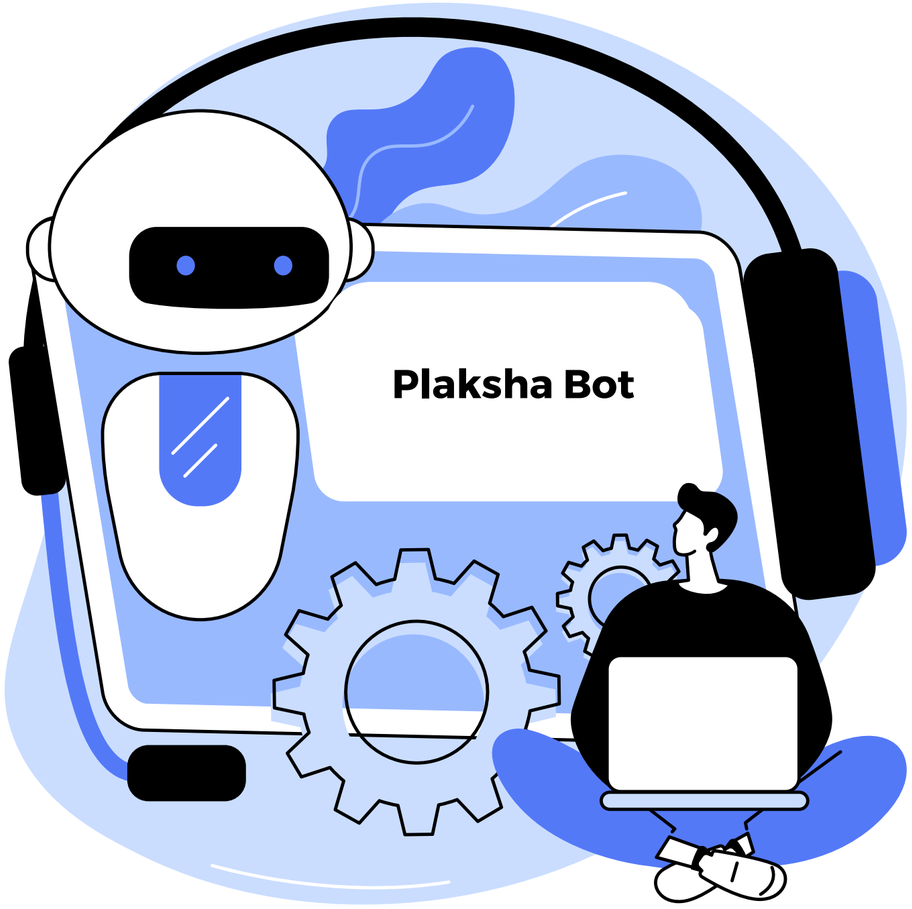


QUESTION AND ANSWER ASSISTANT FOR PLAKSHA'S WEBSITE

Mangesh Singh
Siddharth Sahu
Tushar Garg

PROBLEM STATEMENT

Navigating through Plaksha's website for specific information can be sometimes overwhelming, especially for those who are not part of the Plaksha community.

MOTIVATION

This will make fetching information from the website more convenient for users.

