```
# pip install guidance llama_index langchain chromadb python-dotenv


'''dotenv for loading the OPENAI_API_KEY'''
from dotenv import load_dotenv

'''langchain for loading the data and charachter text splitting'''
from langchain.document_loaders import DirectoryLoader
from langchain.text_splitter import CharacterTextSplitter

'''vector database to use for guidance'''
import chromadb

'''two methods for chatbot creation'''
import guidance
from llama_index import VectorStoreIndex, SimpleDirectoryReader


# Loading the OPENAI_API_KEY
load_dotenv()
```

```
True
```

```
llmcc = guidance.llms.OpenAI("gpt-3.5-turbo")


# Using langchain's directory loader to load the data present in research_papers directory and create its document
data = 'research_papers'          # dir path
loader = DirectoryLoader(data)    # loader obj
docs = loader.load()              # creates list of docs


# Check the docs
print(docs)
```

```
[Document(page_content='ResultsinOptics12(2023)100473\n\nAvailableonline28June20232666-9501/©2023TheAuthor(s).PublishedbyElsevierB.V.Thi
```

```
# Document Structure is ready!
print(type(docs[0]))
```

```
<class 'langchain.schema.document.Document'>
```

```
# Split the document into chunks
splitter = CharacterTextSplitter(chunk_overlap=200, chunk_size=1024) # overlap allows to have some notion of consistency between chunks
splitted = splitter.split_documents(docs)
```

```
2023-09-17 10:37:14,437 - 22360 - text_splitter.py-text_splitter:181 - WARNING: Created a chunk of size 1448, which is longer than the s
2023-09-17 10:37:14,439 - 22360 - text_splitter.py-text_splitter:181 - WARNING: Created a chunk of size 1159, which is longer than the s
```

```
# Prepare the documents and unique ids for ChromaDB collection
uids, docs = zip(*[(f'doc{id+1}', doc.page_content) for id, doc in enumerate(splitted)])


uids
```

```
('doc1',
 'doc2',
 'doc3',
 'doc4',
 'doc5',
 'doc6',
 'doc7',
 'doc8',
 'doc9',
 'doc10',
 'doc11',
 'doc12',
 'doc13',
 'doc14',
 'doc15',
 'doc16',
 'doc17',
 'doc18',
 'doc19',
```

```
 'doc20',
 'doc21',
 'doc22',
 'doc23',
 'doc24',
 'doc25',
 'doc26',
 'doc27',
 'doc28',
 'doc29',
 'doc30',
 'doc31',
 'doc32',
 'doc33',
 'doc34',
 'doc35',
 'doc36',
 'doc37',
 'doc38',
 'doc39',
 'doc40',
 'doc41',
 'doc42',
 'doc43',
 'doc44')
```

docs

```
('ResultsinOptics12(2023)100473\n\nAvailableonline28June20232666-9501/
©2023TheAuthor(s).PublishedbyElsevierB.V.ThisisanopenaccessarticleundertheCCBY-NC-NDlicense(http://creativecommons.org/licenses/by-nc-nd/4.0/).\n\nContents lists available at ScienceDirect\n\nResults in Optics\n\njournal homepage:
www.sciencedirect.com/journal/results-in-optics\n\nFundamental study on optical performance of low-melting-point metal mirrors for space telescopes\n\nEisuke Imaizumi a, Masatoshi Kondo b,*, Katsuya Murakami a, Yutaka Hayano c, Yuichi Matsuda c a Tokyo Institute of Technology, School of Engineering, Department of Mechanical Engineering, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan b Tokyo Institute of Technology, Institute of Innovative Research, Laboratory for Zero-Carbon Energy, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan c National Astronomical Observatory of Japan, 2-21-1, Osawa, Mitaka-shi, Tokyo 181-8588, Japan\n\nA R T I C L E I N F O\n\nA B S T R A C T',
 'A R T I C L E I N F O\n\nA B S T R A C T\n\nKeywords: Low-melting-point metal Liquid metal Reflectance Space telescopes',
 'The installation of telescopes with larger primary mirrors in outer space is required for the study of the early universe. However, the shipping, assembling and maintenance of the primary mirror are technical challenges. The primary mirror made of low-melting-point metal (LMPM) enables to simplify these procedures. The rotating liquid metal surface is shaped into a parabola and solidified on the site. The technological feasibility depends on the dynamic transformation of the parabolic LMPM mirror and its optical reflectance. The dynamic trans- formation of LMPM mirror was clarified by means of experiments with liquid gallium (Ga) which has a melting point of 302.91 K. The parabolic shape of solid Ga mirror was produced by the axial rotation and the solidifi- cation. The parabolic shape agreed with the theoretical curve obtained by a physical model. The reflectance spectra on the surface of various LMPMs solidified (i.e., Ga, lead (Pb), tin (Sn), bismuth (Bi), lead-bismuth eutectic (Pb-Bi) and Wood’s metal (Bi-Pb-Sn-Cd)) were measured. The reflectance of Ga solidified was approx- imately 70 % and was not degraded in the near-infrared wavelength due to the favorable optical characteristics. The reflectance both for p-polarized and s-polarized light of Ga solidified was measured by the spectropho- ◦ tometer with the polarizer at the angle of incidence of 30 . The trend of measured reflectance agreed with theoretical prediction.',
