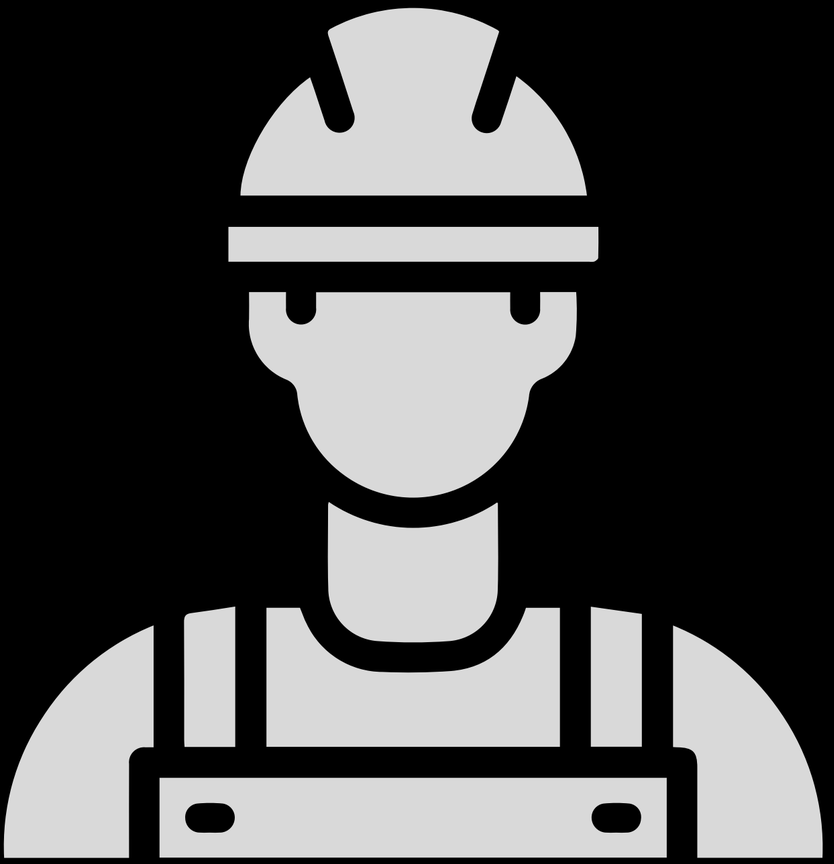
Theme: Building with LLMs

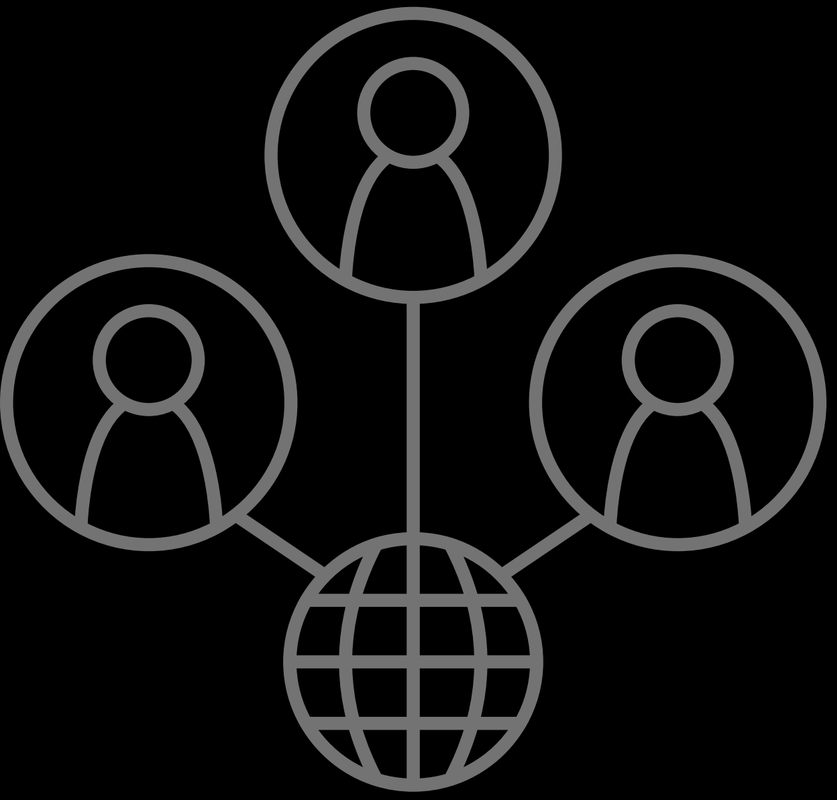
Generative AI Tanzania

By Kalebu Gwalugano

Why this community ?