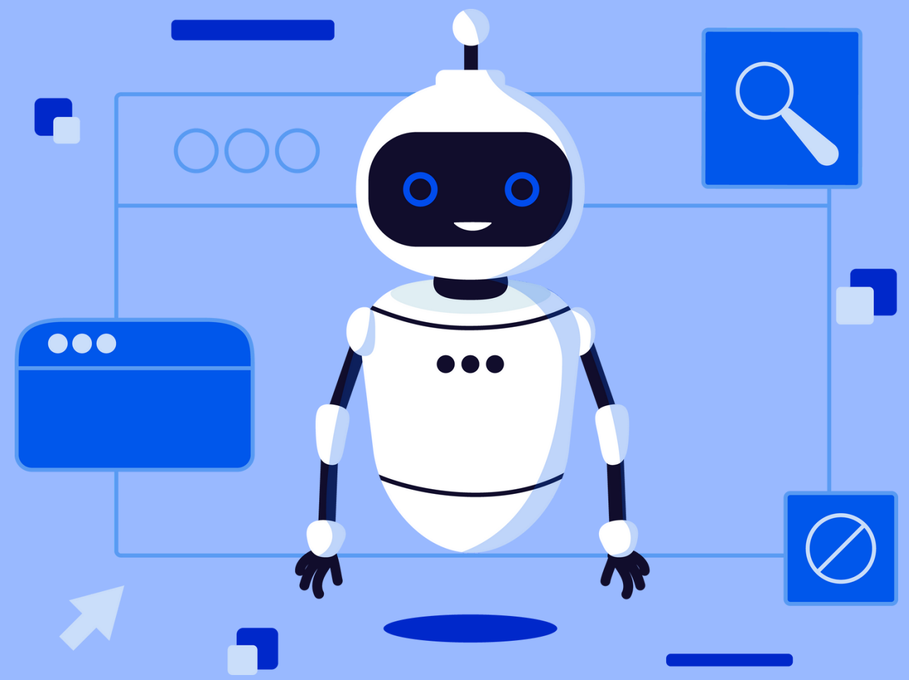
POTENTIAL APPLICATIONS

Question and Answer Assisstant

POTENTIAL IMPACT

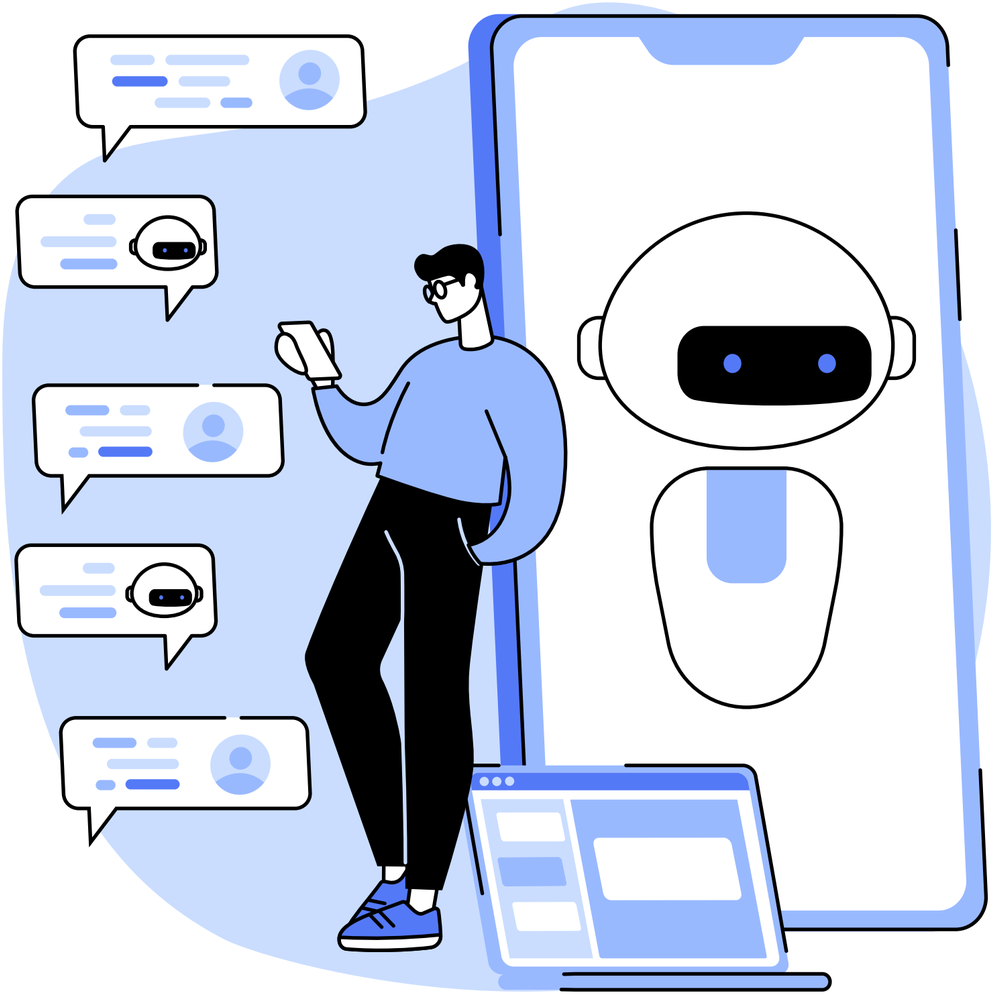
Enhanced User Experience, personalised assistance, innovation and differentiation

