In [1]:
```
#!pip install unstructured chromadb openai langchain-openai langchain-community
```

In [2]:
```
import os
from dotenv import load_dotenv
load_dotenv(override=True)
OPENAI_API_KEY = os.getenv("OPENAI_API_KEY")
```

In [3]:
```
from langchain_classic.document_loaders import TextLoader
documents = TextLoader("AI.txt").load()
```

```
C:\ProgramData\anaconda3\Lib\site-packages\keras\src\export\tf2onnx_lib.py:8: FutureWarning: In the future `np.object` will be defined as the corresponding NumPy scalar.
  if not hasattr(np, "object"):
WARNING:tensorflow:From C:\ProgramData\anaconda3\Lib\site-packages\tf_keras\src\losses.py:2976: The name tf.losses.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

In [4]:
```
from langchain_classic.text_splitter import RecursiveCharacterTextSplitter

# 문서를 청크로 분할
def split_docs(documents,chunk_size=1000,chunk_overlap=20):
  text_splitter = RecursiveCharacterTextSplitter(chunk_size=chunk_size, chunk_overlap=chunk_overlap)
  docs = text_splitter.split_documents(documents)
  return docs

# docs 변수에 분할 문서를 저장
docs = split_docs(documents)
```

In [5]:
```
#OpenAI의 임베딩 모델 사용
from langchain_openai import OpenAIEmbeddings
embeddings = OpenAIEmbeddings(model="text-embedding-ada-002", api_key=OPENAI_API_KEY)

# Chromdb에 벡터 저장
from langchain_classic.vectorstores import Chroma
db = Chroma.from_documents(docs, embeddings, persist_directory="data")
```

In [6]:
```
from langchain_classic.chat_models import ChatOpenAI
model_name = "gpt-4.1-mini"
```

```python
llm = ChatOpenAI(model_name=model_name, api_key=OPENAI_API_KEY)

# Q&A 체인을 사용하여 쿼리에 대한 답변 얻기
from langchain_classic.chains.question_answering import load_qa_chain
chain = load_qa_chain(llm, chain_type="stuff",verbose=True)

# 쿼리를 작성하고 유사성 검색을 수행하여 답변을 생성,따라서 txt에 있는 내용을 질의해야 합니다
query = "AI란?"
matching_docs = db.similarity_search(query)
answer =  chain.run(input_documents=matching_docs, question=query)
answer
```

```
C:\Users\User\AppData\Local\Temp\ipykernel_23600\3009664555.py:3: LangChainDeprecationWarning: The class `ChatOpenAI` was depre
cated in LangChain 0.0.10 and will be removed in 1.0. An updated version of the class exists in the `langchain-openai package a
nd should be used instead. To use it run `pip install -U `langchain-openai` and import as `from `langchain_openai import ChatOp
enAI``.
  llm = ChatOpenAI(model_name=model_name, api_key=OPENAI_API_KEY)
C:\Users\User\AppData\Local\Temp\ipykernel_23600\3009664555.py:7: LangChainDeprecationWarning: This class is deprecated. See th
e following migration guides for replacements based on `chain_type`:
stuff: https://python.langchain.com/docs/versions/migrating_chains/stuff_docs_chain
map_reduce: https://python.langchain.com/docs/versions/migrating_chains/map_reduce_chain
refine: https://python.langchain.com/docs/versions/migrating_chains/refine_chain
map_rerank: https://python.langchain.com/docs/versions/migrating_chains/map_rerank_docs_chain

See also guides on retrieval and question-answering here: https://python.langchain.com/docs/how_to/#qa-with-rag
  chain = load_qa_chain(llm, chain_type="stuff",verbose=True)
C:\Users\User\AppData\Local\Temp\ipykernel_23600\3009664555.py:12: LangChainDeprecationWarning: The method `Chain.run` was depr
ecated in langchain-classic 0.1.0 and will be removed in 1.0. Use `invoke` instead.
  answer =  chain.run(input_documents=matching_docs, question=query)
```

```
> Entering new StuffDocumentsChain chain...


> Entering new LLMChain chain...
Prompt after formatting:
System: Use the following pieces of context to answer the user's question.
If you don't know the answer, just say that you don't know, don't try to make up an answer.
----------------
Artificial intelligence (AI) is the intelligence of machines or software, as opposed to the intelligence of humans or animals.
It is a field of study in computer science that develops and studies intelligent machines. Such machines may be called AIs.

AI technology is widely used throughout industry, government, and science. Some high-profile applications are: advanced web sea
rch engines (e.g., Google Search), recommendation systems (used by YouTube, Amazon, and Netflix), understanding human speech (s
uch as Google Assistant, Siri, and Alexa), self-driving cars (e.g., Waymo), generative and creative tools (ChatGPT and AI art),
and superhuman play and analysis in strategy games (such as chess and Go).[1]

Artificial intelligence (AI) is the intelligence of machines or software, as opposed to the intelligence of humans or animals.
It is a field of study in computer science that develops and studies intelligent machines. Such machines may be called AIs.

AI technology is widely used throughout industry, government, and science. Some high-profile applications are: advanced web sea
rch engines (e.g., Google Search), recommendation systems (used by YouTube, Amazon, and Netflix), understanding human speech (s
uch as Google Assistant, Siri, and Alexa), self-driving cars (e.g., Waymo), generative and creative tools (ChatGPT and AI art),
and superhuman play and analysis in strategy games (such as chess and Go).[1]

Alan Turing was the first person to carry out substantial research in the field that he called Machine Intelligence.[2] Artific
ial intelligence was founded as an academic discipline in 1956.[3] The field went through multiple cycles of optimism[4][5] fol
lowed by disappointment and loss of funding.[6][7] Funding and interest vastly increased after 2012 when deep learning surpasse
d all previous AI techniques,[8] and after 2017 with the transformer architecture.[9] This led to the AI spring of the 2020s, w
ith companies, universities, and laboratories overwhelmingly based in the United States pioneering significant advances in arti
ficial intelligence.[10]

Alan Turing was the first person to carry out substantial research in the field that he called Machine Intelligence.[2] Artific
ial intelligence was founded as an academic discipline in 1956.[3] The field went through multiple cycles of optimism[4][5] fol
lowed by disappointment and loss of funding.[6][7] Funding and interest vastly increased after 2012 when deep learning surpasse
d all previous AI techniques,[8] and after 2017 with the transformer architecture.[9] This led to the AI spring of the 2020s, w
ith companies, universities, and laboratories overwhelmingly based in the United States pioneering significant advances in arti
ficial intelligence.[10]
Human: AI란?

> Finished chain.
```

> Finished chain.

Out[6]: 'AI(인공지능)는 기계나 소프트웨어가 인간이나 동물의 지능과는 구별되는 지능을 갖도록 하는 기술 및 연구 분야를 말합니다. 컴퓨터 과학의 한 분야로, 지능적인 기계를 개발하고 연구하는 데 중점을 둡니다. AI는 웹 검색 엔진, 추천 시스템, 음성 인식, 자율주행차, 생성적 창작 도구(예: ChatGPT), 그리고 체스나 바둑 같은 전략 게임에서 인간을 능가하는 플레이 능력 등 다양한 분야에서 널리 활용되고 있습니다.'