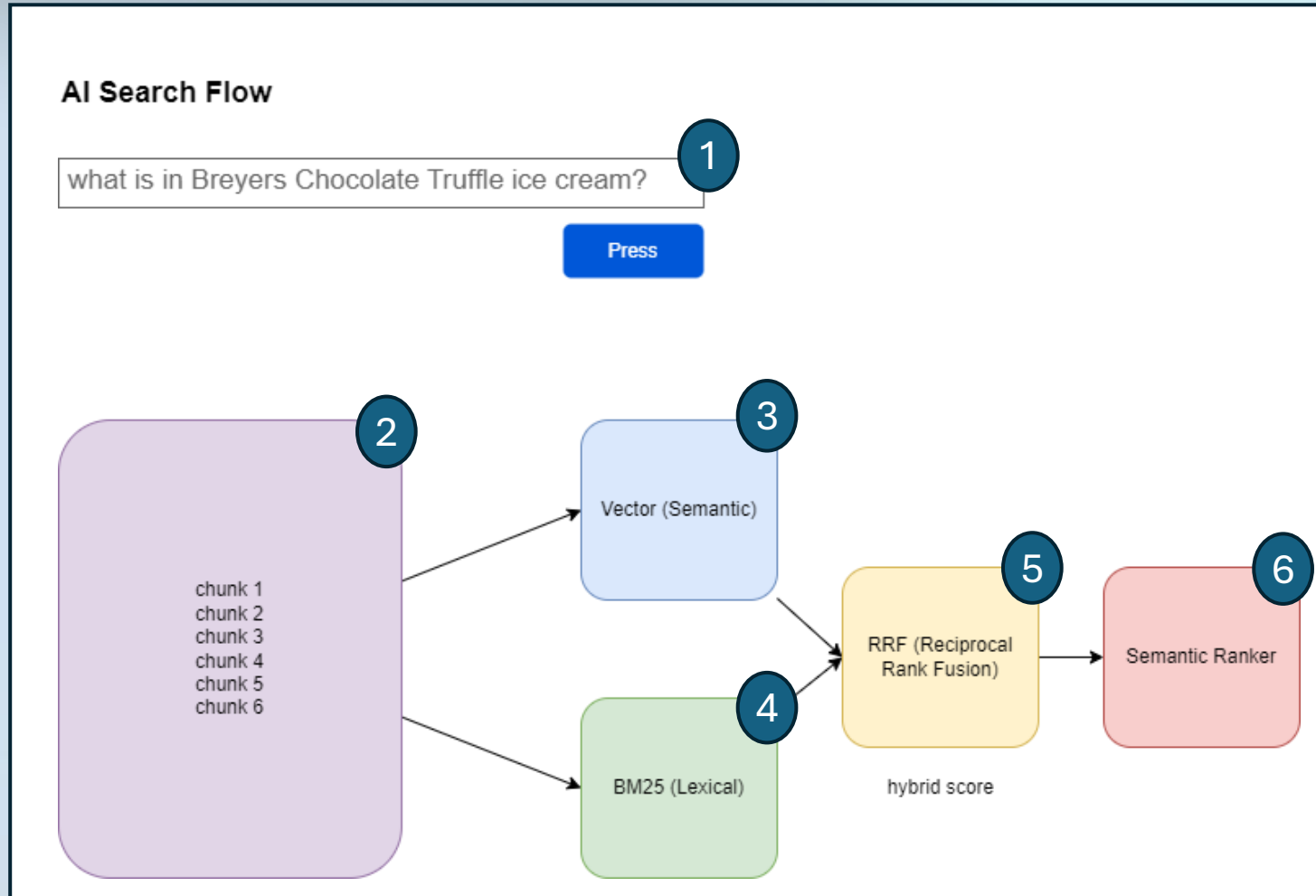


# Parts of AI Search demo

- Overview
  - Properties
  - Usage
  - Monitoring
- RBAC
- Indexer
- Index
- skillset



# How does it work? demo



# Feedback

Do you want us to follow up after the event? Do you have feedback?



<https://aka.ms/summit/feedback>



Microsoft Federal Developer Summit  
Building AI Solutions





# Ice Cream Time!



# THANK YOU!

## CONTACT

NAME |

EMAIL |

PHONE |



Microsoft Federal Developer Summit  
Building AI Solutions