Build capacity of people who are builders and not just consumers of Generative AI



A safe place for AI enthusiasts to share and network together

Generative AI Tanzania

Our community has now about 111 members on WhatsApp group

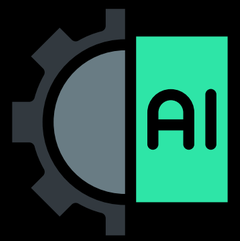
https://bit.ly/44XpwKN

We are just getting started

1st
Theme: Workshop on Prompt Engineering

2nd
Theme: Anything generative AI

3rd
Theme: Building with LLMs



Any Question up to here?

LLMs?

Large language models (LLMs) are deep learning algorithms that can recognize, summarize, translate, predict, and generate content using very large datasets.

source: https://www.nvidia.com/en-us/glossary/data-science/large-language-models/

Llama-2

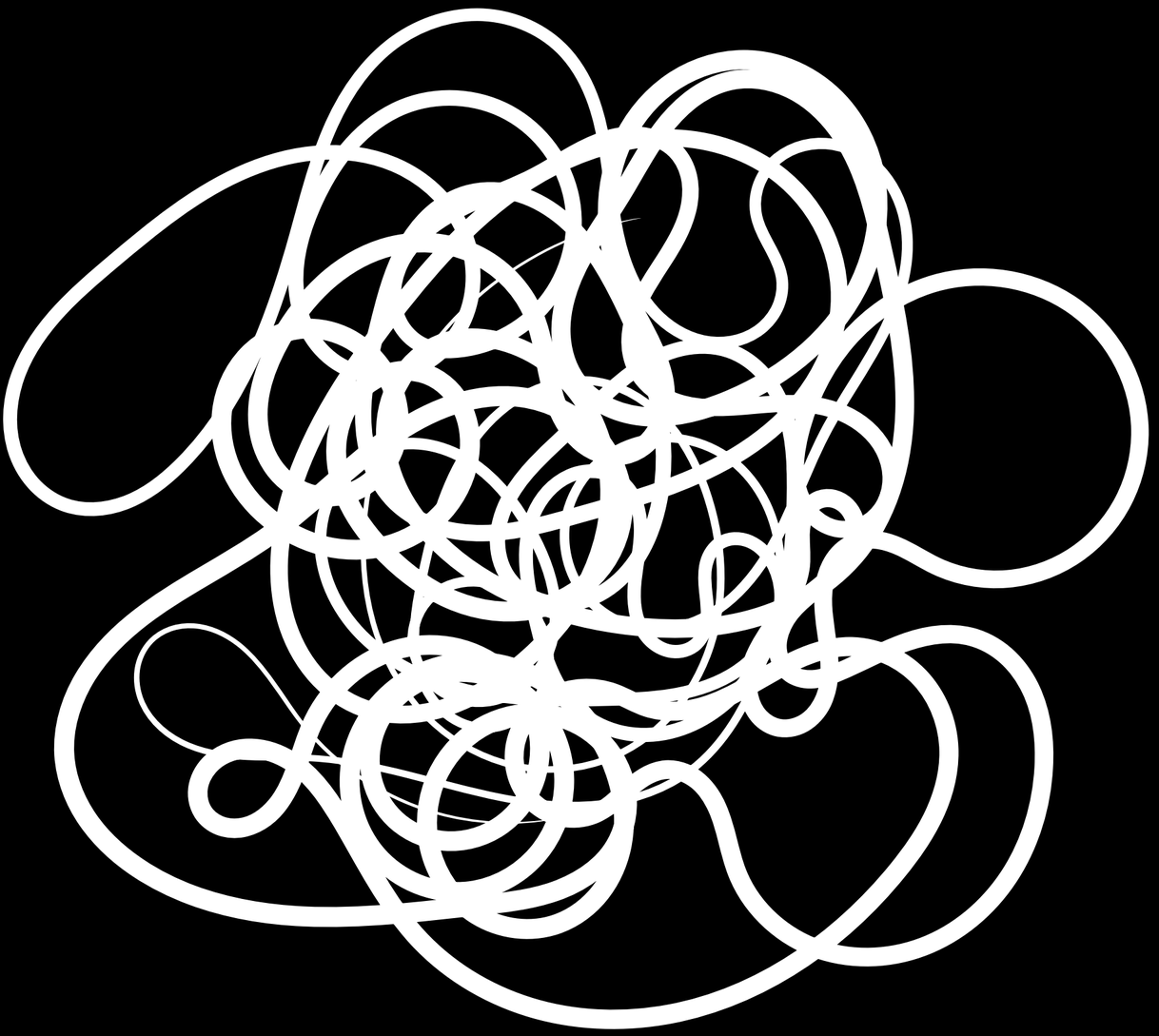
2 trillions tokens

Falcon

3.5-trillion tokens

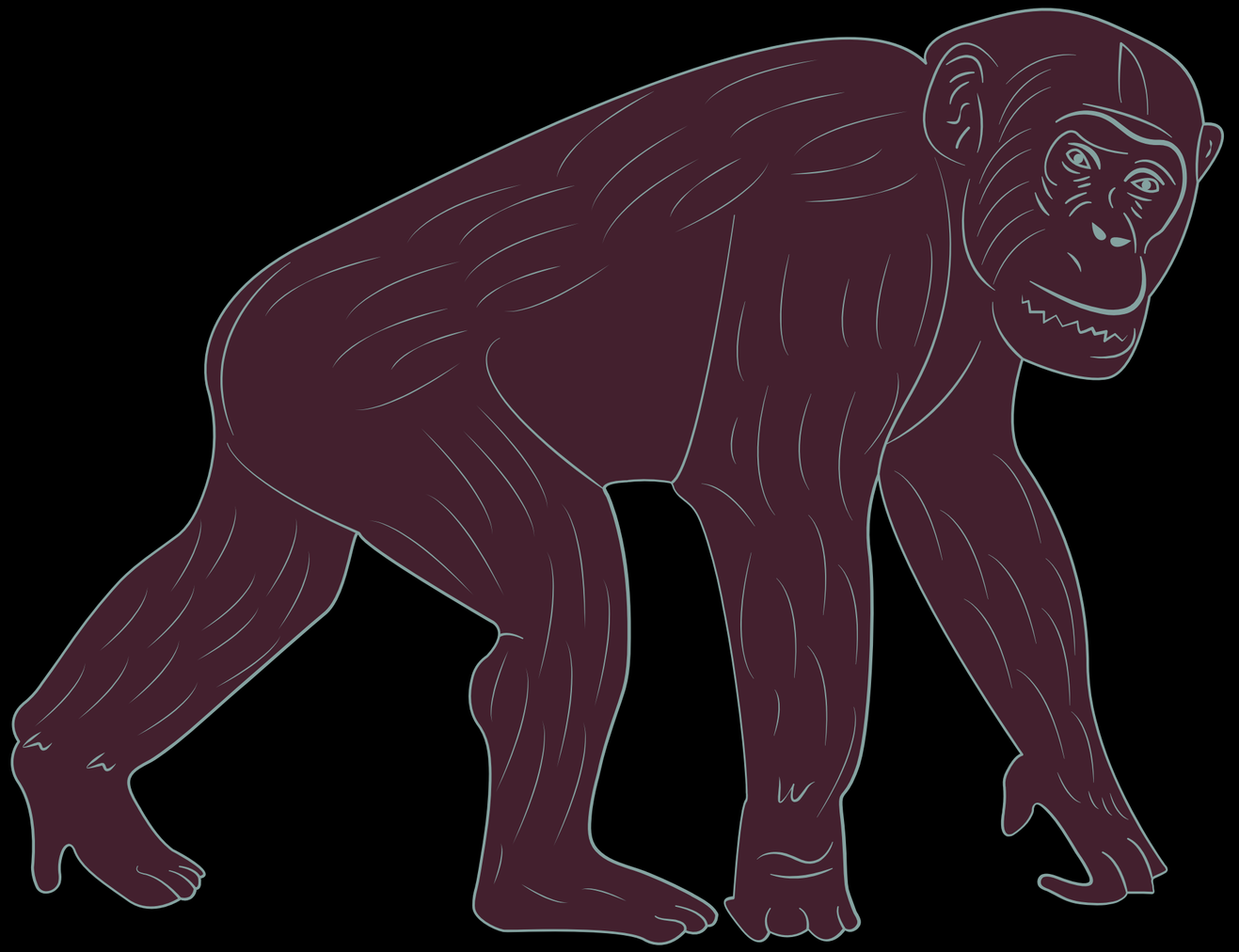
GPT-3

570 gigabytes of text



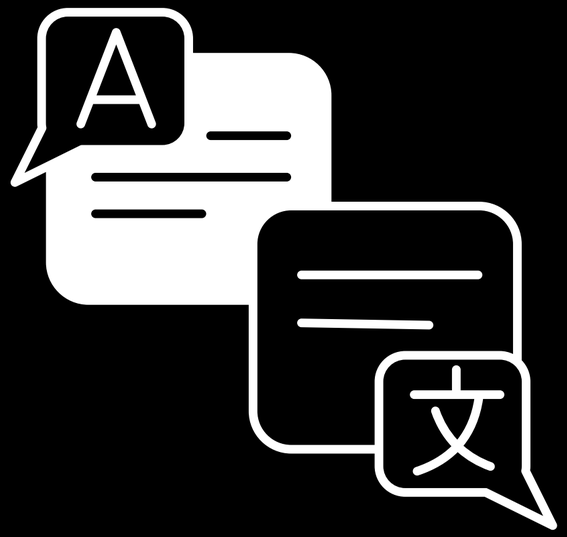
Tokens ?
How does they work ?

Trust me you will get it



Life before LLMS ?

Traditional NLP vs LLMs



