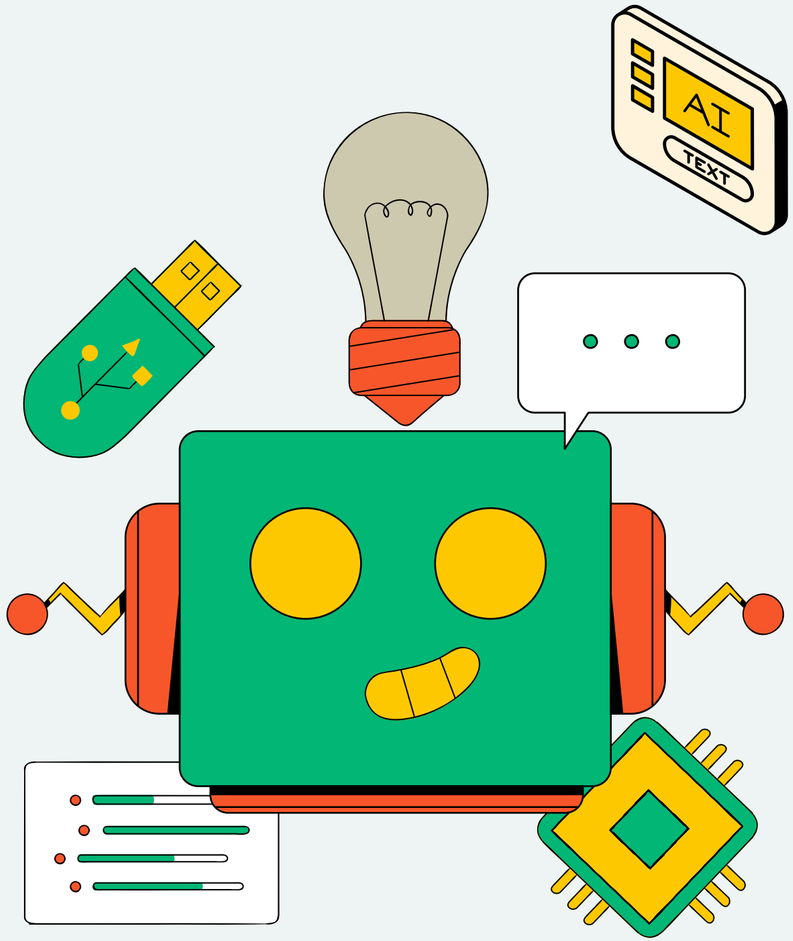
THYNK UNLIMITED

WE LEARN FOR THE FUTURE

MULTI-AGENT SYSTEM FOR QUESTION ANSWERING



PRESENTATION

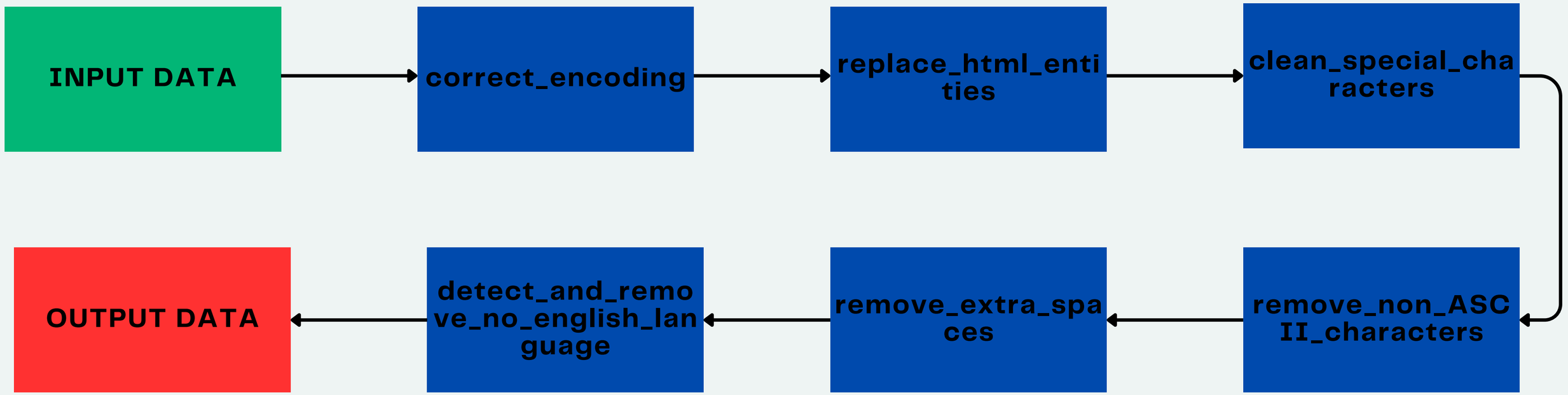
PRÉSENTÉ PAR :

