Entity Extractor

Connected System Plugin for Appian
Overview 3
Features 3
Connected System Configuration 4
  Chat Completion Model: OpenAI 4
  Chat Completion Model: Azure OpenAI 7
Integration 11
  1. Document Input Type 11
  2. Text Input Type 12
Overview

The Entity Extractor Connected System allows developers to extract specific entities with the help of OpenAI services. It allows users to extract all or specific entities from a text or an Appian document.

An entity refers to a specific piece of information that is categorized and extracted from a text. In the context of Natural Language Processing (NLP) and text analysis, entities are typically nouns that have a distinct and well-defined meaning. They often represent real-world objects, such as people, locations, organizations, dates, products, etc.

For example, in the sentence "Apple Inc. was founded by Steve Jobs in Cupertino on April 1, 1976," the entities might be:

"Apple Inc." (Organization)
"Steve Jobs" (Person)
"Cupertino" (Location)
"April 1, 1976" (Date)

Entity extraction, also known as Named Entity Recognition (NER), is the process of identifying and classifying these entities into predefined categories. It's a crucial task in many NLP applications, including information retrieval, question answering, and summarization.

Developers can leverage their entity extraction capabilities through Appian with this connected system by entering the credentials retrieved from either OpenAI or Azure OpenAI Studio. This documentation outlines the process of obtaining and leveraging these credentials within the Appian platform.

Note: This AI Tool can utilize external LLMs such as OpenAI and Azure OpenAI. As with all plug-ins, data sent to these external organizations is subject to the external organization's security policies, procedures, and pricing. Please check with your administrator and your organization's AI policy before sending any sensitive information to external services.

Features

- Extract user specified entities from a given text or document
The Entity Extraction connected system with OpenAI authentication requires the following credentials: **OpenAI API Key** and **Chat Completion model**.
To retrieve your OpenAI API Key:

1. Go to the [OpenAI console](https://openai.com). Make sure that the **API keys** menu is selected.

2. Click on **Create new secret key** to generate a new API key.

3. Copy the value and save it separately as we won’t be able to access it again. Paste the API key in the connected system dialog box.

   ![Create new secret key](image)

   To find the appropriate Chat Completions model:

   1. Visit [https://platform.openai.com/docs/models/model-endpoint-compatibility](https://platform.openai.com/docs/models/model-endpoint-compatibility) and use one of the models listed under `/v1/chat/completions` endpoint. Example: gpt-3.5-turbo
for GPT 3.5 Turbo model, gpt-4 for GPT 4 model.

a. Each model has unique strengths so try to select the most appropriate for your use. If you would like to prioritize consistency in model's behavior, we recommend you use a GPT 4 model. If you need to prioritize speed of generation, GPT 3.5 Turbo might be better suited.
Chat Completion Model: Azure OpenAI

Connected System Properties

Entity Extractor
Extract entities from text using ChatGPT
Version: 1

Name *
EES CS Entity Extractor

Description

Entity Extractor Configuration

Authentication
Azure OpenAI Services

Use the Azure OpenAI services for Chat Completion

Azure Region *
eastus

Provide the Azure region.

Deployment ID *
GPT35_Turbo

Provide the Deployment ID.

Azure API Key *

Provide the API Key obtained from Azure OpenAI

Connection successful

TEST CONNECTION

CANCEL USE IN NEW INTEGRATION SAVE
This authentication requires the following credentials: Azure Region, Azure Deployment ID and Azure API Key. Follow these steps to get the Azure credentials.

Set up your Azure OpenAI Account

1. Navigate to Azure’s OpenAI API docs and ensure you have met the listed prerequisites. View the prerequisites by selecting “Quickstarts.” If you have not already done so, create an Azure subscription.

2. Apply for access to Azure OpenAI services by completing the form here. You will need your subscription ID from the previous step.

3. Create a service and set your domain name.
Create and Access API Keys
4. Within your service, create and access API keys through “Keys and Endpoints” under Resource Management. The "Location/Region" listed in this window will be used as your Azure Region in the Connected System configuration.

Deploy OpenAI Models and Set your Deployment ID
5. Visit the Azure OpenAI Studio to deploy OpenAI models under your created resource.
6. Click into your resource to enter the Azure OpenAI Studio. Navigate to “Models” seen as a tab on the left side of the screen. After selecting the best OpenAI Model for your data and use case, deploy your selected chat completions model. The **deployment ID** you create during deployment will be used when configuring your Connected System.
Integration

Extract entities from a text or document.

ChatGPT Prompt used:
You are a highly intelligent entity extractor that displays only the entities in
[{'entity': 'entity_name', 'value': 'entity_value'}] JSON Array format and return empty JSON Array
if you are unable to find the entities. An entity is a valuable piece of information contained within
the text. Entities will be given as Array of String. Extract <all the entities / entities list> from the
following text

1. Document Input Type

Inputs:

**Document** (Document) - Required - Provide the document from which the entities are extracted.

**Batch Number** (Number(Integer)) – Optional - Provide the batch number. Default: 1

**Number of Pages** (Number(Integer)) – Optional - Provide the number of pages (Batch size) to
be selected for entity extraction. Default: 1

**Entities to extract** (List of Text) – Optional - Provide the entities to be extracted from the
document as a list of text items. Provide null to extract all the entities.

Output: Dictionary

{
2. Text Input Type

**Inputs:**

- **Input Text** (Text) - Required - Provide the text from which the entities are extracted.

- **Entities to extract** (List of Text) – Optional - Provide the entities to be extracted from the text as a list of text items. Provide null to extract all the entities.
Output: Dictionary

{  
  entities: { { value: "OpenAI", entity: "organization" } },  
  success: true  
}