Traditional Natural language processing is focused on understanding language by it's individual components

LLMs try to understand the language as a whole

Traditional NLP

Part of Speech tagging (POS)

Lemmatization

Dependency parsing

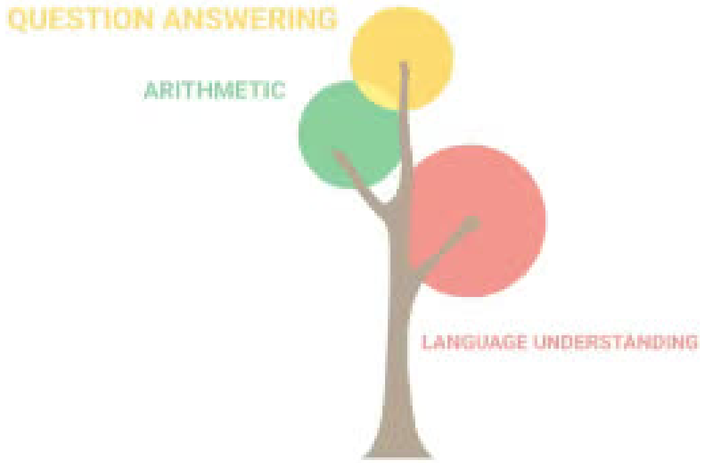
Stemming

Stopwords Analysis

How do LLMs learn?

I’m learning a new _____ language for building AI systems

LLM emergent behaviors



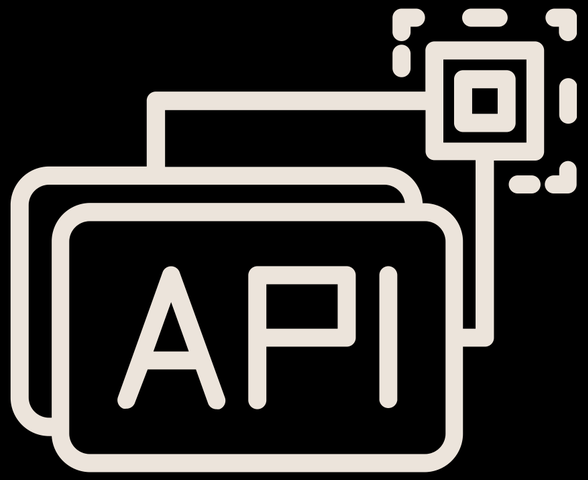
As Language Models grow, novel abilities suddenly appear [8]

What can you do with LLMS ?