YASSINE ZITOUNI

PRESENTATION OUTLINE

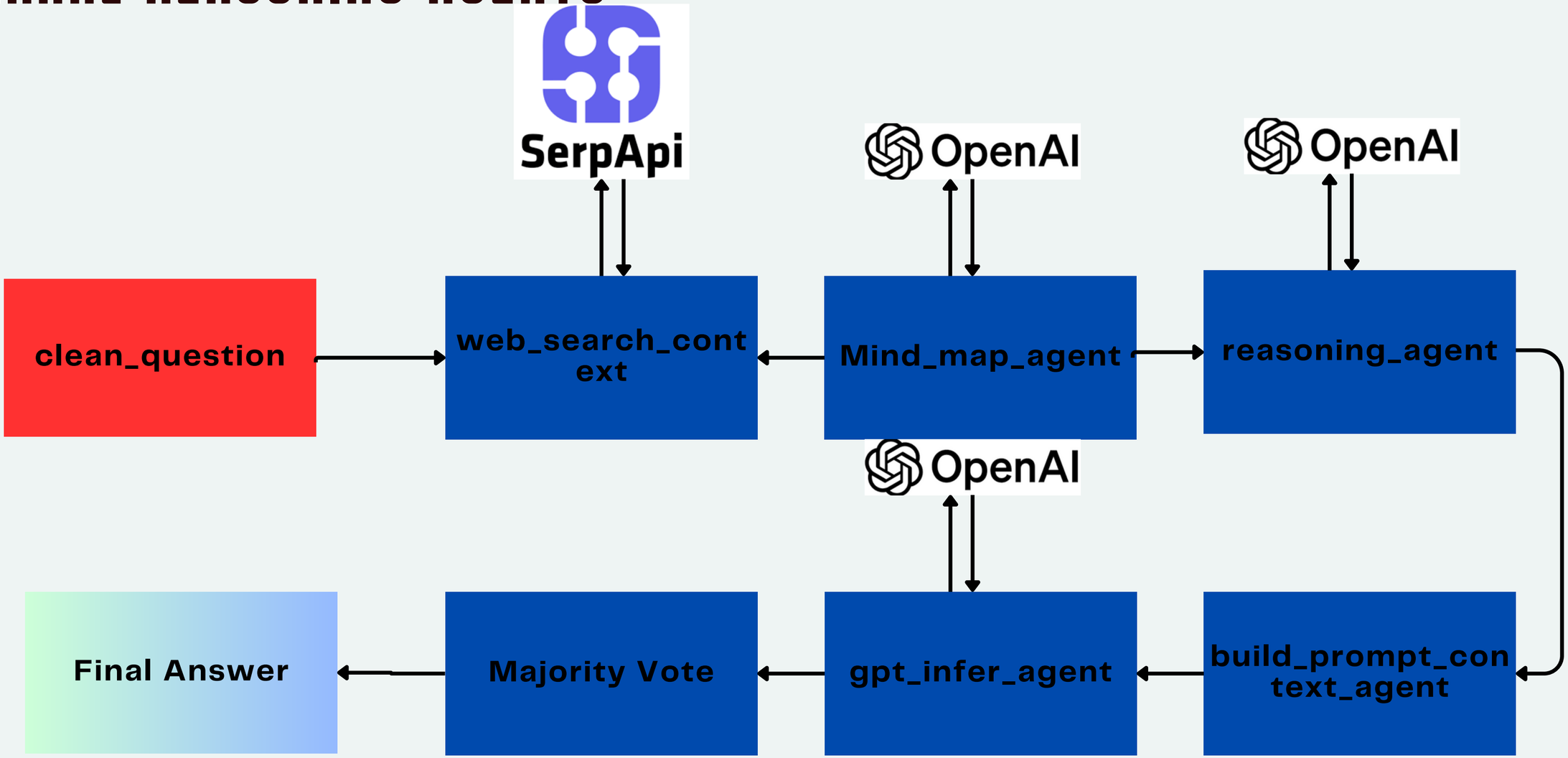
- Data Exploration & Preprocessing
- High-Performance Benchmark: GPT-4.5 with External Reasoning Agents
- Comparative Study of Direct Reasoning
- Fine-Tuning a Lightweight Model: T5 in Action
- Proposed Enhancements for Low-Cost Performance Boost

