

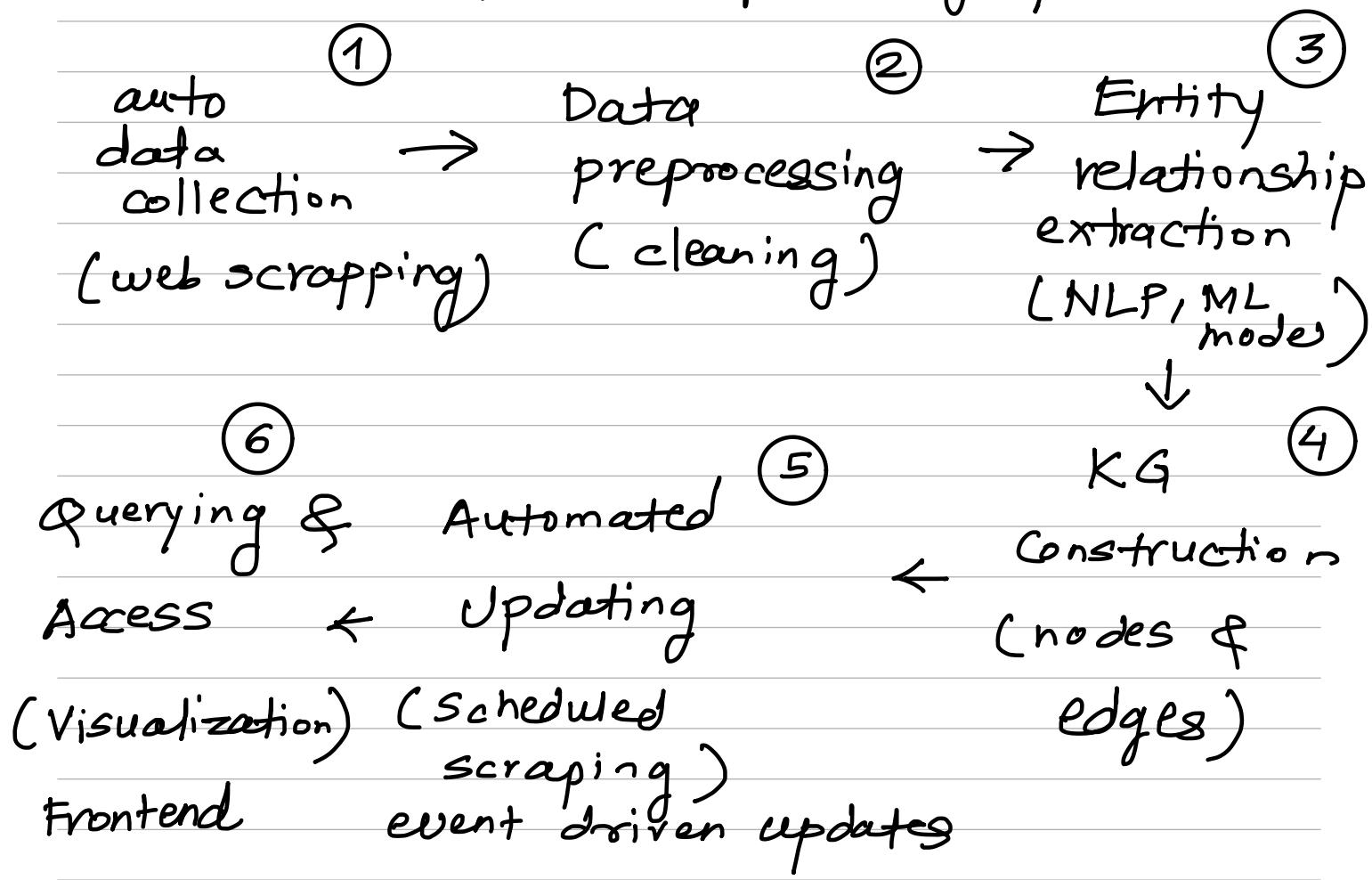
11/8/21

Automated Knowledge Graph Builder with Chat Interface

> Automated KG ?