LITERATURE SURVEY



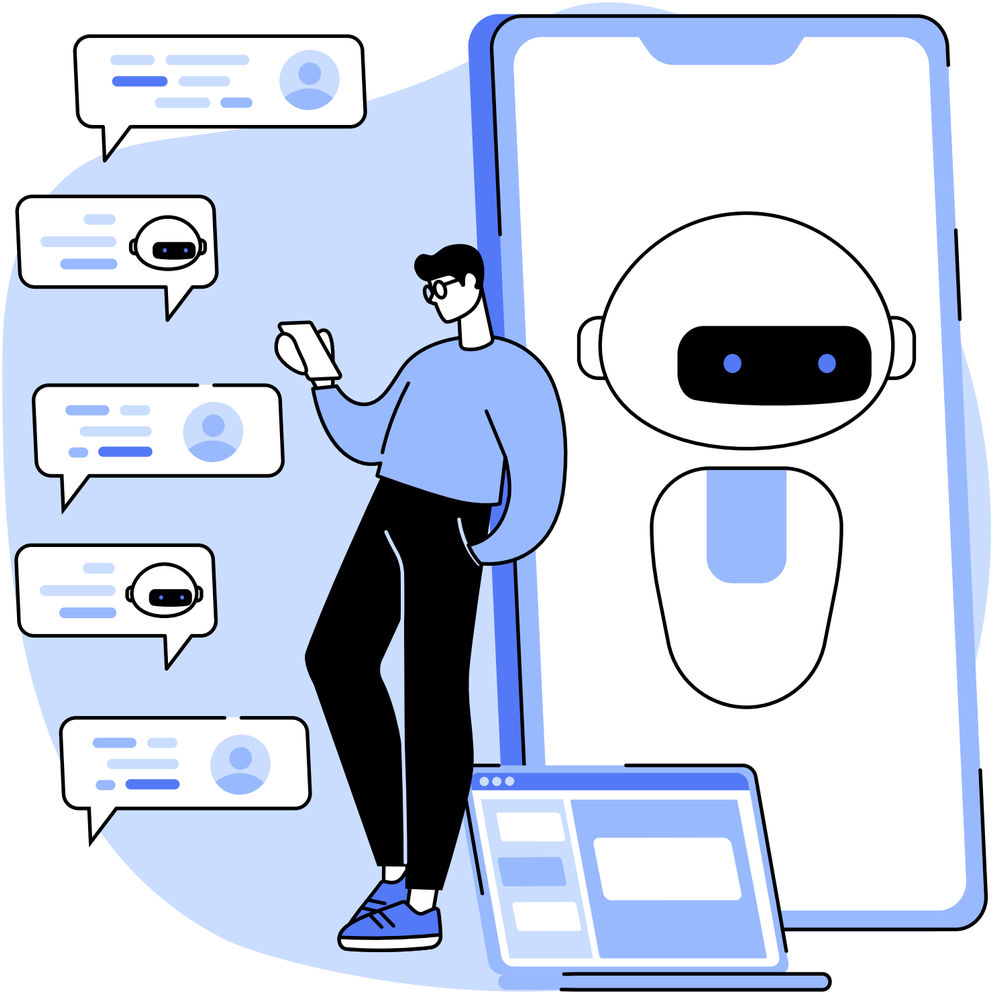
FINETUNING LLM

Finetuning pretrained parameters: Adjusting the weights of pre-trained models on a smaller, task-specific dataset to improve performance on that task.