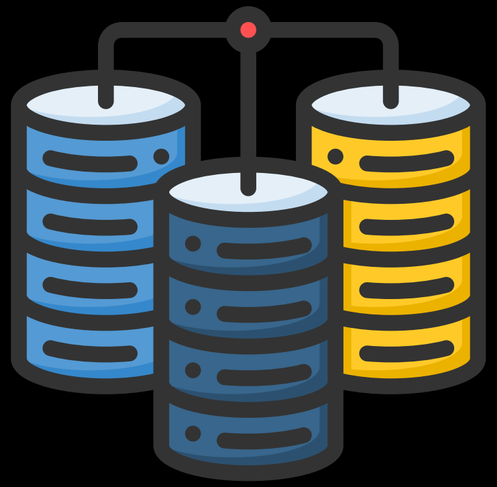
- Semantic Search
- Chatbots Application
- AI assistants
- Automate workflows
- So much more

The limit is your own imagination

Good, Now how can I get started?

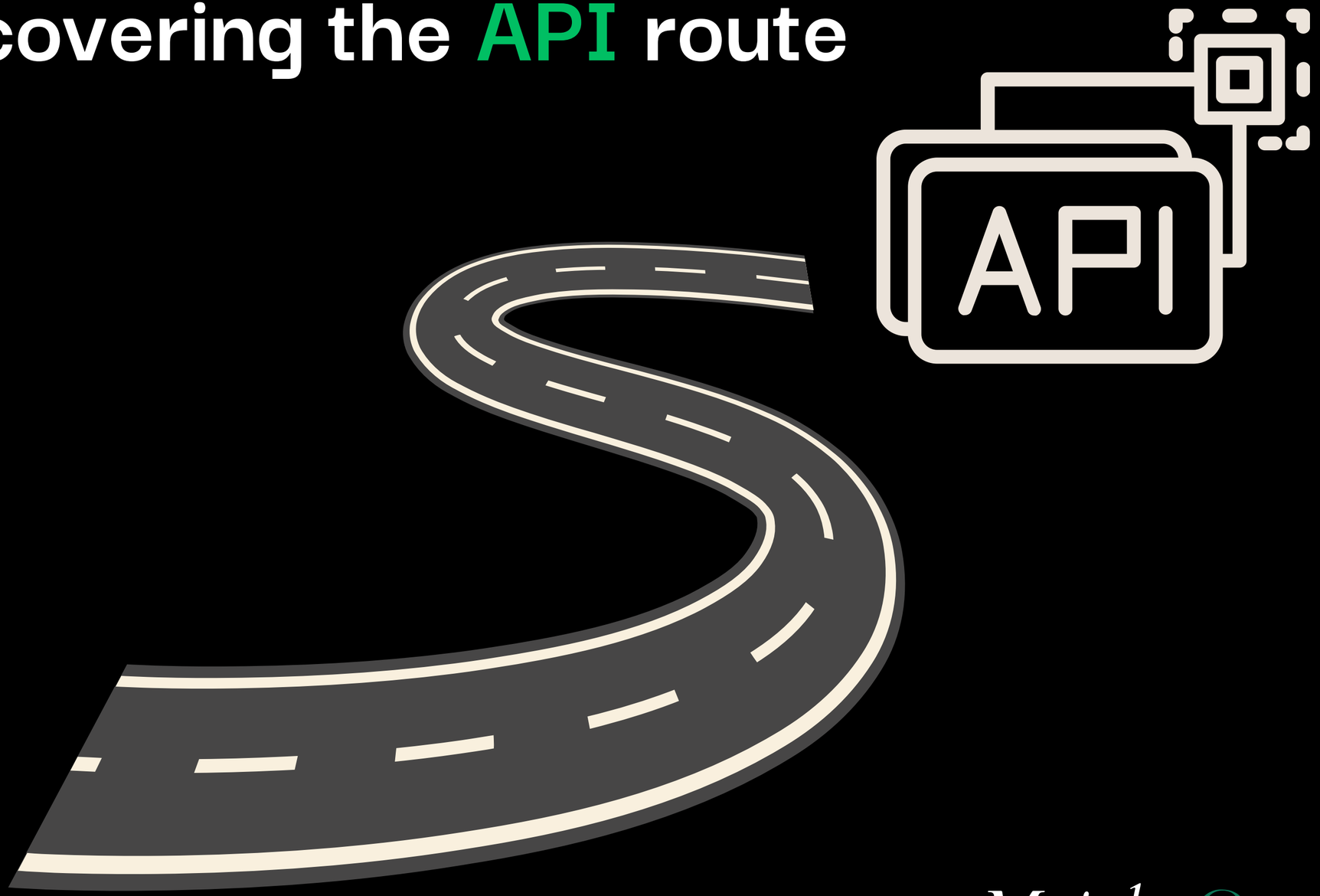


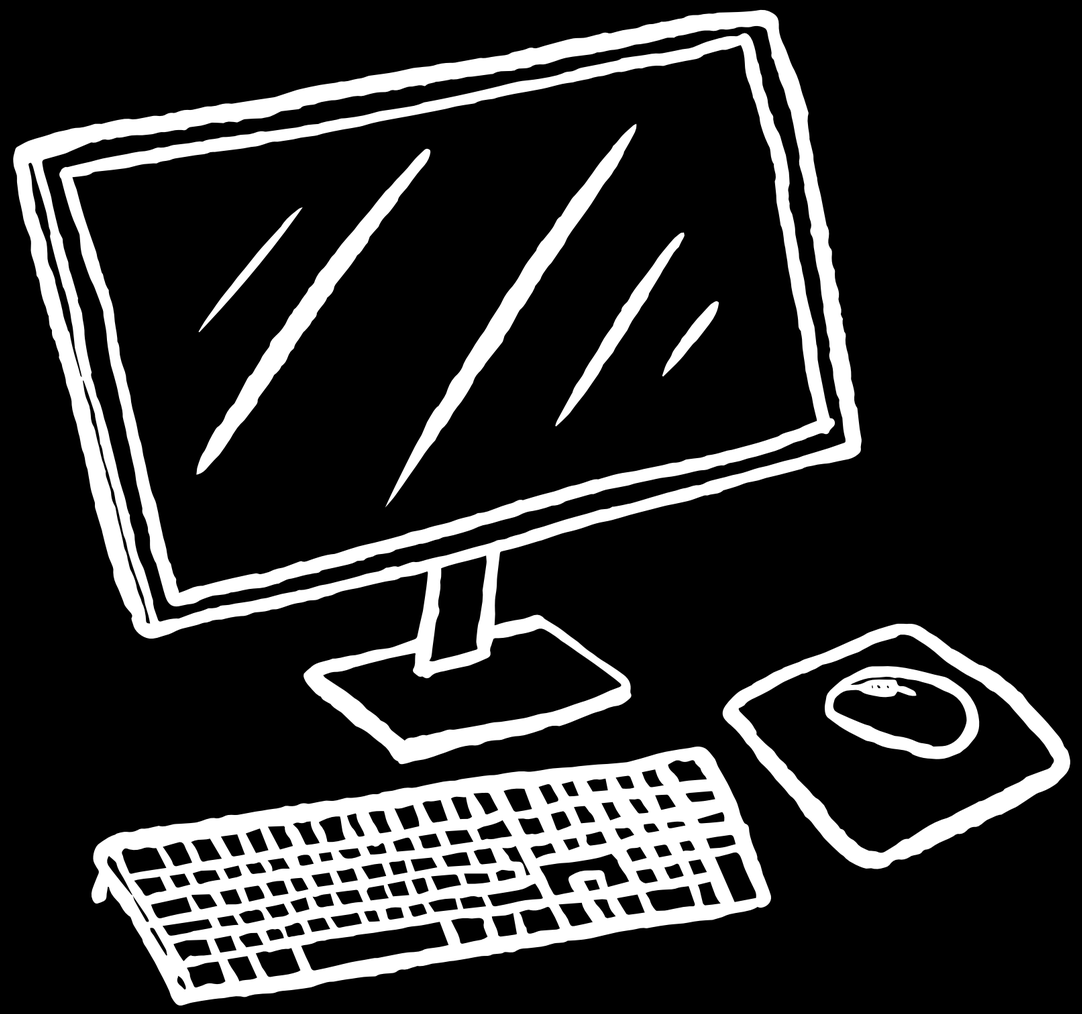
Use API from LLM provider like OpenAI, Cohere, Google, Anthropic and so on