minimum human intervention

auto data gather, extract entities,
relationship & update graph



Q. Think how to make KG scalable
key para that will reduce latency,
scales horizontally/vertically ?

↳ ① Database selection &
Optimization

neo4j / Azure Cosmos DB / Neptune
indexing, caching, sharding →
Amazon Neptune
distribute data
across multiple instances

② Data Ingestion & ETL Pipeline
batch / stream processing

③ Architecture for Scalability

microservices → stateless services
so that it will grow
horizontally

④ Query optimization &
management

⑤ Monitoring

performance metrics tools
ELK stack / Grafana

⑥ Security & access control

⑦ Horizontal & Vertical Scaling

Clustering to distribute data & load
across many servers
Allocate more CPU, memory.
Storage resources to your db

KG

— structured representation of knowledge that captures relationships between entities in graph format.

It consists of nodes & edges & represent relationship bet'n entities.

Assignment

Automated Knowledge Graph KG Builder with Chat Interface

Objective:

Build an automated Knowledge Graph KG that searches the web for information on a topic eg. Altera® FPGA, constructs a KG automatically, and provides a chat interface to query the KG. The chat interface should also be capable of creating a document based on the KG data.

Overview:

1. Web Scraping & Data Extraction: Automatically search the web for relevant information on a given topic.
2. Knowledge Graph Construction: Build the KG automatically, it should capture the entities and relationships.
3. Chat Interface: Create a web app using ReactJS for taking the inputs and presenting the output. The inputs can be "What is the on-chip RAM of Stratix FPGA"? The output would be 256 KB.
4. Document Creation: Generate a document based on user queries and the KG data.

Evaluation Criteria:

1. Functionality: Does the system meet the specified requirements?
2. Knowledge Graph Representation: How good is the representation? What algorithms have been used?
3. Code Quality: Is the code well-organized, documented, and follows best practices?

Problem Statement -

2 KG const

Build an automated Knowledge

Graph KG / that searches the web for
information on a topic eg. Altera®

FPGA, constructs a KG automatically,

and provides a chat interface to

query the KG. The chat interface

should also be capable of creating a

document based on the KG data.

1 web scrapping &
data extraction

3 Interface

4 document
creation

1 web scrapping & data extraction -

Tools used -
BeautifulSoup

Pandas

help in creating
dataframe with
data

pd. DataFrame()

