TEAM 1

HACKATHON

for the many journeys in life

andremehangga.kurniawan@kedgebs.com
nikhilsunil.gaikwad@kedgebs.com
Zhankhaidarabzhan@kedgebs.com
alexandra.mongouewandji@kedgebs.com
gowri.kothakota@kedgebs.com
daniesteven.rodriguezsandoval@kedgebs.com

CONTEXT

OBJECTIVE

Extract and aggregate pollution KPIs (NOx, COx, PM) from various data sources.

CHALLENGES

Multiple formats (PDFs, websites, databases), time-consuming manual processing, scalability needs.

DELIVERABLE

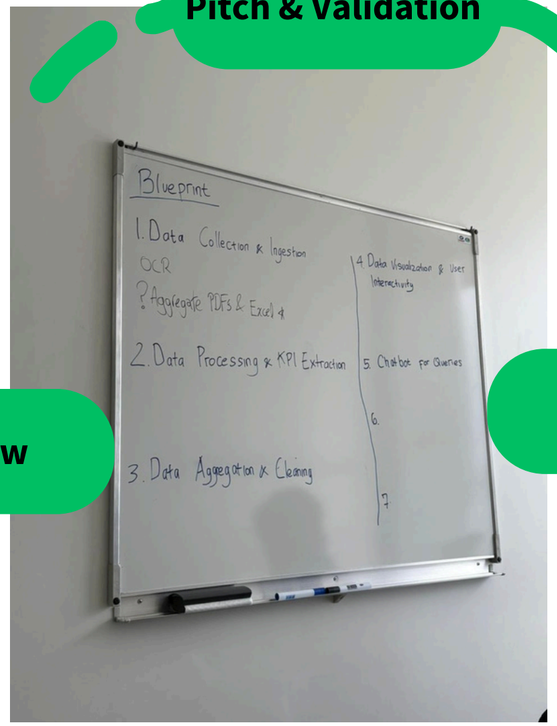
Automated, scalable, and reliable Azure-based solution.