 ', 45\n\n, 60\n\nand 75\n\n1. Introduction\n\nTelescope was recently launched, and the diameter of the primary mirror is 6.5 m. These space telescopes have made great impacts to the astronomical science such as a characterization of exoplanets (Char- bonneau et al., 2002; Hinkly, 2022) and a discovery of galaxy formation in the early universe (Williams, 1996; Welch et al., 2022) and highly-redshifted objects. The discovery can be performed by the space obser- vation in infrared wavelength. The space telescopes with larger primary mirrors are being developed, and the improvements of thermal design (Li et al., 2016; Bannon et al., 2021) and vibration isolation system (Qin et al., 2020) have been studied. However, the mirrors are extremely vulnerable and the diameter to be transported is restricted by the size space ships. Therefore, the space transportation of larger primary mir- rors provides a critical challenge from a technological viewpoint.',
 'The lunar surface is an attractive location for long-term astronomical observations using telescopes, since night-time hours are long as approximately 14 days and the astronomical observations are not affected by the atmosphere and weather conditions. The concept of lunar telescope (LT) has been proposed in the previous study, and the LT with 15 cm aperture mirror has been operated from 2013 (Meng et al., 2016). Artemis program will promote the exploration of the moon (Kumari et al., 2022; Evans and Graham, 2020).',
 'The exploration of the early universe is one of the important topics to clarify the evolution of the universe. The infrared observations using space telescopes located in the outer space and/or on the moon at very low temperature will play an important role in the study of the early universe. The space telescopes with larger primary mirrors have higher angular resolution and light collecting power. Hubble Space Telescope is the optical/infrared space telescope operating on the orbit of the earth, and the diameter of the primary mirror is 2.4 m. James Webb Space\n\nThe concept of liquid metal telescope (LMT) has been studied (Ols- sonsteel, 1986; Borra, 1982), and the LMT is operated in Canada (Kumar et al., 2018; Hickson et al., 2007). The spin of liquid mercury (Hg) which has a melting point of 234.32 K forms its parabolic mirror. The aperture diameter of LMT can be increased more easily than the\n\nCorresponding author.',
 'Corresponding author.\n\nE-mail addresses: kondo.m.ai@m.titech.ac.jp (M. Kondo), y.hayano@nao.ac.jp (Y. Hayano), yuichi.matsuda@nao.ac.jp (Y. Matsuda).\n\nhttps://doi.org/10.1016/j.rio.2023.100473 Received 21 March 2023; Received in revised form 11 May 2023; Accepted 26 June 2023\n\nResultsinOptics12(2023)1004732\n\nE. Imaizumi et al.',
 'https://doi.org/10.1016/j.rio.2023.100473 Received 21 March 2023; Received in revised form 11 May 2023; Accepted 26 June 2023\n\nResultsinOptics12(2023)1004732\n\nE. Imaizumi et al.\n\ntelescopes with solid mirrors (Li et al., 2019), since the LMT is not restricted by the technical issues of on-site assembling with high preci- sion and space transportation as explained in next chapter. The opera- tion of LMT on the lunar surface (Klimas, 2010) enables to observe distant objects which are not followed up with ground-based telescopes (Angel, 2006; Angel et al., 2008). However, the LMT cannot point at off- zenith (out-of-plane tip/tilt) angles since the liquid mirror cannot be tilted. The Zenith program of the Defense Advanced Research Projects Agency (DARPA) is going to investigate and develop technologies to facilitate demonstration of LMT and liquid mirror that can be pointed at off-zenith (out-of-plane tip/tilt) angles.',