INTRODUCTION



HIGH-PERFORMANCE BENCHMARK: GPT-4.5 WITH EXTERNAL REASONING AGENTS

Model Name	Public score	Private score	Test Data Accuracy
GPT4.5	0.75	0.75	0.75



. COMPARATIVE STUDY OF DIRECT REASONING

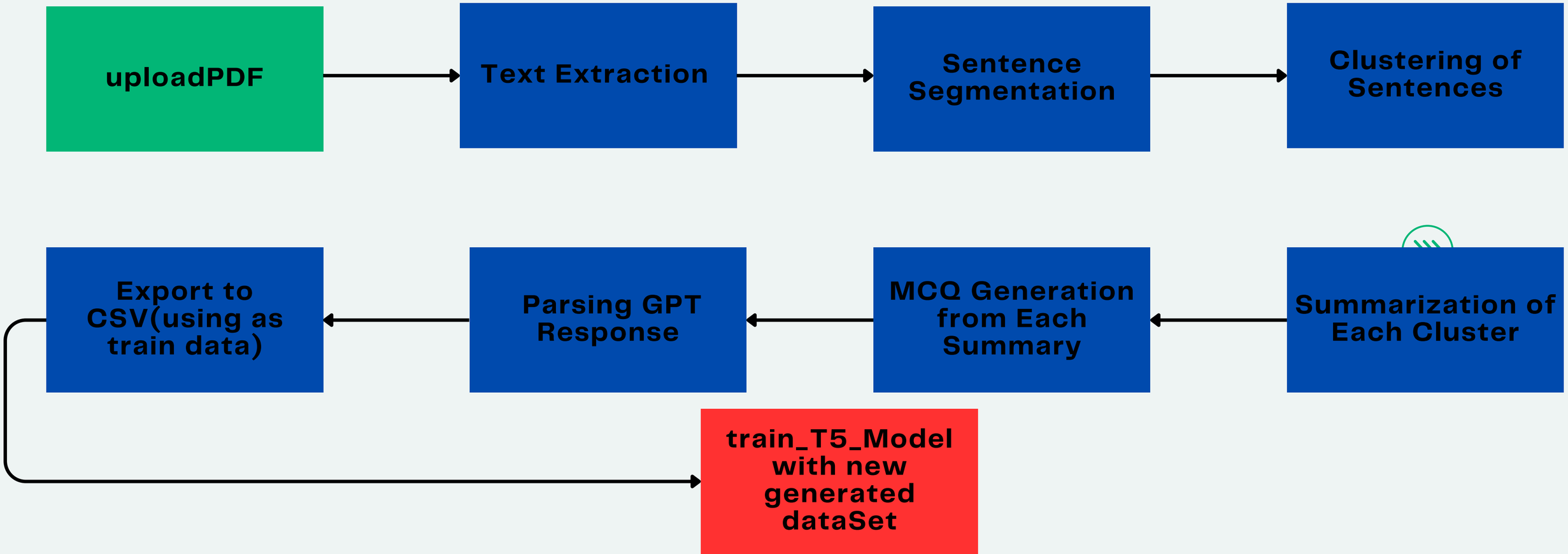
Model Name	Train Data Accuracy	Public score	Private score	Test Data Accuracy
GPT-3.5-TURBO	0.44	0.75	0.4333	0.5915
GPT-4	0.48	0.5833	0.4833	0.533
GPT-4-TURBO	0.48	0.600	0.51666	0.558
GPT-4O-MINI	0.44	0.5888	0.5666	0.5777
GPT-4O	0.48	0.533	0.588	0.56
O1	0.48	0.488	0.588	0.538

FINE-TUNING A LIGHTWEIGHT MODEL: T5 IN ACTION

- input data
- construcut few-shot prompt
- tokenisation
- Fine-tuning with Seq2SeqTrainer
- generate prediction
- output

Modele	public score	private score	learning rate	epochs	zero shot	one shot	few shot
google/flan-t5-small	0,4166	0,5833	3,00E-05	30	no	no	yes

PROPOSED ENHANCEMENTS FOR LOW-COST PERFORMANCE BOOST



📚 MCQ Generator from PDF

📤 Upload one or more PDF files

Drag and drop files here
Limit 200MB per file • PDF

Browse files

s41598-025-96043-1.pdf 4.0MB

📄 Uploaded files:

✅ s41598-025-96043-1.pdf

📖 Preview of extracted text

✅ MCQs successfully generated!

	question	option_A
0	What does a θ value greater than 1 represent in the context of failure types?	A. Failures due to wear.
1	What are the purpose of expanding the Weibull distribution family according to the re	A. To make it more comp
2	What does the resilient estimation technique for the three-parameter Weibull distribu	A. Probability integral tra
3	What software version was used for simulations in the statistical modeling and reliab	A. R 4.4.131

📥 Download MCQs in CSV format

Resources

Agentic Reasoning : Reasoning LLMs with Tools for Deep Research

LEVERAGING LARGE LANGUAGE MODELS FOR MULTIPLE CHOICE QUESTION ANSWERING 🔗 https://arxiv.org/pdf/2210.12353