DOCUMENTATION

Project Objective

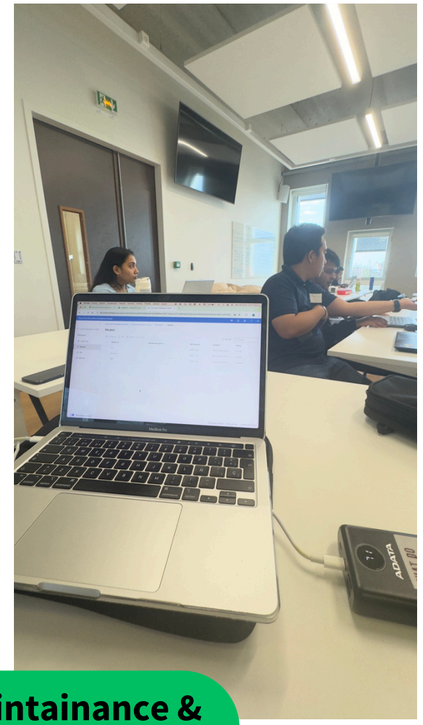
System Architecture

Workflow



Pitch & Validation

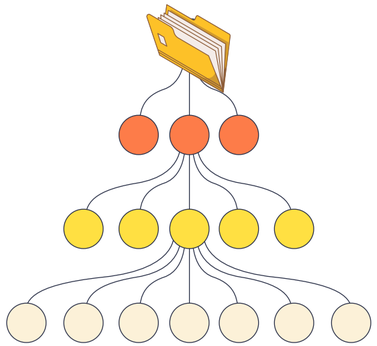
Challenges



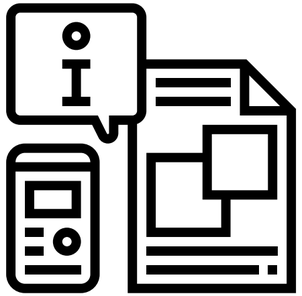
Maintainance & Conclusion

OUR PIPELINE

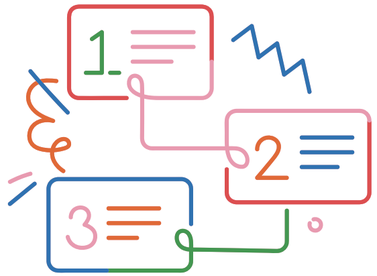
COLLECTION



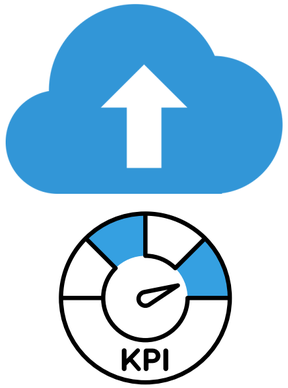
DOCUMENT CLASSIFICATION



INFORMATION EXTRACTION

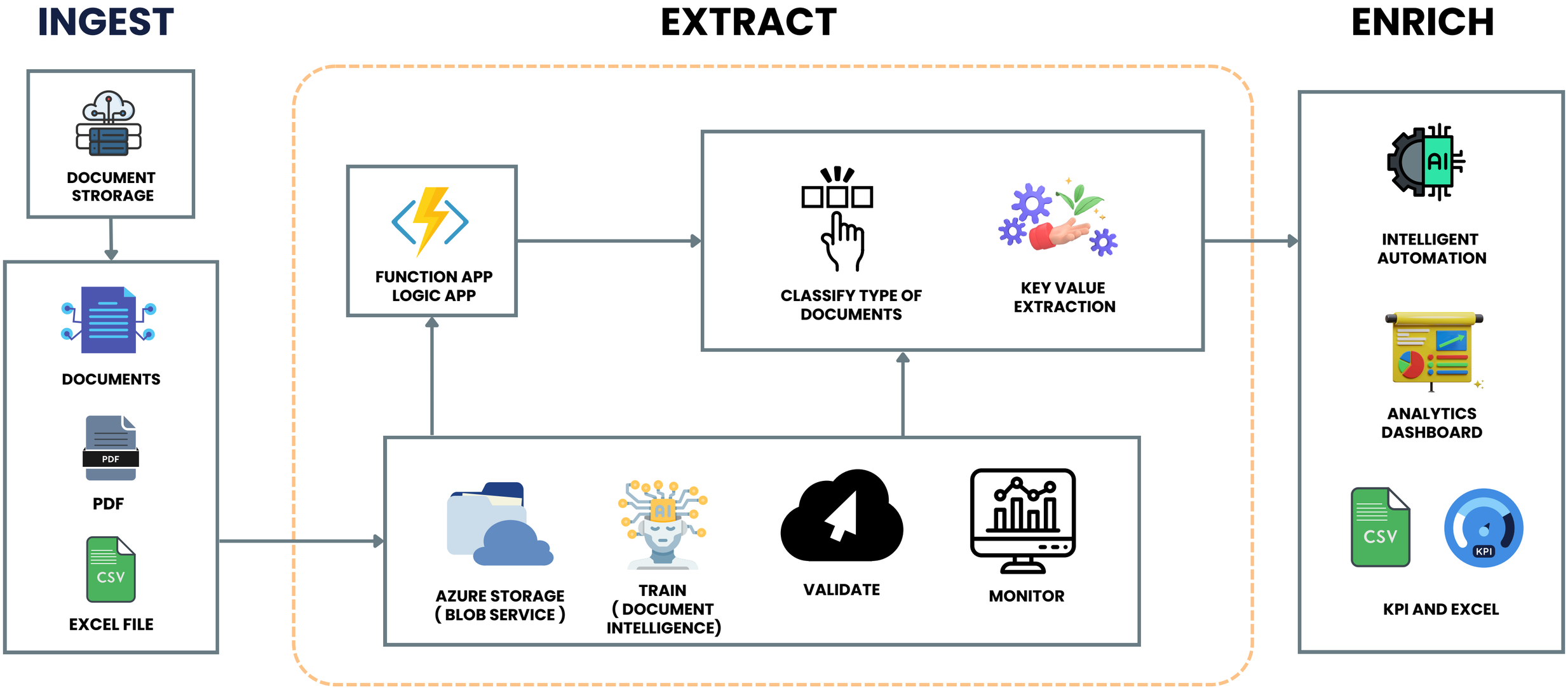


POST PROCESSING



RESULT

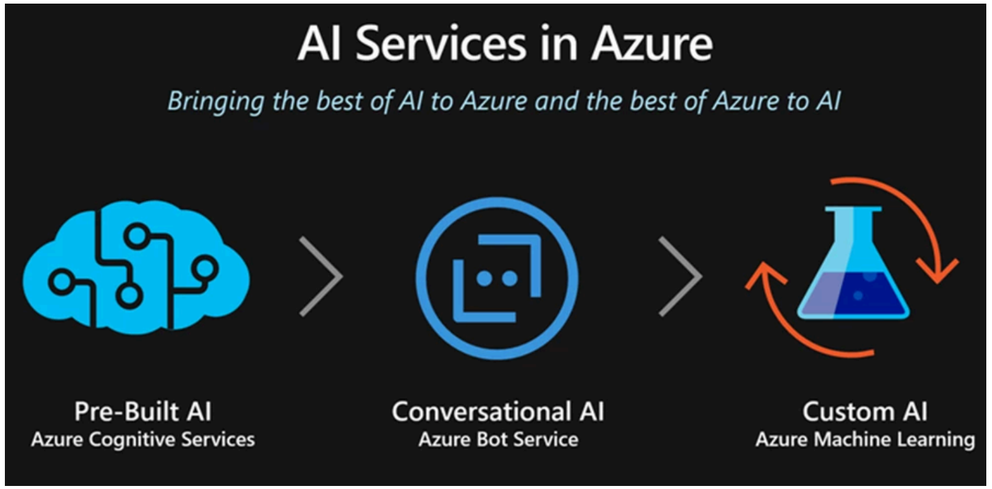
PROCESSING


INGEST
EXTRACT
ENRICH
DOCUMENT STRORAGE
DOCUMENTS
PDF
PDF
CSV
EXCEL FILE
FUNCTION APP
LOGIC APP
CLASSIFY TYPE OF DOCUMENTS
KEY VALUE EXTRACTION
AZURE STORAGE (BLOB SERVICE)
TRAIN (DOCUMENT INTELLIGENCE)
VALIDATE
MONITOR
AI
INTELLIGENT AUTOMATION
ANALYTICS DASHBOARD
CSV
KPI
KPI AND EXCEL

FINAL OUTPUT

List of vehicle emission types selected for and results of PEMS testing

PEMS-test-family	08-SKx-715W_5E_DADA_FD7_0_A_1_0-000								
Vehicle emission type	715W_5E_DADA_FD7_0_A_1_0								
Commercial name	Octavia								
Combination of fuel [1]	Petrol								
PMR_H [2]	- -								
PMR_L [3]	- -								
Transmission type [4]	FD7								
Four-wheel drive	- -								
Engine volume [cm^3]	1498								