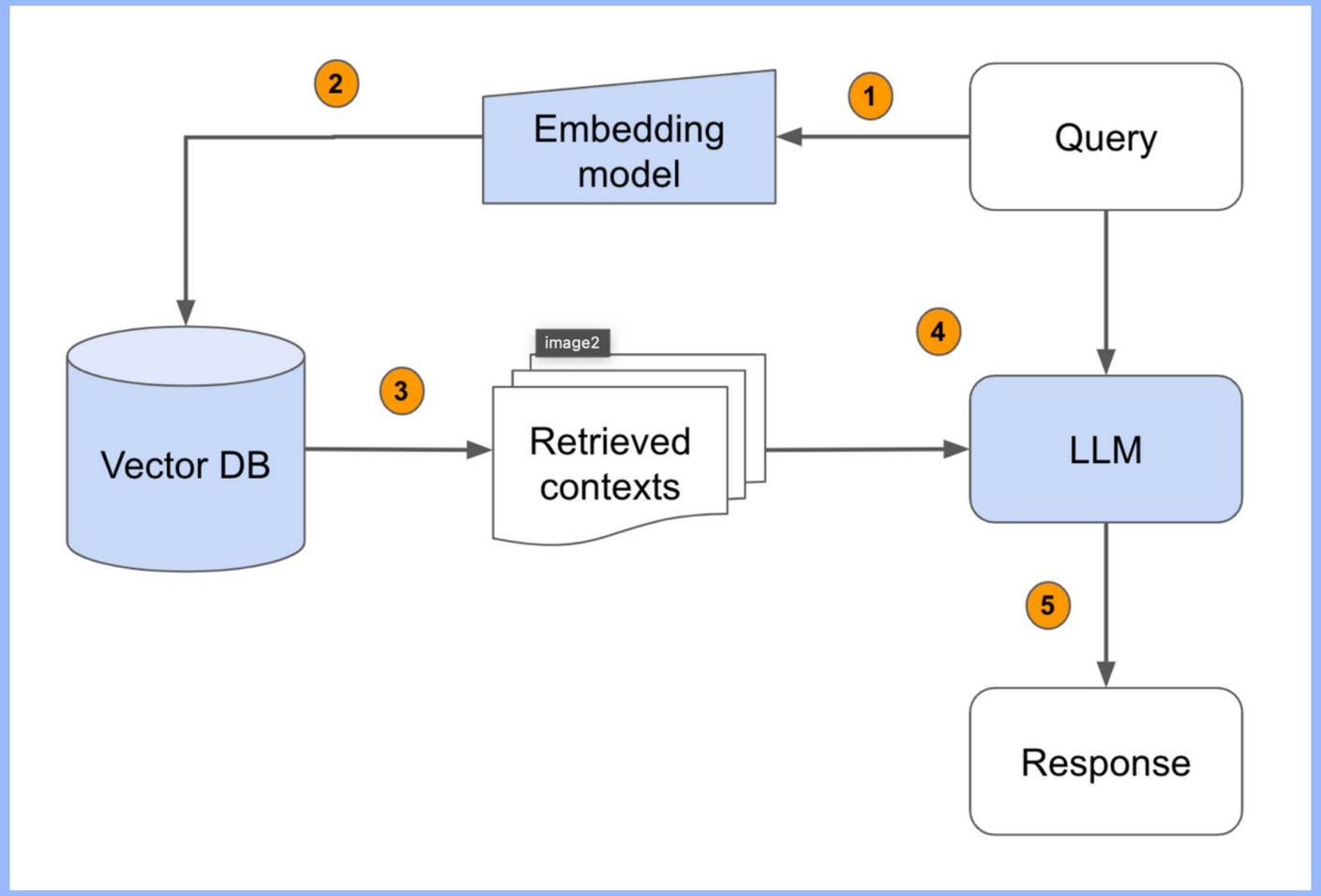
RETRIEVAL AUGMENTED GENERATION

Retrieval-Augmented Generation (RAG) fetches relevant data from outside the foundation model and enhances the input with this data, providing richer context to improve output.



RETRIEVAL AUGMENTED GENERATION

2. Pass the embedded query vector to our vector DB.



1. Pass the query to the embedding model to semantically represent it as an embedded query vector.

3. Retrieve the top-k relevant contexts measured by distance between the query embedding and all the embedded chunks in our knowledge base.