 'LMPM layer can reduce the total weight of the mirror which is major driver of the shipping cost.\n\nThe building of space telescopes with the parabolic LMPM mirror is rather simple. The LMPM melted is placed on the shell assembled, and axial rotation of
```

```
the shell with the LMPM melted shapes the parabolic mirror on the site. The heat and electricity are required only in the shaping procedure of the parabolic mirror. High precision is not required for the assembling of the shell with support system, though it is required for the assembling of segmented mirrors. The transformation on the orbit of the earth is also possible by the use of thrust
```

```
# For those who didn't understand the code above, here is using longer method

# uids = []
# docs = []

# for id, doc in enumerate(splitted):
#     uids.append(f'doc{id+1}')
#     docs.append(doc)


# Store the above processed data in chromadb
client = chromadb.Client()                                  # create chromaDB client
collection = client.create_collection("sample_collection")  # using the client create sample_collection
collection.add(documents=list(docs), ids=list(uids))        # add the ids and respective documents to database


# Retrieve the top-n information based on user query from this database which we be the knowledge base for LLM (gpt-3.5-turbo)
def knowledgebase(query):

    # Use the ChromaDB collection to perform a query with the provided 'query' text and request up to 3 results
    results = collection.query(query_texts=[query], n_results=3)  # result is the list of top 3 chunks based on similarity

    # Comment this code line later after demonstration is over
    print(results)

    # Join these 3 chunks and seperate them using new line charachter
    chromadb_output = "\n".join(results['documents'][0])

    # Return this string output
    return chromadb_output



knw = knowledgebase('can you list of specimens used for measurement of optical reflectance?')
```

```
{'ids': [['doc22', 'doc36', 'doc19']], 'distances': [[0.887841522693634, 0.9419827461242676, 0.959404706954956]], 'metadatas': [[None, N
```

```
knw
```

```
'3.3. Measurement of reflectance on surface of low-melting-point metal specimens\n\nThe reflectance of solidified LMPM specimens was evaluated by relative specular reflectance measurement with the spectrophotometer (SolidSpec-3700, SHIMADZU CORPORATION, Kyoto-shi, Kyoto, Japan) as shown in Fig. 3(a), which is installed in National Astronomical Ob- servatory of Japan (NAOJ). The reflectance of the LMPM specimens was based on the intensity ratio between the incident light and the light detected by the detector. The wavelength of incident light was varied in the range between ultraviolet and near infrared (300–2000 nm). The ◦ angle of incidence was 15 in all the measurements of optical reflectance.\nFig. 6. Optical reflectance spectra obtained by measurement with LMPM specimens : (a) Polished surface of metals, (b) Naturally solidified surface of LMPMs.\n\nResultsinOptics12(2023)1004737\n\nE. Imaizumi et al.\n\nFig. 7. Relationship between surface roughness and optical reflectance of LMPMs: (a) wavelength of incident light was 600 nm and (b) wavelength of incident light was 1500 nm.\n\nAcknowledgement\n\nThis research was supported by the grant of Joint Research by the National Institutes of Natural Sciences (NINS). (NINS program No, 01112204).\n\nReferences\n\nAdachi, S., 2012. The Handbook on Optical Constants of Metals In Tables and Figures.\n\nWorld Scientific, Singapura.\n\nAngel, R., et al., 2006. A Lunar Liquid Mirror Telescope (LLMT) for deep-field infrared observations near the lunar pole. Proceedings of SPIE. 6365, 62651U. https://doi.org/10.1117/12.669994.\nTable 2 presents the list of specimens used for the reflectance mea- surements. The specimens of solidified LMPMs (i.e., Wood’s metal, lead–bismuth eutectic (LBE), Pb, Bi, Sn and Ga) were prepared. The liquid LMPMs were melted by a plate heater and solidified under natural\n\nTable 2 List of specimens used for measurement of optical reflectance.\n\nSurface observation and optical reflectance measurement\n\nAfter polished\n\nAfter solidification (Static condition)\n\n–\n\n316L austenitic steel Wood’s metal LBE Pb Bi Sn Ga\n\nYes Yes Yes Yes Yes Yes No\n\nTable 3 Arithmetic mean surface roughness of low-melting-point metals solidified under natural cooling at static condition (Unit: µm).\n\nYes Yes Yes Yes Yes Yes\n\nWood’s metal\n\nLBE\n\nPb\n\nBi\n\nSn\n\nGa\n\n0.748\n\n5.636\n\n0.074\n\n0.534\n\n0.072\n\n0.049\n\nResultsinOptics12(2023)1004734\n\nE. Imaizumi et al.\n\n4. Results and discussions\n\nTable 4 Arithmetic mean surface roughness of metal specimens after various polishing procedures (Unit: µm).\n\n4.1. Dynamic formation of parabolic LMPM mirror\n\nPb'
