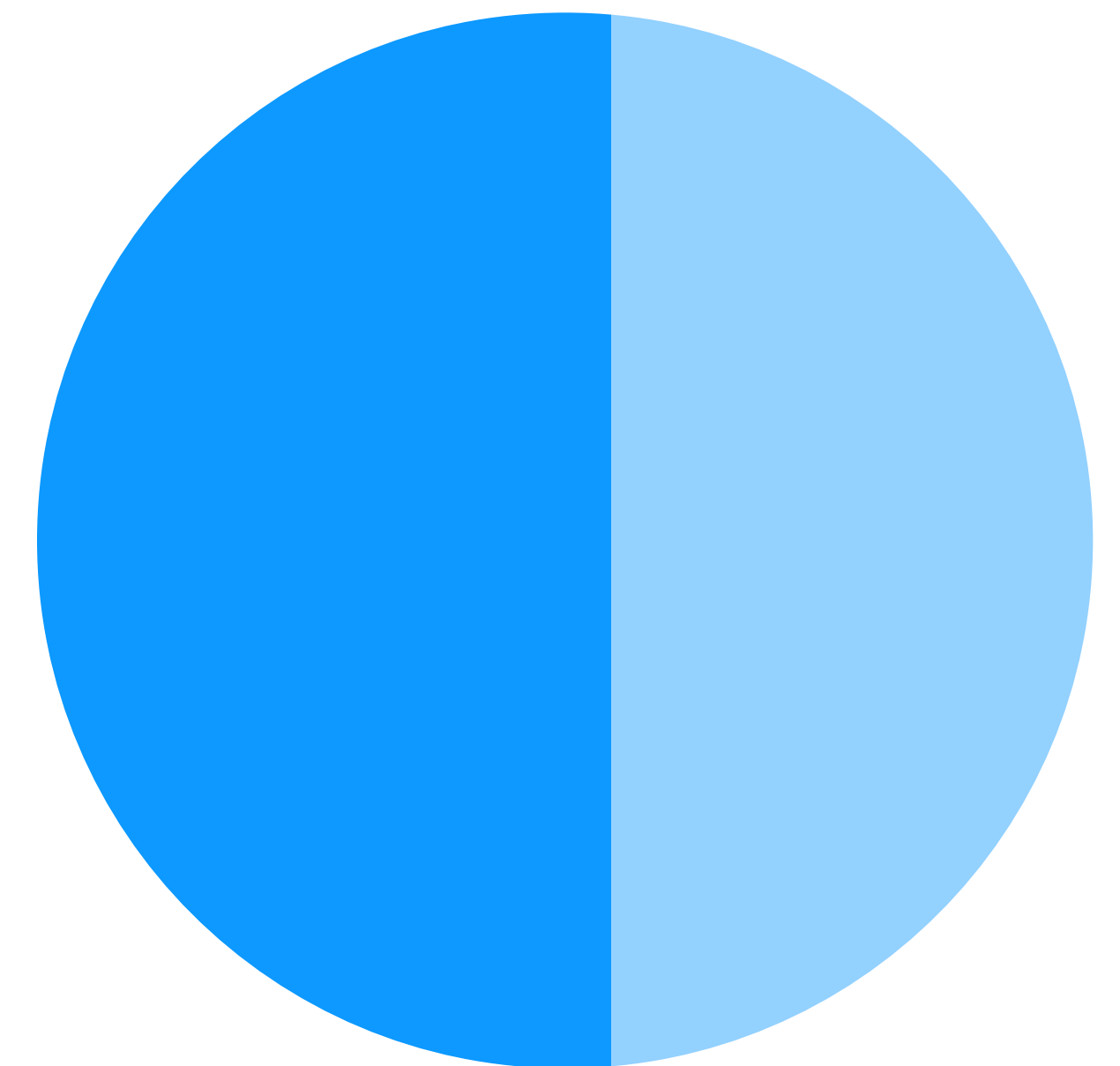


Sprint 4 - Fine tuning, RAG e Front-end

Mockingjay [Brastel Co., Ltd.](#)



Recapitulando

- Análise exploratória e pré-processamento dos dados

SPRINT 1
16 AGO

- Vetorização, data augmentation e modelo de classificação de intenções

SPRINT 2
30 AGO

- Modelo LSTM, aprimoramento da pipeline de pré-processamento, Back-end e Front-end

SPRINT 3
13 SET

- Modelo de Geração de Respostas (LLM)