Rated power [kW]	110								
Exhaust after-treatment	Particular Filter; 2 Three-Way-Catalysts								
cold testing (c) or hot testing (h)	c								
Test driven by [5]	Technical service								
Method	Moving Averaging Window *						Pc		
Exhaust emissions	CO_{total} mg/km	$NO_{X\ total}$ mg/km	PN_{total} #/km	CO_{urban} mg/km	$NO_{X\ urban}$ mg/km	PN_{urban} #/km	CO_{total} mg/km	$NO_{X\ total}$ mg/km	PN_{t} #/kr
Test results – value (M_{si})	49,9	10,6	1,93E+10	24,9	18,5	1,99E+10	52,2	10,5	1,54E
K_i factor (acc. to 2017/1347 WLTP)									
K_i offset (acc. to 2017/1347 WLTP)									
Value calculated with Ki (Mpi)	49,9	10,6	1,93E+10	24,9	18,5	1,99E+10	52,2	10,5	1,54E
Declared Maximum RDE Values	- -	126	6E+11	- -	126	6E+11	- -	126	6E+



Microsoft Azure | databricks

Search data, notebooks, recents, and more... CTRL + P

gobigorgohome

New

Workspace
Recents
Catalog
Workflows
Compute
Marketplace

SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses

Engineering
Job Runs
Data Ingestion
Delta Live Tables

hine Learning
Playground
Experiments
Features

Untitled Notebook 2025-01-28 22:15:00

Python

File Edit View Run Help Last edit was 38 minutes ago

Run all | KURNIAWAN ANDREM... | Schedule | Share

Workspace

andremehangga.kurniawan@kedge...

