ESDC Curam Assist Me

Introduction to Assist Me

What is Assist Me ?

Assist Me is a **virtual assistant** for OAS representatives, delivered through a **chatbot** interface. When prompted with a question on guidelines, policies, or procedures pertaining to the OAS benefit and/or *Cúram (Pensions)* platform, it will provide **relevant information** to **assist representatives** in the **case review process**.

Where is the information in Assist Me coming from?

The information within Assist Me comes **from OAS on BDM knowledge articles** that were developed in support of **OAS on BDM Release 1 training**.

Why was Assist Me created?

The goal is to significantly **reduce the amount of time** agents spend **searching** the knowledge management database for information, allowing them to **focus on higher-value tasks** and provide a **more efficient service to Canadians**.

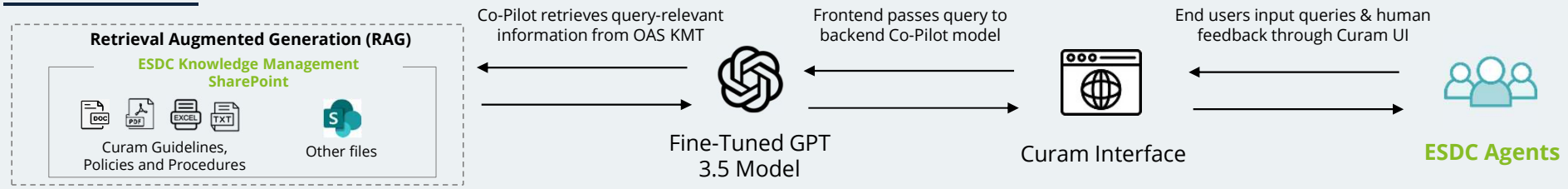
Overview of Assist Me Solution

Background: The OAS project anticipated transition to a new system will present challenges to all end user based on R1 lessons learned. To enhance the Agent experience + manage onboarding/training requirements, an intuitive knowledge retrieval solution is required.

Assist Me is a virtual assistant for OAS Agents, delivered through a chatbot interface. When prompted with questions on guidelines, policies, or procedures pertaining to the OAS benefit and/or Curam platform, Assist Me provides relevant information to assist agents in the case review process.

The goal is to significantly reduce the amount of time Agents spend searching the Knowledge Management database for information, allowing them to focus on higher-value tasks and provide a more efficient service to Canadians.

How it Works:



AI Components: Virtual assistant interface built using React JS, automated data ingestion pipeline for SharePoint built using Azure Logic Apps, integration layer for Curam built using JS, GenAI model from OpenAI (GPT 3.5 turbo)

Azure Services: Logic Apps, Blob Storage, Cognitive Service, AI Search, App Service, Cosmos DB, Azure OpenAI

End-user Journey

The journey of integrating Assist Me into your day-to-day work experience may not come naturally, but the illustration below will assist you in making the transition seamless.

Interacting with the Chatbot

Begin by opening Assist Me in *Cúram (Pensions*). Utilize the built-in chatbot feature by asking *Assist Me* specific questions. This interactive feature is designed to simulate a conversation with a “virtual knowledgeable agent”.

Evaluating Reference (Citation) Sources

After receiving a response from *Assist Me*, review the references and sources it used to generate the answer. By evaluating these sources, you can assess how well the response aligns with what you are looking for.

Refining the Answer with Follow-Up Prompts

If the initial response requires further refinement, use follow-up prompts to fine-tune the answer. *Assist Me* is designed to adapt and refine its responses based on these follow-up interactions, ensuring a high degree of customization and precision.

Provide Feedback

Once you have finished using Assist Me, provide feedback at the end of each interaction. This helps identify areas where the tool can be enhanced to better meet your needs.

High-Level Solution Architecture on Azure

The following diagram shows a proposed high-level generative AI architecture for ESDC Co-Pilot on Azure Service with GPT-3.5 or other Large Language Models (LLMs).

