

Single Day Event

From Text to Insights: Lessons Learned from Large Language Model Application Construction

Xiaochang (Jack) Liu
Johnson & Johnson

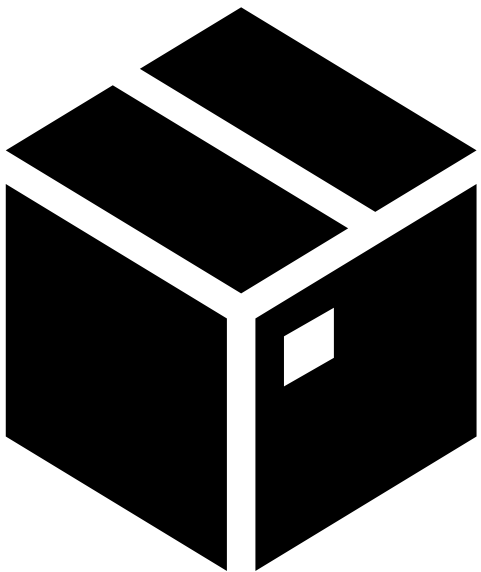


Disclaimer

This presentation is for informational purposes only and does not represent professional guidance or advice. Any views and opinions expressed during this presentation are those of the presenters and do not necessarily reflect the views or policies of Johnson & Johnson. This presentation may not be reproduced, modified, or copied, in whole or in part, without the express written permission of Johnson & Johnson.



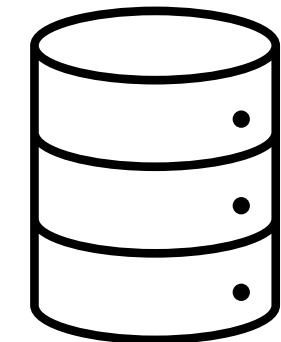
Large Language Model



A Black Box



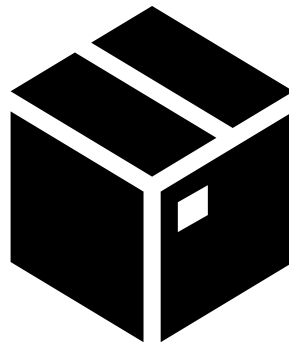
The Birth of an LLM



Tons of Data



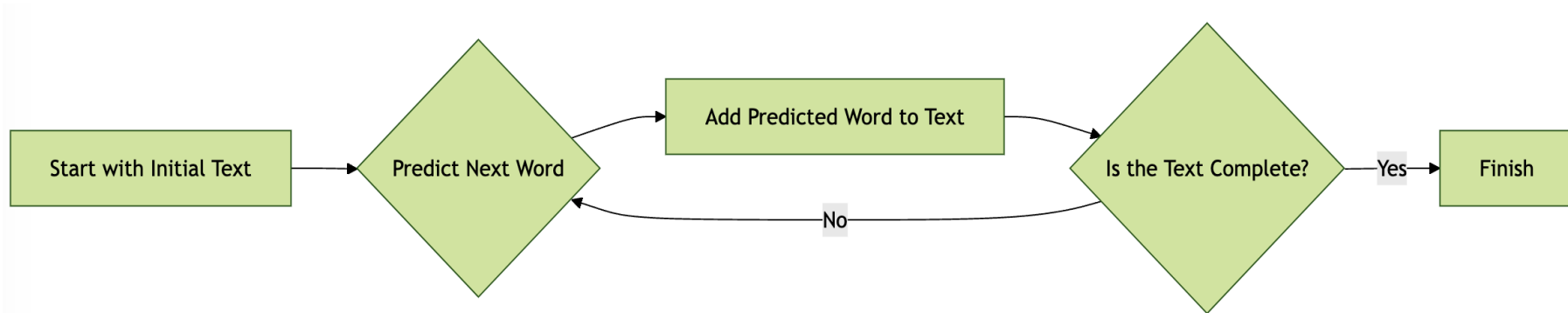
Pre-training



LLM



How does an LLM speak? (1/2)



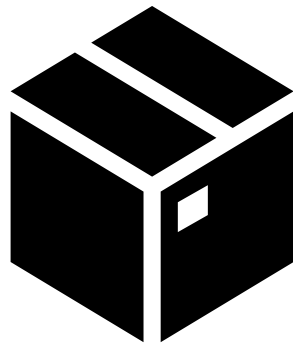
Workflow



How does an LLM speak? (2/2)

"I am Carl and I
use Python for"

Initial Text



LLM



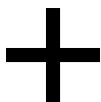
Candidate Word(s)	Probability
Work	0.891
Cooking	0.007
Eating	0.008



Retrieval-Augmented Generation

Carl is someone at a Company... He has extensive experience in utilizing **ML/Generative AI** across various stages of clinical and statistical programming and is proficient in programming using both R and **Python**. ...