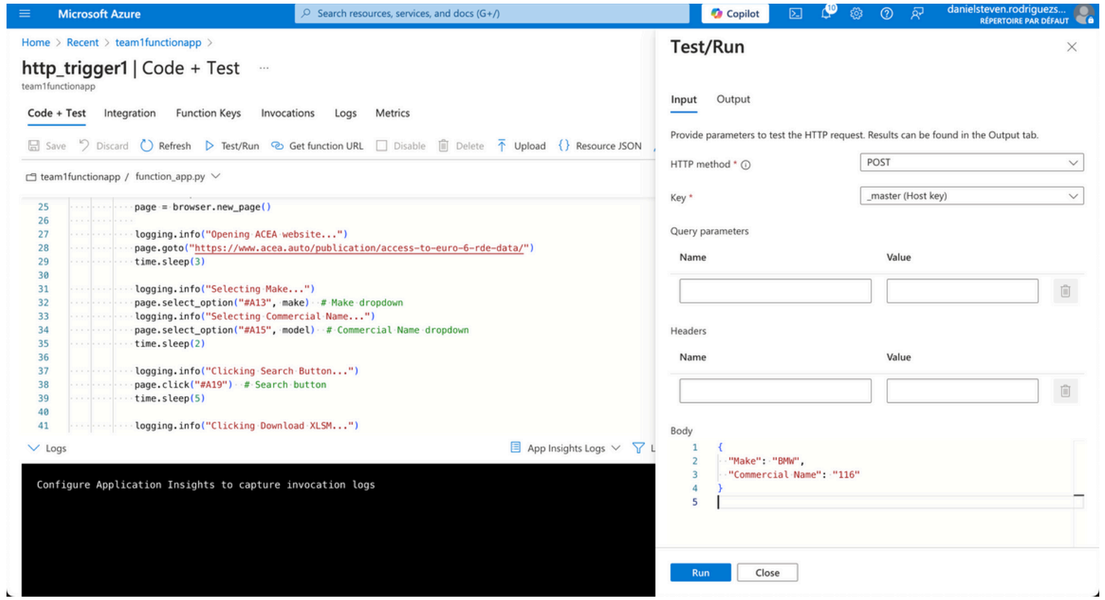
PDF SOURCE
RESULTS
Untitled Notebook 2025-01-28 22:15:00

12:04 PM (3s) 11 Python

```python
import openai

# Azure OpenAI credentials
AZURE_OPENAI_ENDPOINT = "https://grp1trial.openai.azure.com/openai/deployments/gpt-4o/chat/completions?api-version=2024-08-01-preview"
AZURE_OPENAI_KEY = "FdzBu9g1IX6wJ9b5sWtoWnoKXOmC3hBucP4fF9tBrWuwHKZUJYVWJQQJ99BAACYeBjFXJ3w3AAABACOGGKo2"
DEPLOYMENT_NAME = "gpt-4o"

# Configure the OpenAI client
client = openai.AzureOpenAI(
    api_key=AZURE_OPENAI_KEY,
    api_version="2024-08-01-preview",
    azure_endpoint=AZURE_OPENAI_ENDPOINT
)

# Send a prompt to GPT-4o
def get_ai_response(prompt):
    response = client.chat.completions.create(
        model=DEPLOYMENT_NAME,
        messages=[{"role": "system", "content": "You are an AI assistant. "},
                  {"role": "user", "content": prompt}],
        max_tokens=200
    )

    return response.choices[0].message.content

# Example usage
prompt = f"""
Extract the following information from the text:
    - Total NOx (mg/km)
    - Total CO (mg/km)
    - Total PM (mg/km)
    - Source Type
    - PEMS Test Family
```