SPRINT 4
27 SET

- Implementação final

SPRINT 5
11 OUT

LLMs - Possíveis abordagens

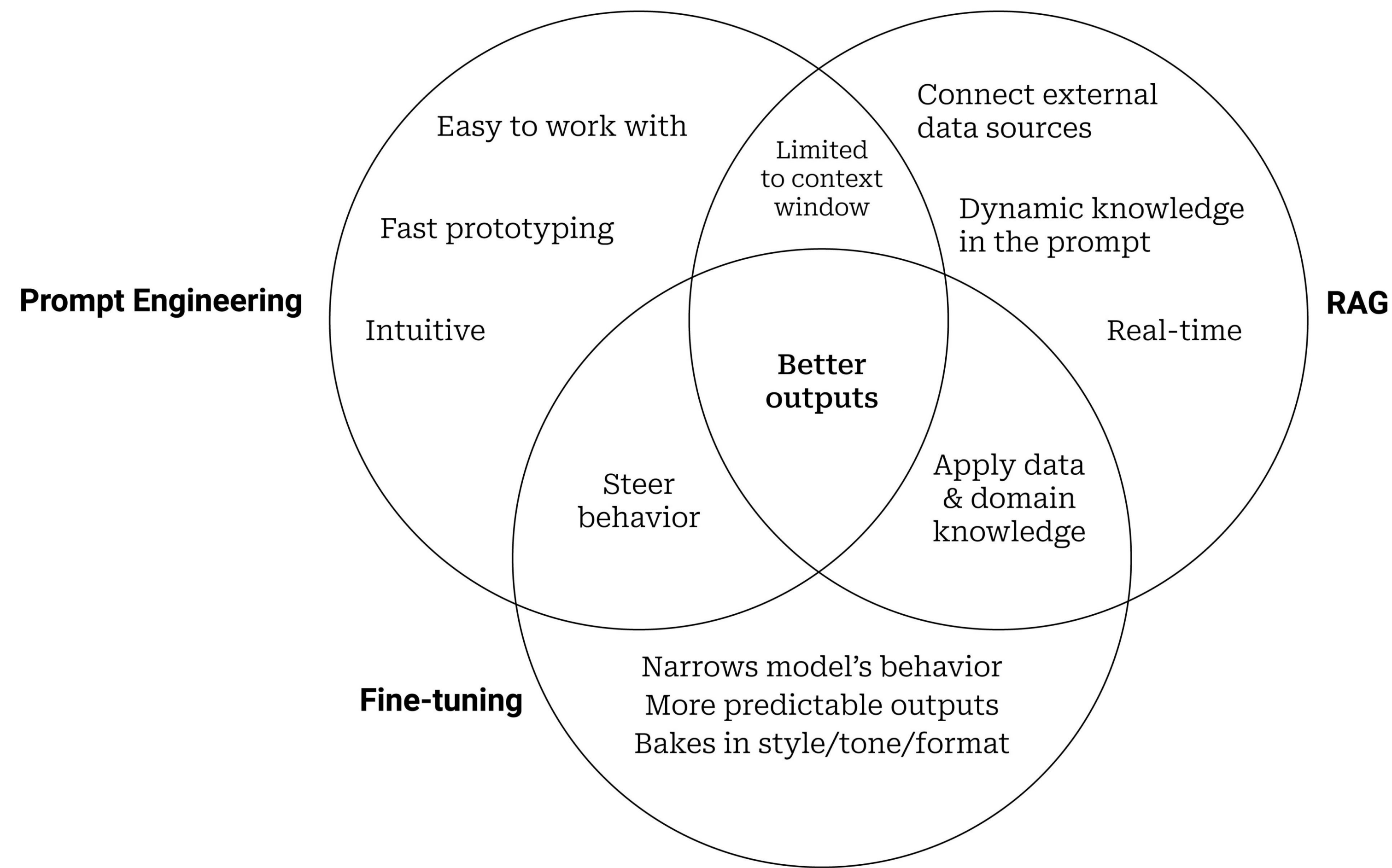


Figura 1 – Ilustração - Abordagens com LLMs
Disponível em <https://www.prompthub.us/blog/fine-tuning-vs-prompt-engineering>

Prompt Engineering



Figura 2 – Ilustração - Técnicas de Prompt Engineering
Disponível em <https://www.scribbledata.io/blog/prompt-engineering-introduction-techniques-limits-and-future-perspective/>