Key Capabilities

Robust and Scalable

Develop a robust GenAI solution on Azure with a GPT-4 or open source LLMs model, and other add-on services, including Azure AI Search, Open AI Embeddings, auto-scalable Azure Storage, etc.

Open Source LLM Models Adoption and Fine-Tune

use Azure ML Studio and Azure Function to build, fine-tune, and deploy open source LLM models from Hugging Face, LangChain, etc.

OAS Data Enrichment

Provide AI data enrichment in backend services, interact with OAS data through Azure OpenAI GPT-4 API or Azure Function, generate, and process human-like summarization, notes, text insights, etc.

Indexing & Searching

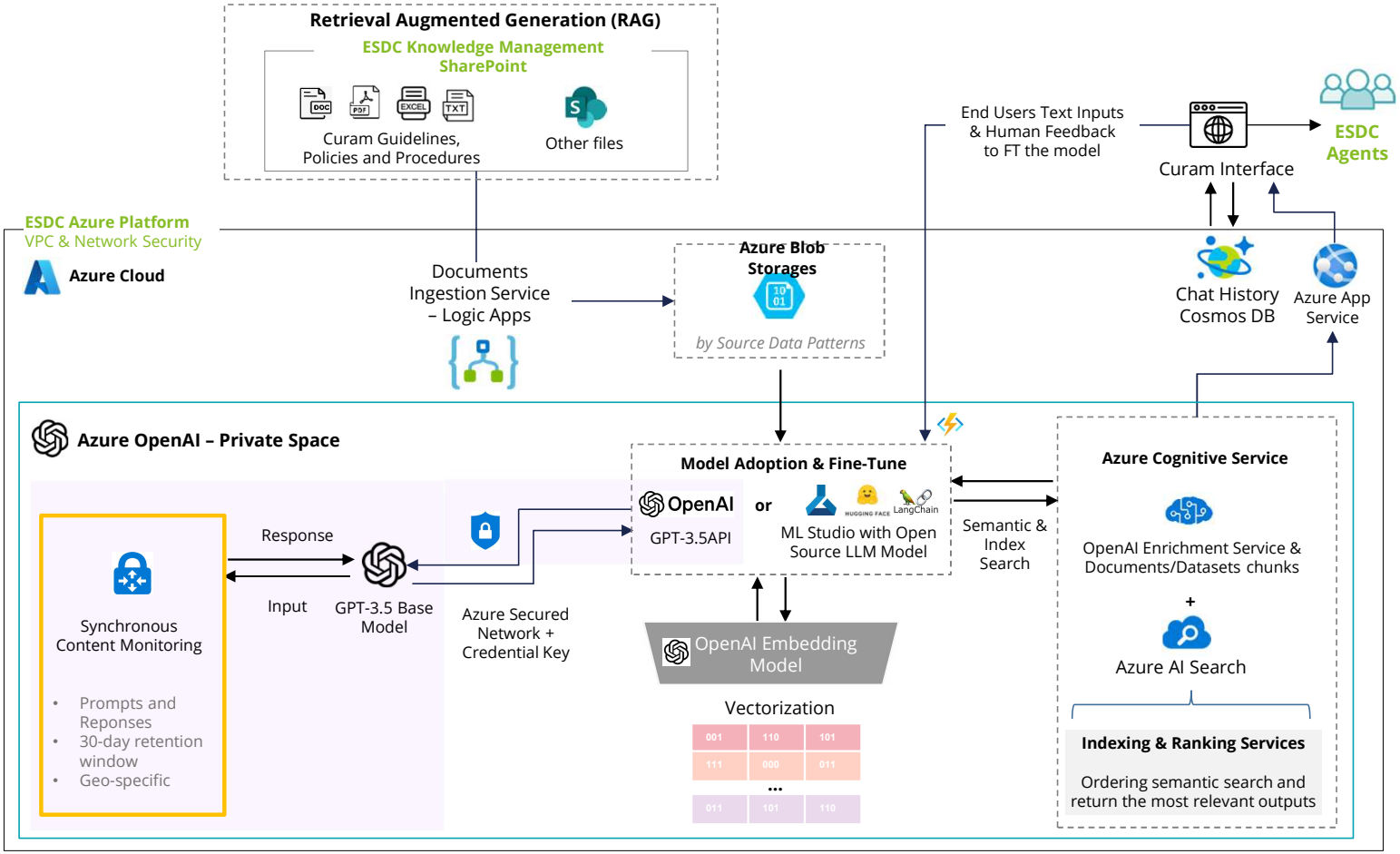
Use Cognitive AI Search supports semantic search and use an indexer for driving the process end-to-end based on the vector fields to identify relevance with confidence scores