Context Text



"I am Carl and I use Python for"

Initial Text



LLM



Candidate Word(s)	Probability
LLM Application Development	0.982
Work	0.891
Cooking	0.008



Retrieval



Vector Database

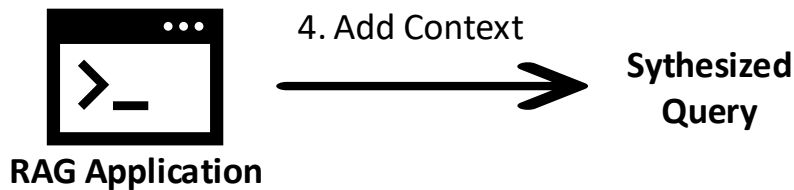
1. Embedding Creation
2. Embedding Matching
3. (Top-K)
4. (Reranking)

Graph Database

1. NER for input query
2. Graph Query Generation (e.g., Cypher)
3. Traversal of Graph Database
4. (Filtering and Scoring)



Augment



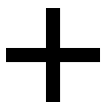
- "Context: [context here] Question: [query here]"
- "Given the following details: [context here]. What can you tell me about [query here]?"
- "In the context of [specific industry or domain], [context here]. Now, what is [query here]?"
- "Considering the following scenario: [context]. Based on this, how would you respond to [query]?"



Generation

Carl is someone at a Company... He has extensive experience in utilizing **ML/Generative AI** across various stages of clinical and statistical programming and is proficient in programming using both R and **Python**. ...

Context Text



"I am Carl and I use Python for"

Initial Text



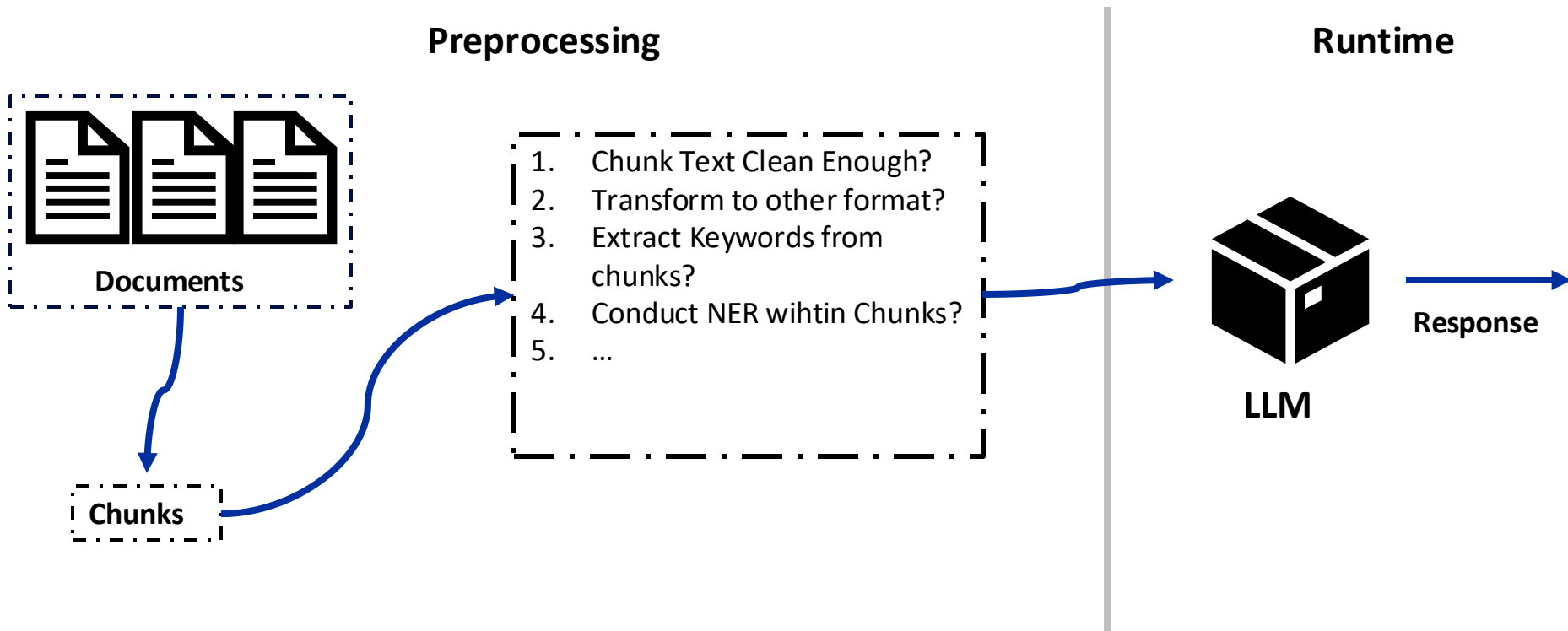
LLM



Candidate Word(s)	Probability
LLM Application Development	0.982
Work	0.891
Cooking	0.008



All About Data





Lesson Learned #1

Data Quality Matters

1. Incomplete Data
2. Inaccurate Data
3. Duplicated Data
4. Outdated Data



Lesson Learned #2

When NOT to use RAG frameworks?

1. When you have a limited choices of LLM.
2. When you have designed a clear and mature pipeline.
3. When you are tired of digging into the source code and found out there are so many pre-defined prompts but none of them could be applied to your project.



Summary

**I can RAG
You can RAG
We can RAG
Let's RAG!**

Thank you!



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitla](https://twitter.com/phusetwitla)
📘 [/phusebook](https://facebook.com/phusebook)
📺 [/phusetube](https://youtube.com/phusetube)
🌐 [/company/phuse](https://linkedin.com/company/phuse)