RAG

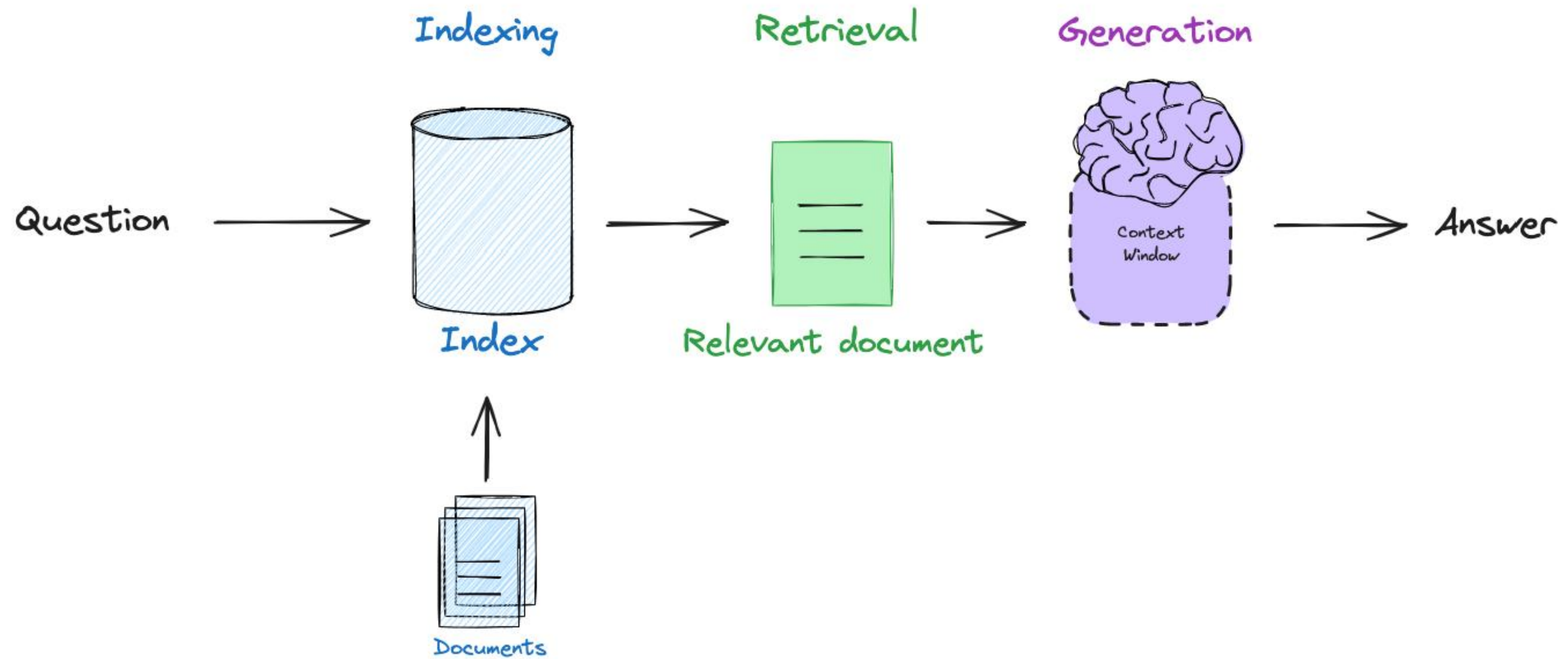


Figura 3 – Ilustração - Fluxo RAG
Disponível em <https://blog.langchain.dev/agent-rag-with-langgraph/>