Interactive User Interface

The solution uses Azure Bot Service to provide conversational experiences and integrated development with OAS systems, e.g., Curam

Security

The solution deploys on ESDC Azure OpenAI private space applied security policies, such as Key Vault, Firewall, IP whitelist, etc.



Evaluation Approach

The evaluation of Assist Me ’s responses during development is intended to rigorously assess the quality and effectiveness of the generated outputs. This evaluation is conducted through a qualitative approach, where testers from both the development team and ESDC are engaged to review and rate the responses based on specific criteria. The evaluation focuses on five key dimensions:

RELEVANCE

This criterion measures how well the response aligns with the specific query posed by the IO Agent. It ensures that the information provided is directly applicable and pertinent to the topic at hand, thereby enhancing the practical value of the response.

ACCURACY

This measure evaluates the factual correctness of the response. It is essential that the information provided is reliable and trustworthy, as inaccuracies can lead to misguided decisions and delay case processing.

COMPLETENESS

This aspect assesses whether the response comprehensively addresses all elements of the query. It is important that the response lists out all critical details, providing a full understanding of the topic.

CLARITY

The clarity of the response is evaluated to ensure that the information is presented in an understandable and coherent manner. Clear communication is key to ensuring that the insights provided can be effectively utilized in decision-making processes.

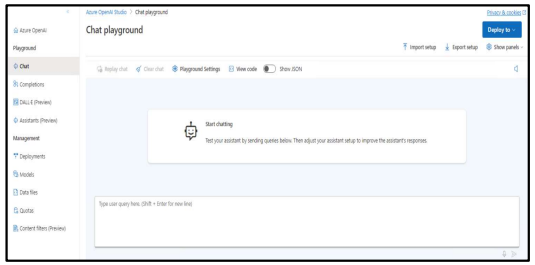
TONE

The tone of the response is reviewed to ensure it is appropriate for a professional government environment. The right tone helps in maintaining a respectful and engaging interaction, which is crucial for the user experience.

Each of these dimensions is scored on a scale from 1 to 5, with 1 indicating a poor performance in the respective category and 5 representing an excellent performance. A response must score at least a 3 in each category for the test query to be considered satisfactory. This scoring system allows for a nuanced assessment of the model's outputs, providing clear feedback for continuous improvement.

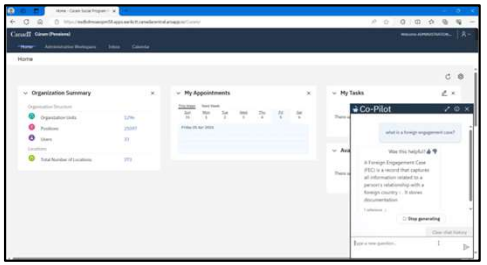
Project Timeline

Evolution of Assist Me



Iteration 1: Prototype in Azure Studio

Iteration 2: Separate Web Application



Iteration 3: Curam UI Integration

What's Next

Current State

Dev and UAT Complete: UI Integration for all R1 Knowledge Articles

Next Step (July)

Production Release and Go-Live for 10 users

Complete Go-Live (Aug)

Deployment to Prod Environment for All R1 Users