

# INFOSYS FINAL REPORT

August 2024

# — TABLE OF CONTENT

01

Project overview

02

My role in the project

03

Project agenda

04

Demonstration

05

Futur work

# PROJECT OVERVIEW

- **LLM** (Large Language Model) powered by **RAG** (Retrieval-Augmented Generation) can generate **accurate** and contextually relevant responses by leveraging a vast knowledge base.
- Highly interesting for Infosys due to significant opportunities to enhance information **accuracy** and relevance.
- Problem : hard to extract **multimodal data** from corporate documents (text, tables, and images) and convert this data into **vectors** for developing LLM-powered RAG systems.
- Project : Create a solution and demonstrate it using an open-source library like **deepdoctection**.

# MY ROLE IN THE PROJECT

Investigate various open-source and cloud-based solutions

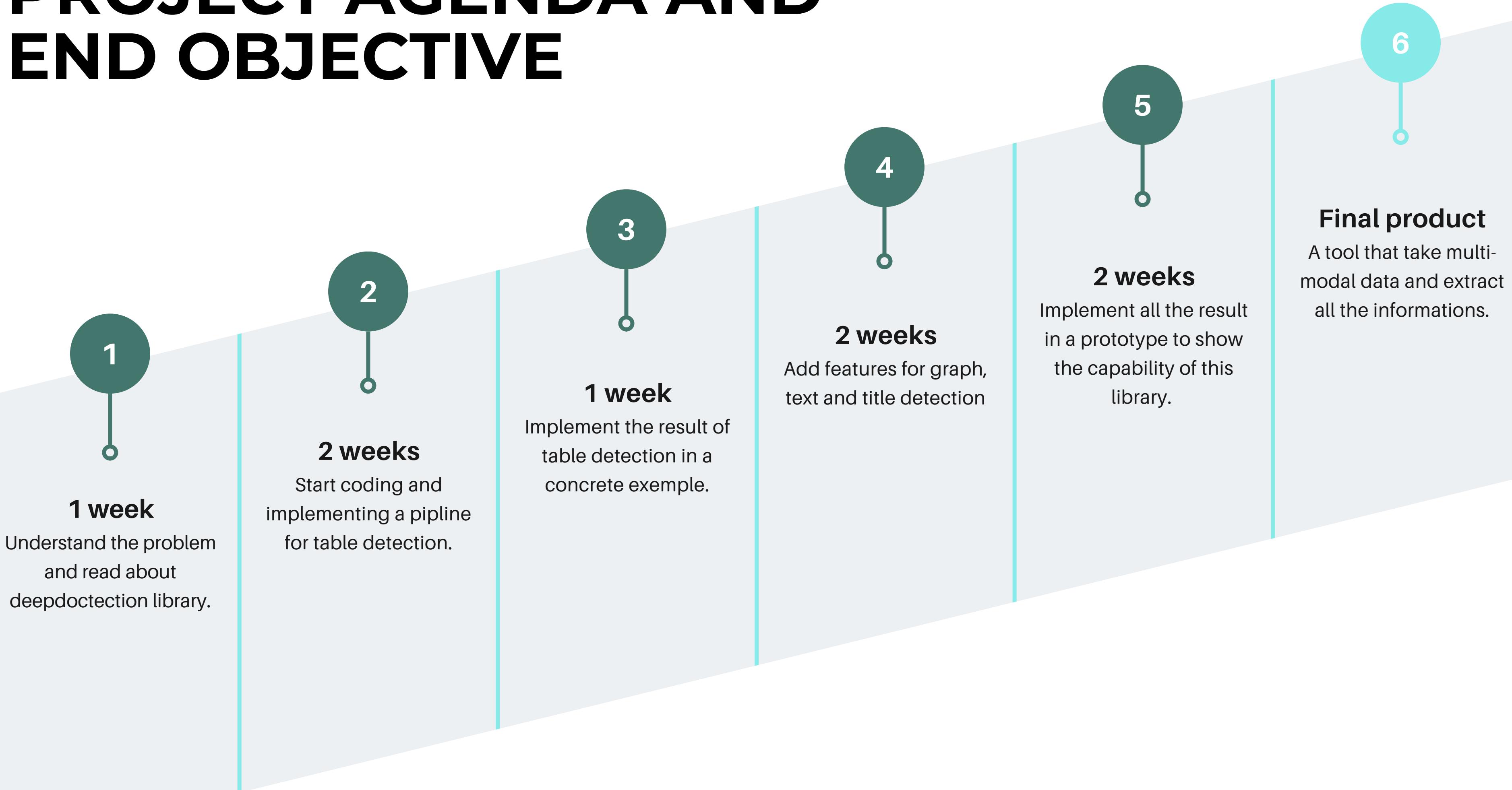
Understand the deepdoctection library

Have a first pipeline of extracting table from documents.