You can either run load the entire models

I will be covering the API route





Everyone will have their private AI node that is free from third parties,
constantly feeding your data from a variety of angles

Terminologies

prompt engineering

Temperature

Embeddings

Tokens

Hallucination

Guardrails

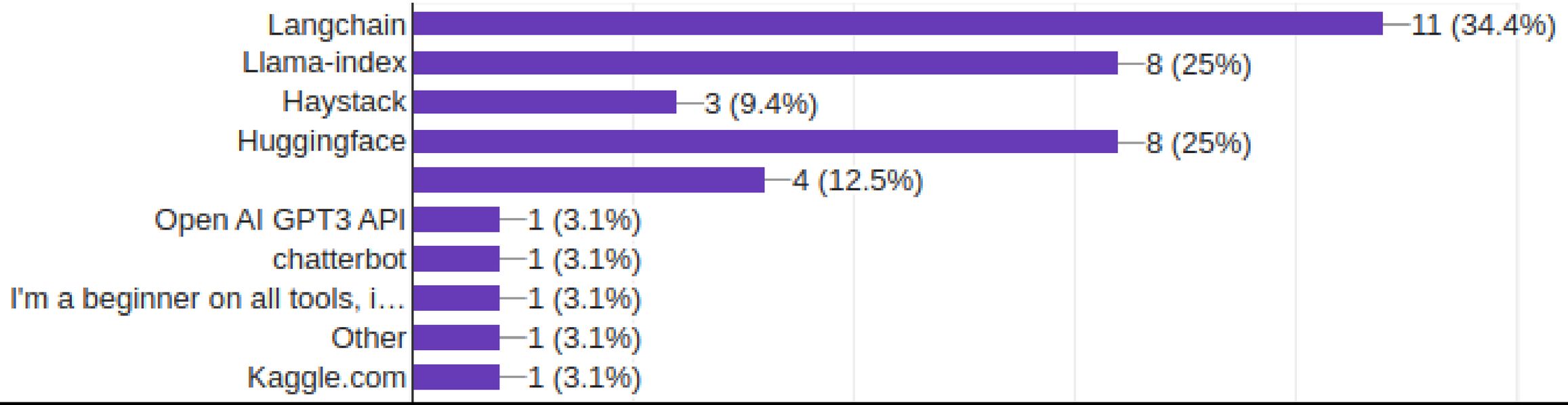
Tools we are going to look

- Langchain
- LLama-index
- hugging-face
- haystack

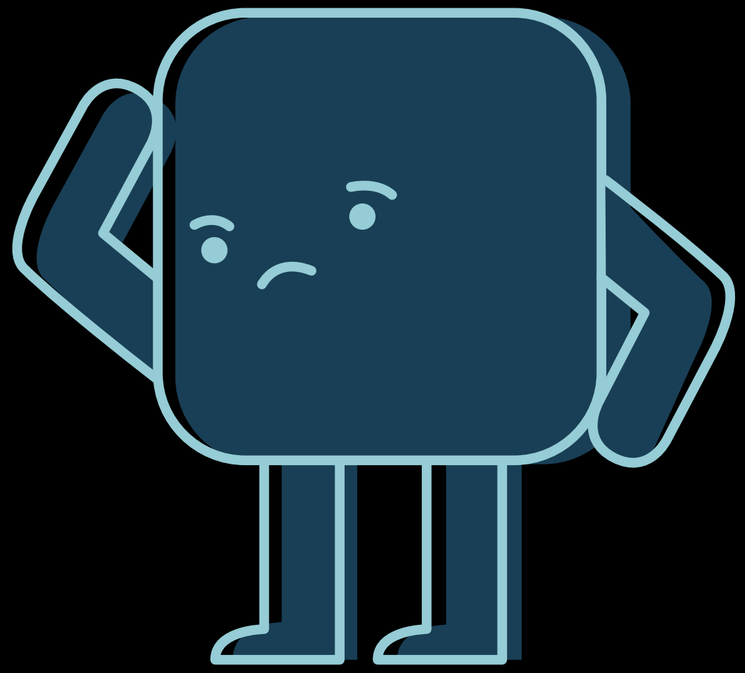
Which of the following tools have you previously worked with or explored? (Multiple choices: Langchain, Llama-Index, Haystack, Hugging Face, Others - Please specify)