4. Pass the query text and retrieved context text to our LLM.

5. The LLM will generate a response using the provided content.

WHY RAG OVER OTHER RAA'S LIKE REALM

Retrieval Augmented Architectures

Retrieval-Augmented Language Model Pre-Training

1. Due to its excellent open-source documentation and availability.

2. Less computationally expensive due to partial end-to-end training

REFERENCES AND CITATIONS

How To Choose Perfect LLM For The Problem Statement Before Finetuning https://www.labellerr.com/blog/how-to-choose-llm-to-suit-for-use-case/

A Beginner's Guide to Large Language Models https://resources.nvidia.com/en-us-large-language-model-ebooks

Fine-Tuning Llama-2: A Comprehensive Case Study for Tailoring Models to Unique Applications

https://www.anyscale.com/blog/fine-tuning-llama-2-a-comprehensive-case-study-for-tailoring-models-to-unique-applications

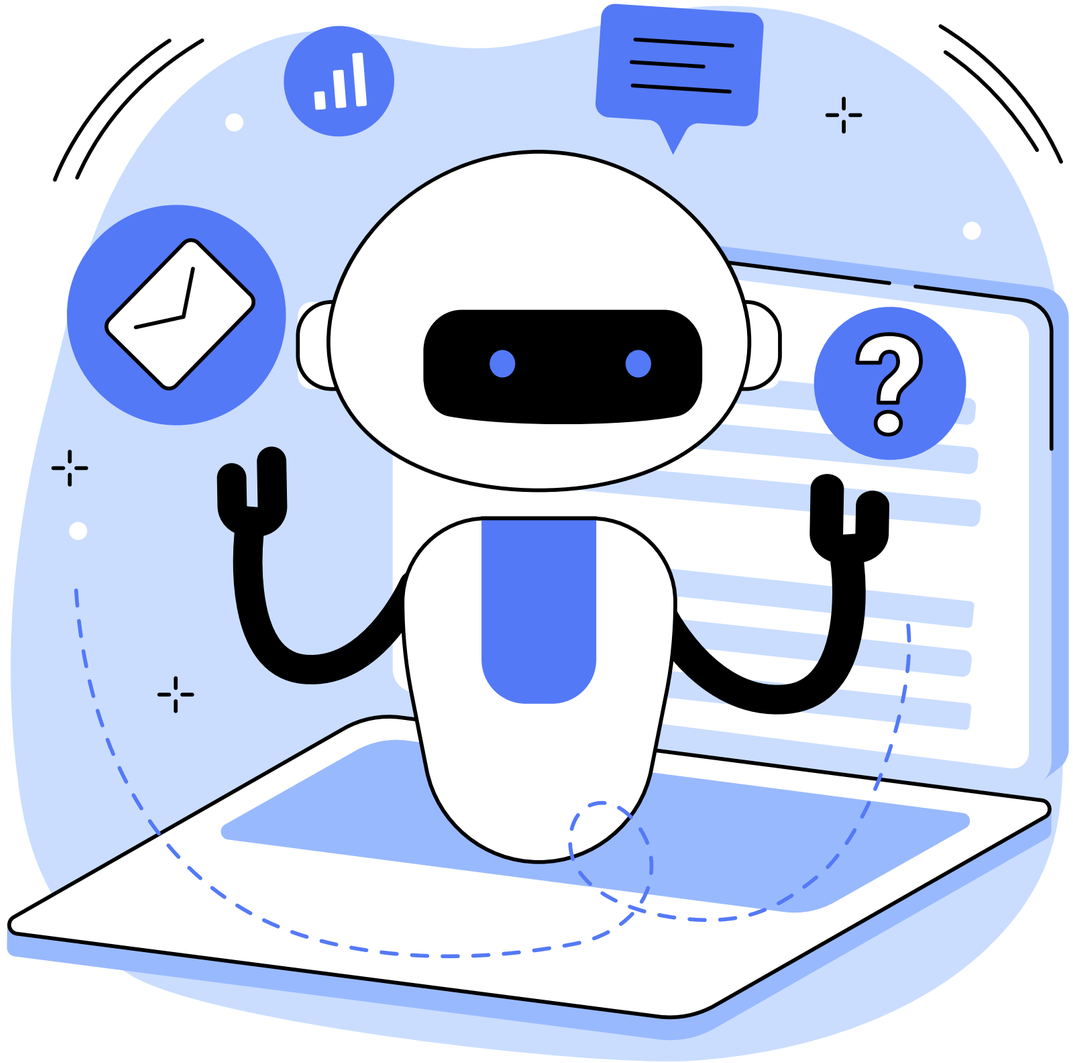
Building RAG-based LLM Applications for Production (Part 1)

https://www.anyscale.com/blog/a-comprehensive-guide-for-building-rag-based-llm-applications-part-1

DATA COLLECTION

Considerations

- Authentic
- Accurate
- Reliable



DATA TYPE?