https://www
link provided

to user

↓
provided

.CSV file

2 KG Construction

Tools Used

NetworkX

Spacy

Code asks
for path

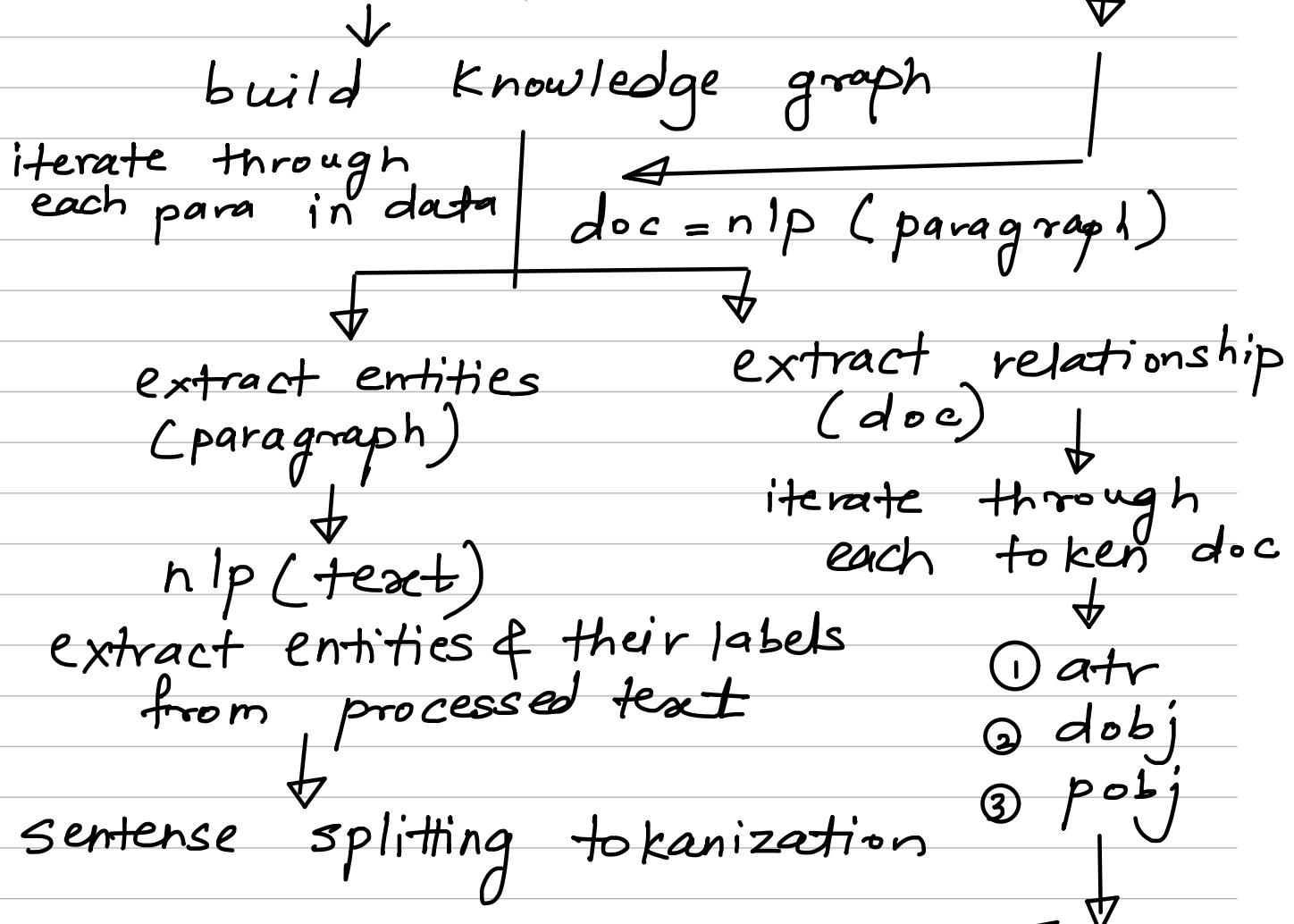
of .CSV file

user
enters
path

Pretrained
model

en_core_web_lg

CSV loaded = paragraphs

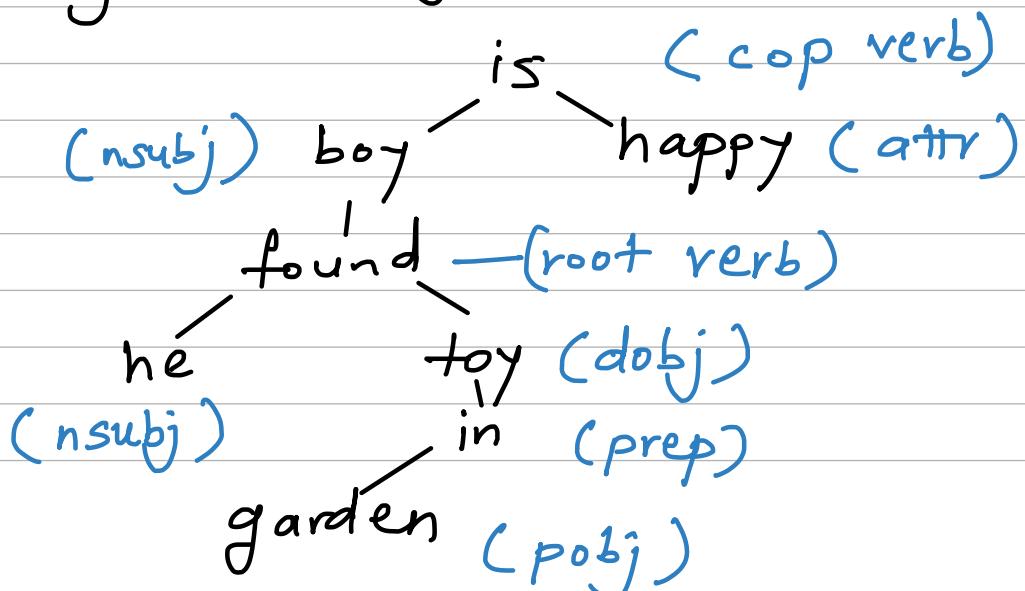


e.g "The boy is happy because he found a toy in the garden"