WHAT WENT WRONG

AZURE FUNCTIONS



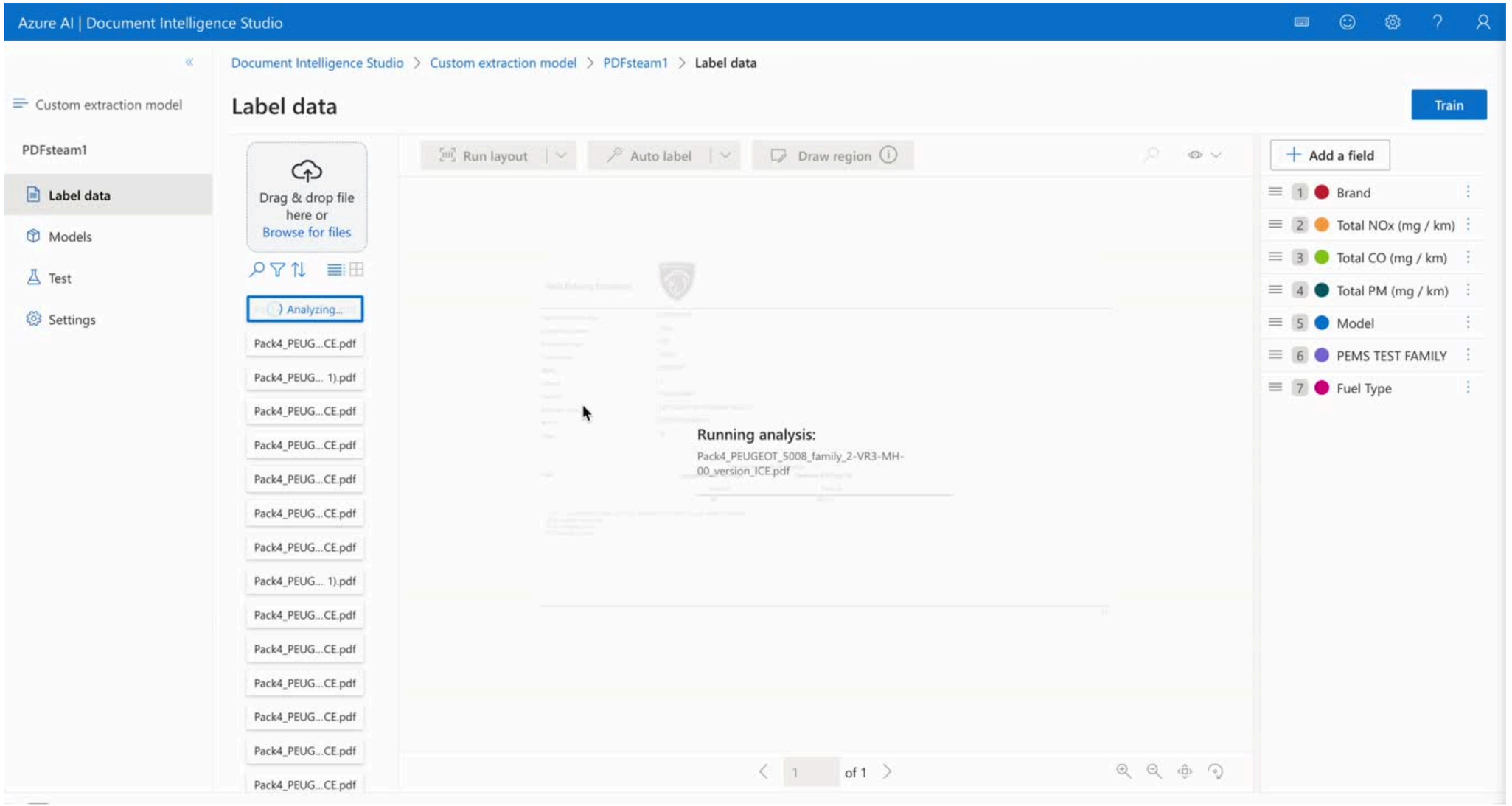
KEY VALUES

Extraction of particular information becomes complex for data extraction

DOCUMENT INTELLIGENCE

PDF documentation was modeled and trained, but Azure lagged out and deleted files several times due to increased call value.