Text

ETHICAL ISSUES?

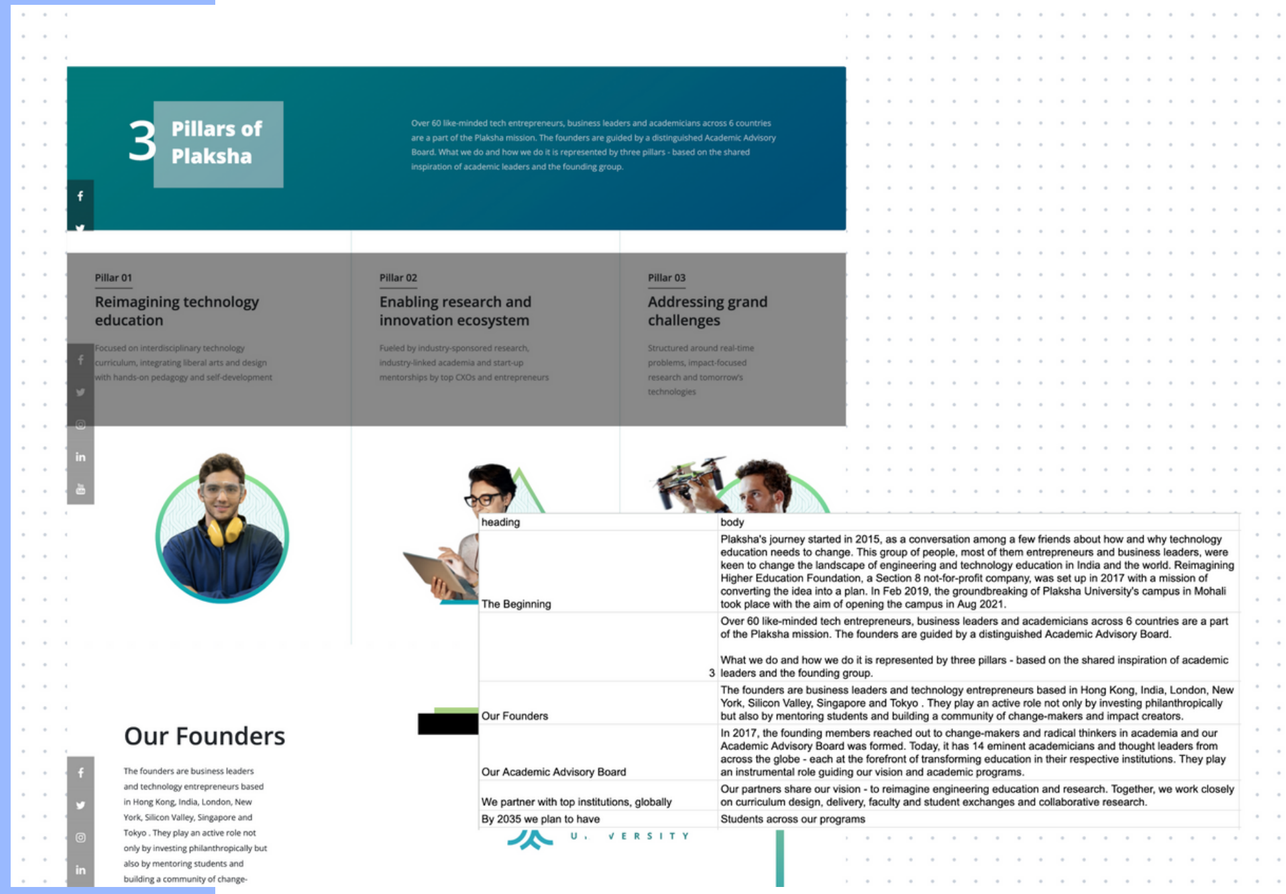
Not Really

DATA COLLECTION: HOW?

1. Web Scraping Tool
2. Manually

DATA COLLECTION: HOW MUCH?

