**This document implements few NLP tasks using GenAI plugins step using Pentaho Data Integration (PDI) which calls OpenAI based LLM. The tests are done via “gpt-4o-mini” model.**

**Below is a snapshot of various Natural Language Processing (NLP) tasks are categorized as below. Few of these NLP tasks we will address using PDI and LLM (OpenAI).**

A diagram of a flowchart

AI-generated content may be incorrect.

**Data Model:**

ERD: <https://dbdiagram.io/d/ChatGpt_PDI-67c71518263d6cf9a030e12c>

A screenshot of a computer

AI-generated content may be incorrect.

**Pentaho Data Integration jobs and transformtations:**A screenshot of a computer

AI-generated content may be incorrect.

**#Generate Data (multiple natural languages) using OpenAI and store in a database for processing NLP**

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**#Testing to see SQL works**

A screenshot of a computer

AI-generated content may be incorrect.

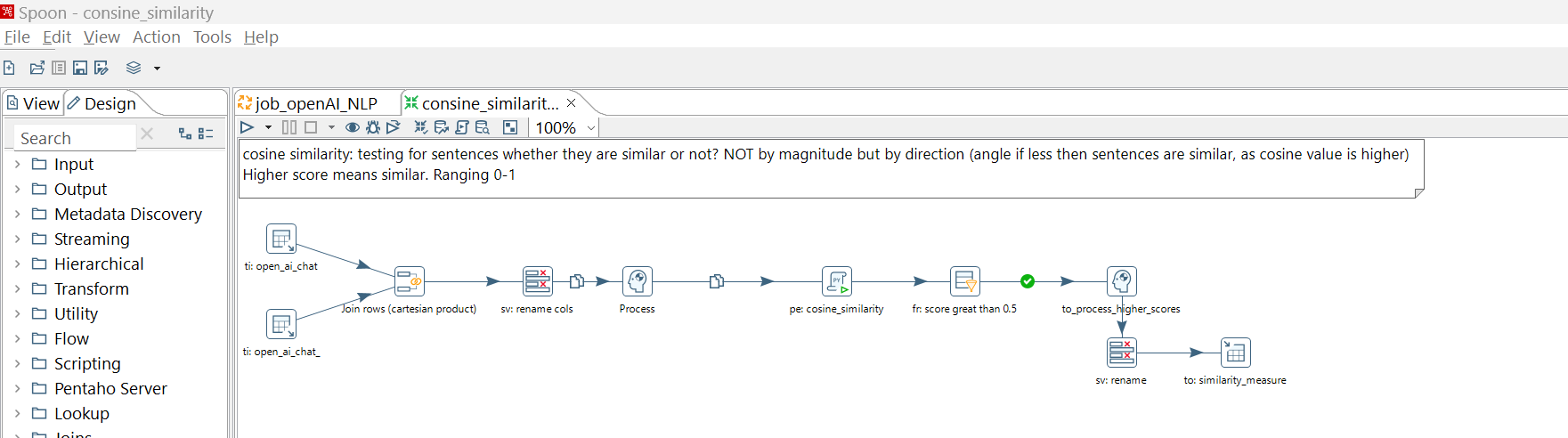
**#Translation, Sentiment, NER, bigrams, PoS**

A screenshot of a computer

AI-generated content may be incorrect.

**#Sentence Similarity (How similar are two sentences?)**

**ChatGPT couldn’t provide a score through prompt, so having this via Python while using data processing in PDI transformation (data in flight).**



**--Co-Reference**

A screenshot of a computer

AI-generated content may be incorrect.