LLM Fine Tuning

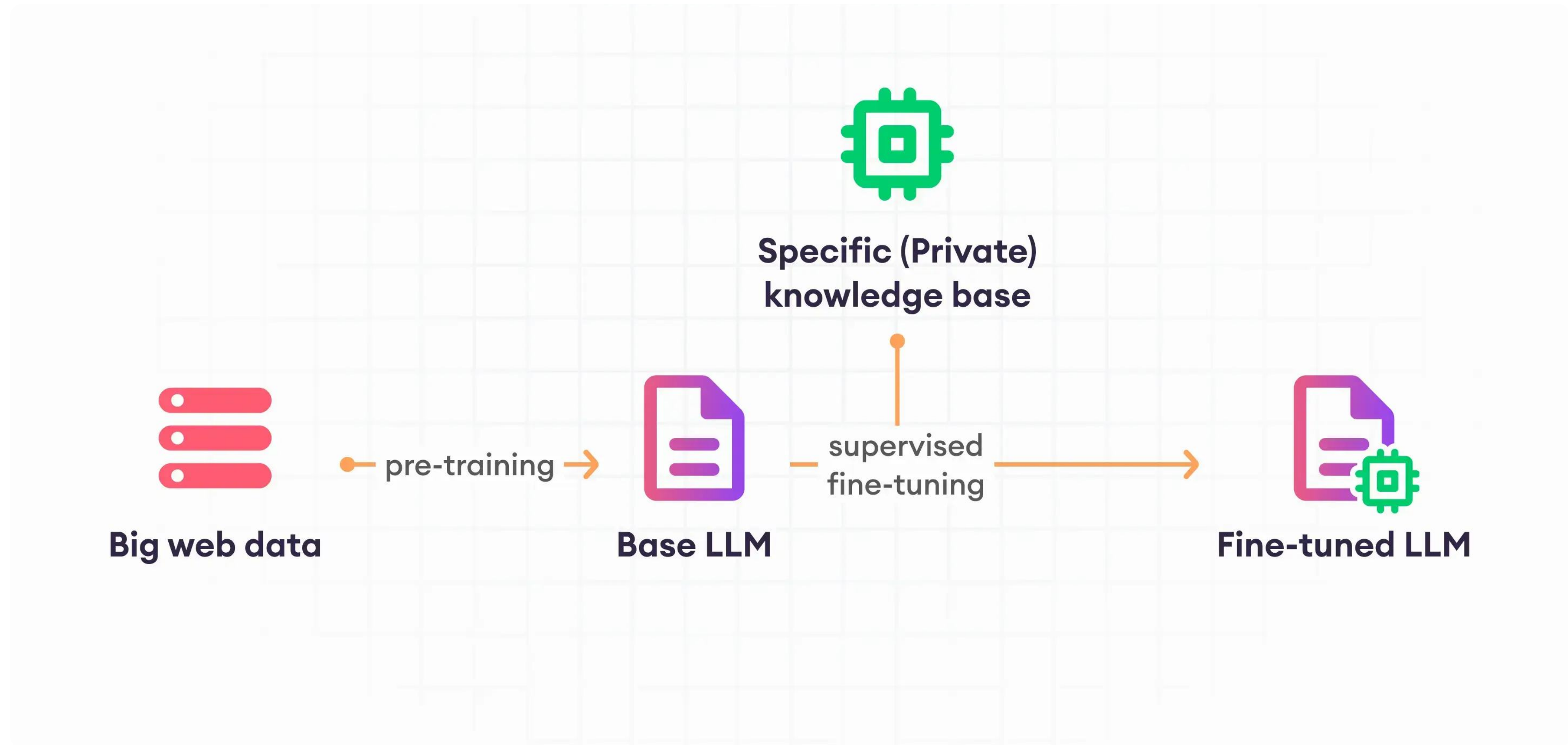


Figura 3 – Ilustração - LLM Fine Tuning
Disponível em <https://www.superannotate.com/blog/llm-fine-tuning>

RAG vs Fine Tuning

	RAG	Finetuning
External knowledge req'd?	✓	✗
Changing model behaviour req'd?	✗	✓
Minimise hallucinations?	✓	✗
Training data available?	✗	✓
Is data (mostly) dynamic?	✓	✗
Interpretability req'd?	✓	✗

Figura 4 – Ilustração - RAG vs Fine Tuning
Disponível em <https://www.ampcome.com/articles/rag-vs-finetuning-which-is-best-for-your-llm-applications>

Abordagens adotadas no projeto

Figura 4 – Ilustração - RAG vs Fine Tuning
Disponível em <https://www.ampcome.com/articles/rag-vs-finetuning-which-is-best-for-your-llm-applications>

Live demo

Demonstração da solução construída.

Próximos passos



- Análise exploratória e pré-processamento dos dados

SPRINT 1
16 AGO

- Vetorização, data augmentation e modelo de classificação de intenções

SPRINT 2
30 AGO

- Modelo LSTM, aprimoramento da pipeline de pré-processamento, Back-end e Front-end

SPRINT 3
13 SET

- Modelo de Geração de Respostas (LLM)

SPRINT 4
27 SET

- Implementação final

SPRINT 5
11 OUT

13 de setembro de 2024

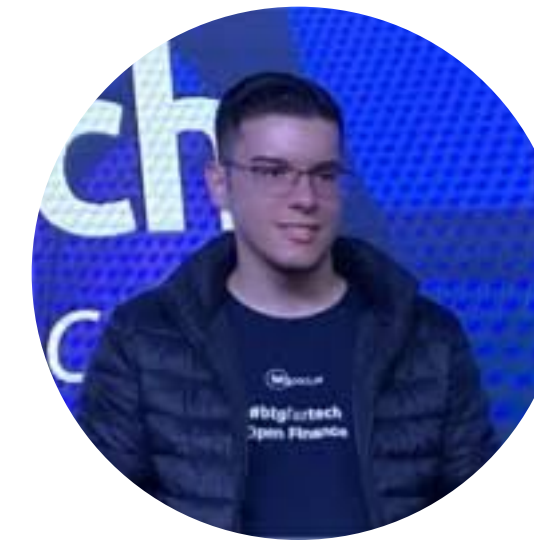
Instituto de Tecnologia e Liderança

Nosso time

Grupo Mockingjay



Allan



Elias



Giovana



Rafael



Gábrio



Cristiane



Melyssa

13 de setembro de 2024

Instituto de Tecnologia e Liderança

Obrigado.