1. Website: 222 Data Points
2. QnA Pairs: ~ 1000



heading	body
The Beginning	Plaksha's journey started in 2015, as a conversation among a few friends about how and why technology education needs to change. This group of people, most of them entrepreneurs and business leaders, were keen to change the landscape of engineering and technology education in India and the world. Reimagining Higher Education Foundation, a Section 8 not-for-profit company, was set up in 2017 with a mission of converting the idea into a plan. In Feb 2019, the groundbreaking of Plaksha University's campus in Mohali took place with the aim of opening the campus in Aug 2021.
3	Over 60 like-minded tech entrepreneurs, business leaders and academicians across 6 countries are a part of the Plaksha mission. The founders are guided by a distinguished Academic Advisory Board. What we do and how we do it is represented by three pillars - based on the shared inspiration of academic leaders and the founding group.
Our Founders	The founders are business leaders and technology entrepreneurs based in Hong Kong, India, London, New York, Silicon Valley, Singapore and Tokyo . They play an active role not only by investing philanthropically but also by mentoring students and building a community of change-makers and impact creators.
Our Academic Advisory Board	In 2017, the founding members reached out to change-makers and radical thinkers in academia and our Academic Advisory Board was formed. Today, it has 14 eminent academicians and thought leaders from across the globe - each at the forefront of transforming education in their respective institutions. They play an instrumental role guiding our vision and academic programs.
We partner with top institutions, globally	Our partners share our vision - to reimagine engineering education and research. Together, we work closely on curriculum design, delivery, faculty and student exchanges and collaborative research.
By 2035 we plan to have	Students across our programs

DATA PREPROCESSING

Chunking

Synonym Augmentation

Text to vector conversion

ML METHODOLOGY

Retrieval Models

Generation Models

Integration Techniques

Implemented Guardrails

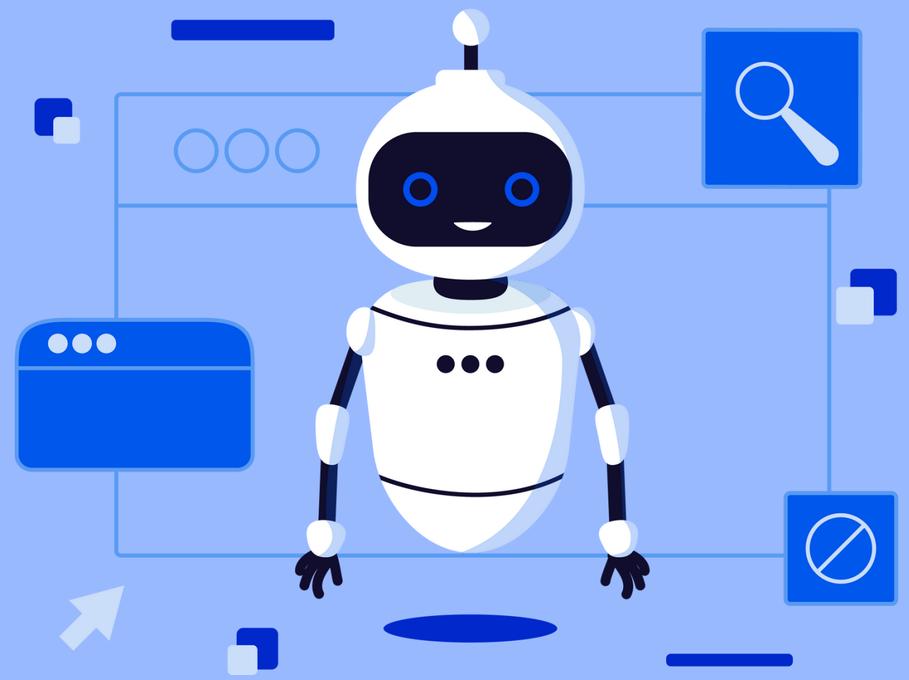
Domain Specific Adaptation

CHALLENGES

1. Hardware: Unavailability of computational resources

2. Software: Lack of proper documentation

PERFORMANCE METRICS



GOOGLE PALM

Semantic Similarity Score

84.2%

Rouge-1 Score

38.05%

DEPLOYABILITY CHALLENGES

Increased Latency

Resource Limitations

Model Optimization

THANK YOU FOR LISTENING!

Feedback and questions, please:)

We’re grateful to Prof. Siddharth for challenging us, and pushing us to more.