attr = happy — attribute

dobj = toy — Direct object

pobj = garden — object of preposition



iterate through
loop

entity = node
relationship = edge

data
in
folding
json
graph
return

save graph ()

file.json

i/p

③ Chat Interface

Tools —

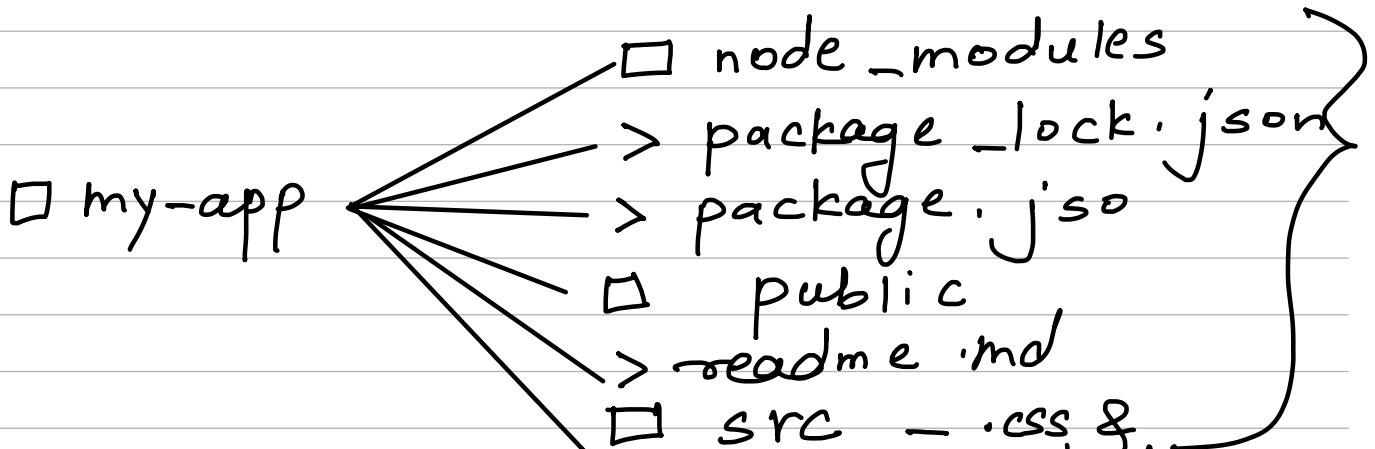
flask
(connectivity)

app.py

React JS
(frontend)