```

```
def chatbot(input, chat_loop):

    knowledge = knowledgebase(input)
    if chat_loop is None:
        chat_loop = guidance(
                '''{{#system}}You are a helpful research assistant that answers any questions related to Astronomy research papers.{{/system

                        {{~#geneach 'conversation' stop=False}}

                        {{#user~}}

                        If user query is very generic like greeting or not at all scientific and doesn't need CONTEXT KNOWLEDGE then respond

                        Query: {{set 'this.user_text' (await 'user_text') hidden=False}}
                        Use this CONTEXT KNOWLEDGE below to frame your answer very intelligently and don't print any output that maybe unnec

                        CONTEXT KNOWLEDGE:-
                        {{knowledge}}

                        {{~/user}}

                        {{#assistant~}}
                        {{gen 'this.response' temperature=0 max_tokens=500}}
                        {{~/assistant}}

                        {{~/geneach}}''', llm = llmcc, silent = True, caching = False)

    chat_loop = chat_loop(user_text = input, knowledge = knowledge)
    return chat_loop['conversation'][-2]['response'], chat_loop



response, chat_loop = chatbot('Who are the authors of this paper on LMPMS?', None)
print(response)


chat_loop


response, chat_loop = chatbot('write a summary of abstract of this paper', chat_loop)
print(response)


chat_loop
```

```
# test case to check chat loop: works

chat_loop =  guidance(
                '''{{#system}}You are a helpful research assistant that answers any questions related to Astronomy research papers.{{/system

                        {{~#geneach 'conversation' stop=False}}

                        {{#user~}}

                        If user query is very generic like greeting or not at all scientific and doesn't need CONTEXT KNOWLEDGE then respond

                        Query: {{set 'this.user_text' (await 'user_text') hidden=False}}
                        Use this CONTEXT KNOWLEDGE below to frame your answer very intelligently and don't print any output that maybe unnec

                        CONTEXT KNOWLEDGE:-
                        {{knowledge}}

                        {{~/user}}

                        {{#assistant~}}
                        {{gen 'this.response' temperature=0 max_tokens=500}}
                        {{~/assistant}}

                        {{~/geneach}}''', llm = llmcc, silent = True, caching = False)


# chat loop (in a loop, both function (chat_loop1) and returned value (chat_loop2) should have same name
print('Type "Quit" to exit!\n')
while True:
    user_text = input()
    if user_text.lower() == 'quit':
        print('User:- Quit')
        print('Assistant:- I hope I was able to help you with your queries. Have a great day!')
        break
    knowledge = knowledgebase(user_text)
    #print('Knowledge:-')
    #print(knowledge)
    chat_loop = chat_loop(user_text = user_text, knowledge=knowledge)
    print(f'User:- {user_text}')
    response = chat_loop['conversation'][-2]['response']
    print(f"Assistant:- {response}")
```

```
documents = SimpleDirectoryReader('research_papers').load_data()
index = VectorStoreIndex.from_documents(documents)
query_engine = index.as_query_engine()
response = query_engine.query('What conclusions can be drawn from this results of the Arithmetic mean surface roughness of low-melting-point r
print(response)
```

```
    e arithmetic mean surface roughness of low-melting-point metals solidified under natural cooling at a static condition for different LMPM
     - Lead-bismuth eutectic (LBE): 5.636 µm
     - Pb: 0.074 µm
     - Bi: 0.534 µm
     - Sn: 0.072 µm
     - Ga: 0.049 µm

     These values were obtained through measurements using a 3D scanning laser microscope.

     User:- What conclusions can be drawn from this results?
     Assistant:- From the results obtained, several conclusions can be drawn:
```