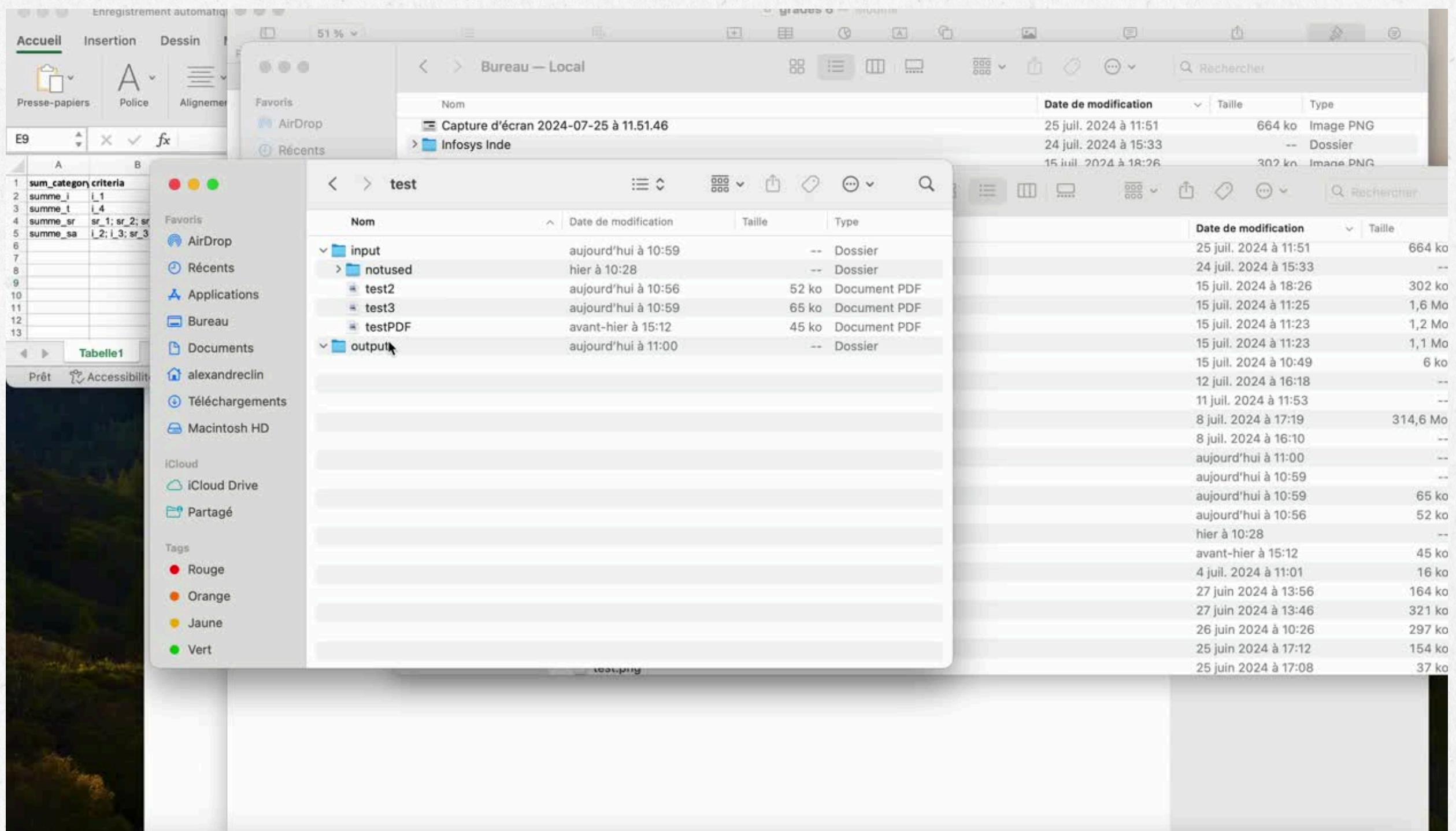
Add features (text/title detections)



# PROJECT AGENDA AND END OBJECTIVE



# DEMO



<https://github.com/alexxelaalexxela/Multi-modal-detection-deep-learning>

# WHAT'S NEXT ?

Improve the features and add graph detection

Implement it in a RAG model to see the result

# THANK YOU