Copy

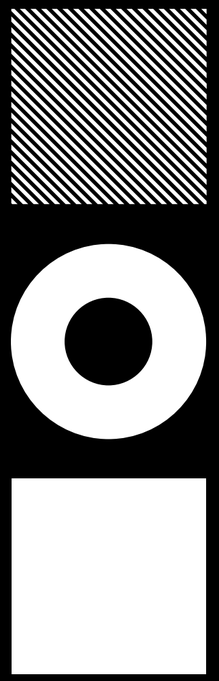
32 responses



Survey Response



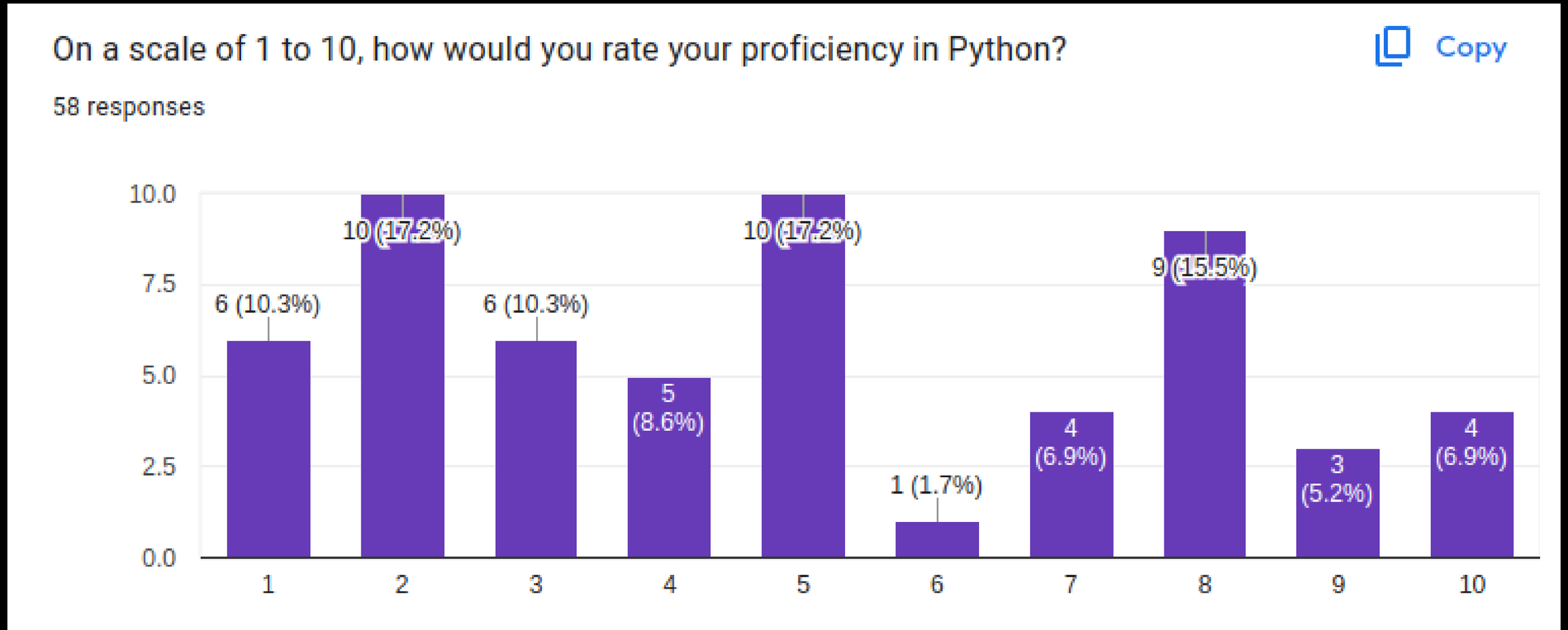
What do these tools actually do?



Think of them as web frameworks but for AI applications, They are architected in a way that simplifies adding complex features like memory, external knowledge bases and so on

—

If your use case is too simple, you might not need them at all

We are going to use Python

You should be able to easily replicate across other languages