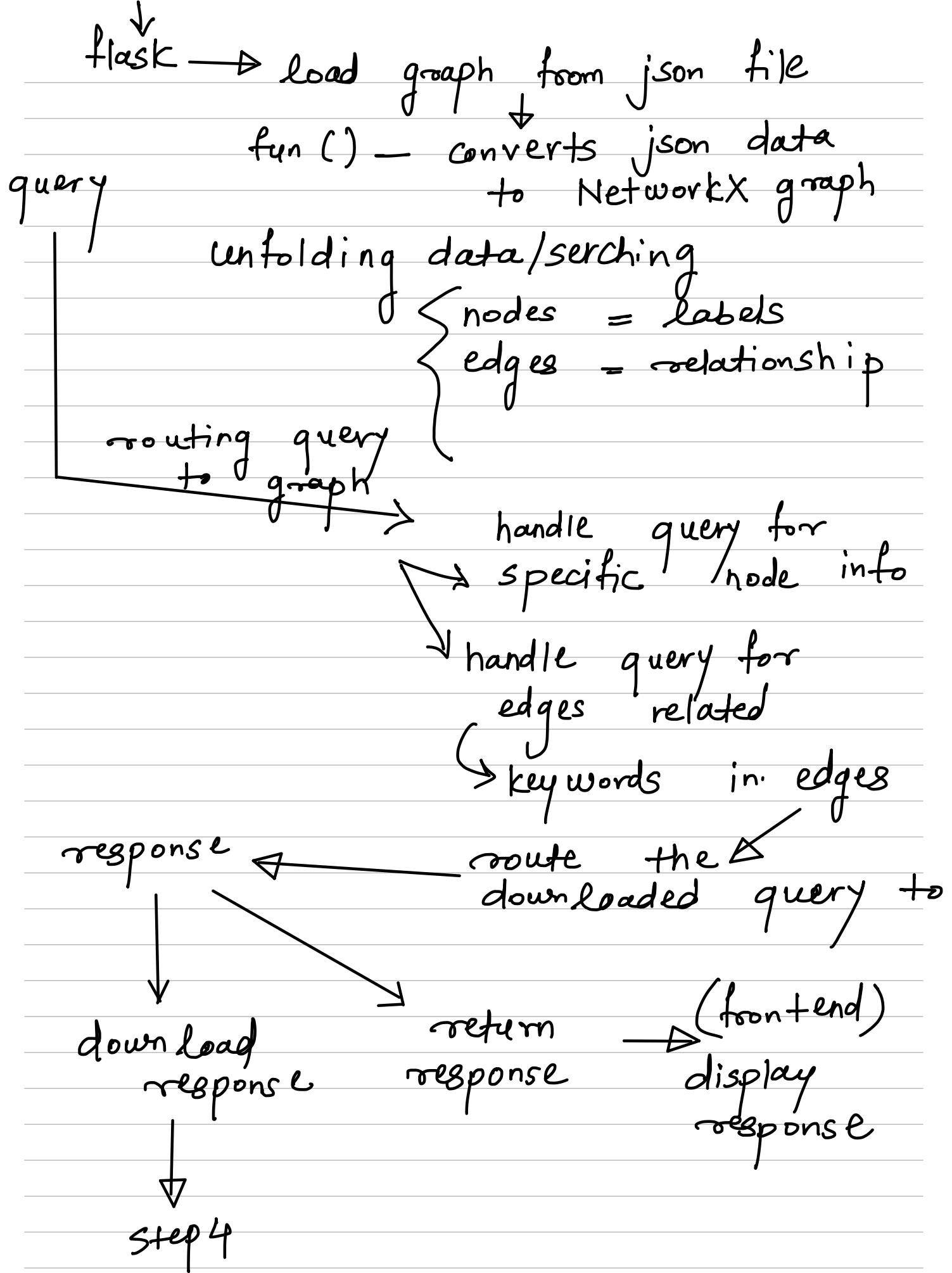
app.js

folder
structure



takes query from user

backend → derived KG → response



④ document creation :

after user entering query &
getting response
download response button pops up
& you can able to save file
on any path

UI

x	-	React App
localhost : 3000		
Knowledge Graph Chat		
Query :		
<input type="text"/>		Submit
Response :		
<input type="text"/>		
<input type="button" value="Download Response"/>		

