

Enhancing User Experience with RAG and GPT-4o Using Text and Visual Information from Websites.



Introduction.

Leveraging **Retrieval-Augmented Generation (RAG)**, combined with the power of GPT-4o, we can now build conversational systems over web sites that utilize both **text** and **visual information** as inputs and outputs. This approach allows businesses to provide their users—customers, employees, or general visitors—with highly informative and context-rich engagement.

Using **GPT-4o with Vision**, we ensure that when we index both textual and visual content, the system extracts contextual image descriptions in a useable format and aligns it with the webpage content. This results in more accurate and relevant conversational responses.

About the author.

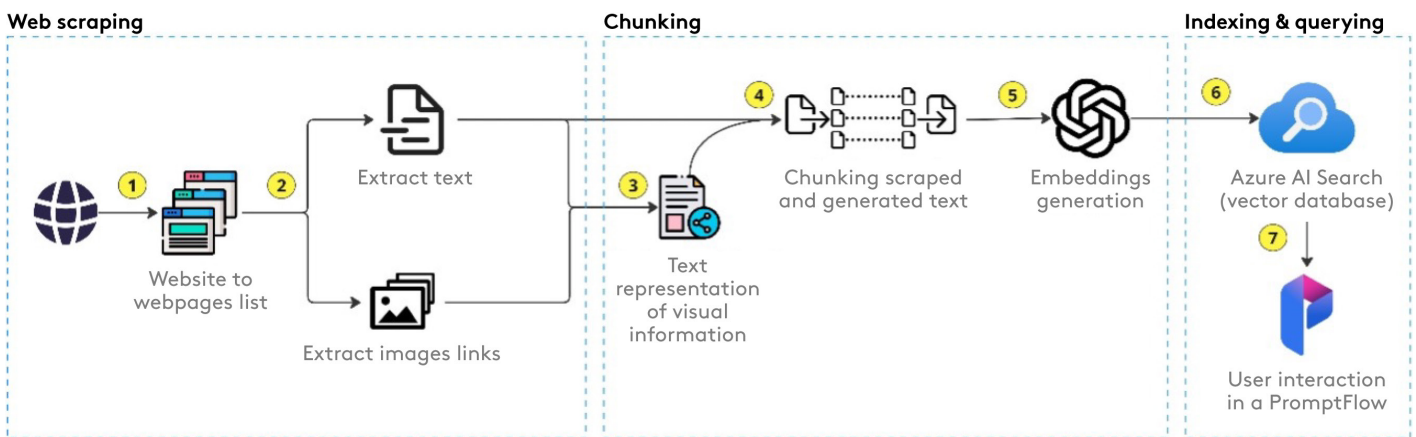


Mariia Razno

Mariia Razno is an AI and NLP engineer with 4 years of experience specializing in natural language processing and machine learning. She expertly utilizes Microsoft AI Foundry, Azure services, and other Microsoft products to develop robust, scalable AI solutions. With a strong focus on NLP technologies and large language models (LLMs), Mariia delivers high-performance systems that drive innovation and enable data-driven decision-making across industries.



How It Works: Indexing with Contextual Image Descriptions.



The process begins by converting a website into a list of webpages (1) and scraping **both text and image** links from each one (2). The image links and webpage text are then passed to **GPT-4o**, which generates text representations of the visual information (3). These descriptions, along with the original text, are chunked together, ensuring the image link is included within the chunk (4). Next, embeddings are generated from these chunks (5), which are then populated into an Azure Search index with the webpage links included as metadata (6). Finally, PromptFlow is utilized to implement Retrieval-Augmented Generation (RAG) for more efficient communication (7).

When accessed from an application, the system retrieves the relevant text and visuals (links) from the indexed knowledge base. By combining these elements, the chatbot delivers enriched responses that includes both **accurate text** and **relevant images**.

Conversation Example.

In the example presented below, the user asks a question about Ascent's thoughts on data platform architecture. The chatbot pulls relevant **textual information** from Ascent's web site and provides the answer with **relevant visual aids** and **links** to the source webpage. The result is a richer, more engaging



user experience, where the answers are not only accurate but also backed by informative images and citations from the website.

Chat Example:

What is Ascent's go-to data platform architecture?
K

Ascent's go-to data platform architecture leverages a modern, composable data platform built on Microsoft's Azure technologies. This integrated approach aims to deliver rapid return on investment by consolidating core infrastructure, ETL, BI, AI, and data science within a unified environment. Here's a detailed look at the architecture:

1. **Data Integration and Storage:**
 - **Real-time and Batch Data:** Integrates data from SaaS, PaaS, and on-premise systems.
 - **Azure Data Lake:** Utilizes a medallion architecture for data storage.
 - **Additional Storage Solutions:** Includes services like Snowflake and Synapse.
2. **Data Processing and Orchestration:**
 - **Compute Services:** Azure Functions and Azure Kubernetes Services for various compute needs.
 - **Orchestration:** Managed through Azure Data Factory and other orchestration tools.
3. **Data Enrichment and Analysis:**
 - Leverages advanced analytic tools and AI capabilities, including services like OpenAI.
4. **Data Serving and Consumption:**
 - **PowerBI, APIs, and Apps:** Serve and make data accessible for applications and end-users.
 - **Business Intelligence and Advanced Analytics:** Empower organizations with data models, KPIs, and integrations.
5. **Operational Components:**
 - **Security and Governance:** Enhanced through continuous operations tools focusing on security, cost management, monitoring, and compliance.
 - **DevSecOps Practices:** Ensures security is integrated throughout the lifecycle.
6. **Insights and Decision-Making:**
 - **Ascent Insights:** Offers valuable insights around cost, security, and sustainability, driving informed decision-making.

For more information, you can explore their [Modern Data Platform](#) and [Modern App Platform](#).

Backend

Store
 Cosmos DB, Azure SQL

Compute
 Azure Functions, Azure Kubernetes Services, App Services, Container Registry

Integrate
 Service Bus, Event Grid, API Hub

AI
 Cognitive Services, Open AI

Monitor
 Azure Monitor, App Insights, Log Analytics

Presentation

Contentful, User Interface, API, Figma

Consumption

Mobile Applications, Web Applications, Users

Operate

Microsoft Entra ID, Azure AD B2C, Cost Management, Key Vault, Azure Monitor, Microsoft Defender, Azure DevOps, GitHub, Azure Policy

ASCENT Insights

Cost Insights, Security Insights, Change Detection, Opportunity Management, Sustainability

Rate your experience: ☆☆☆☆☆



Conclusion.

By combining **RAG** with **GPT-4o** and integrating text and visual data from websites, businesses can create more advanced, contextually aware conversational systems. These systems not only answer questions accurately but also provide visual content that enhances understanding. For companies, this means offering more value to users through detailed, context-rich conversations that incorporate the latest advancements in AI.

This approach opens new opportunities for businesses looking to leverage their existing content and present it through a more modern interface, ensuring that users, customers, and employees get the most relevant and visually supported answers possible.

Explore how RAG-based conversational systems can revolutionize your website's user experience with enhanced text and visual data. Let's get started!

ascent.io/services/data-ai/artificial-intelligence