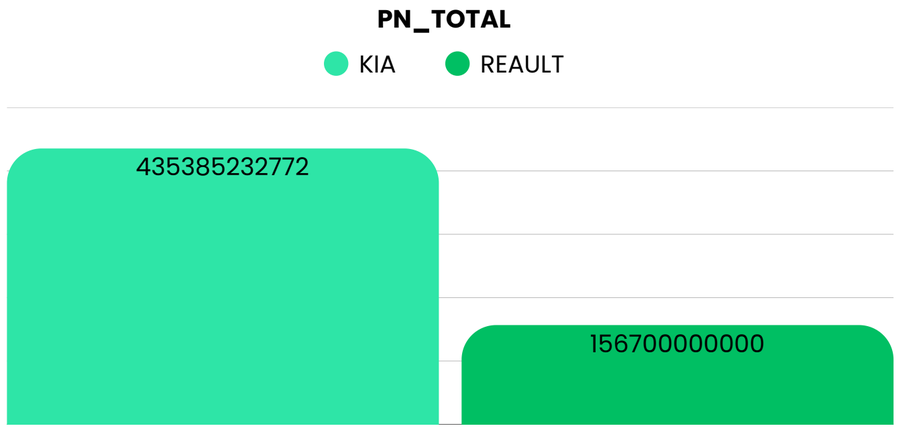
WHAT WENT WRONG



KPIS

WHAT WE WANTED TO ACCOMPLISH

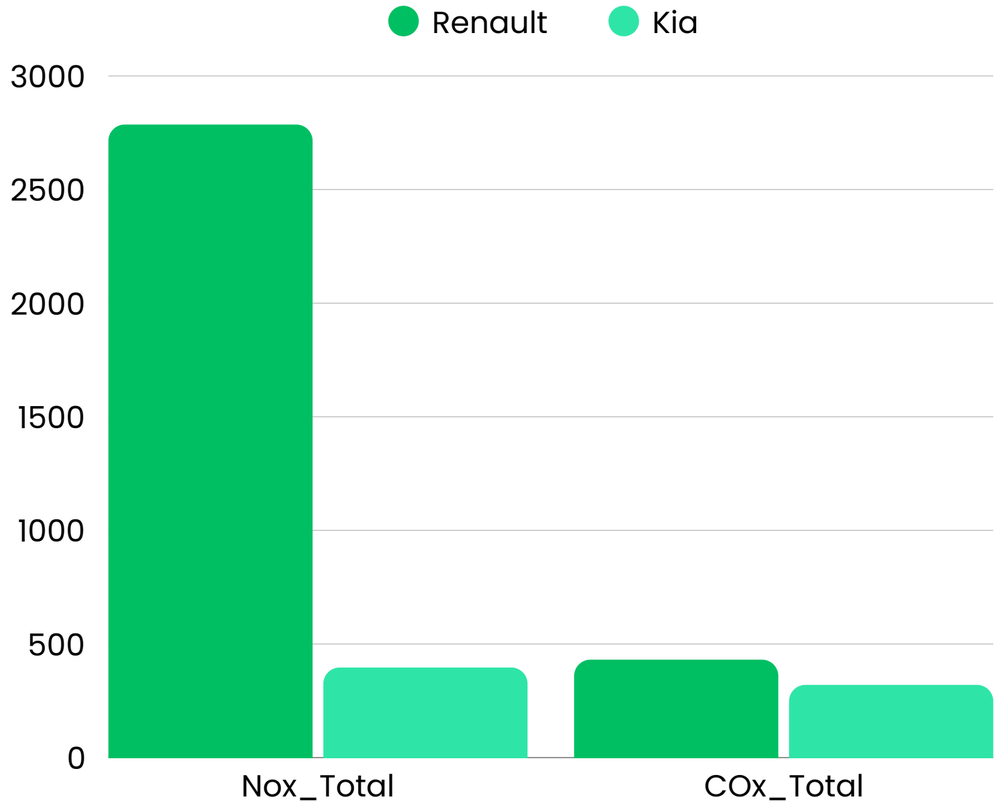
MOCK EXAMPLE



NOx (Nitrogen Oxides) Emissions: Renault has significantly higher total NOx emissions**.**

PN (PARTICULATE MATTER) EMISSIONS: KIA’S PARTICULATE EMISSIONS ARE MUCH HIGHER.

COX Values: Renault's values are much higher than Cupra’s



ADVANTAGES OF OUR IDEA

Model	Pricing	Pricing with Batch API*
gpt-4o	$0.00250 / 1K input tokens $0.00125 / 1K cached** input tokens	$0.00125 / 1K input tokens

PROMPT ≈ 30 WORDS
9000 TOKENS | 22 $

TIME EFFICIENCY

FASTER DATA RETRIEVAL

INCREASED ACCURACY & DATA RELIABILITY

ELIMINATION OF MANUAL ERRORS & INCONSISTENCIES

COST SAVINGS

AUTOMATED DATA EXTRACTION

RECOMMENDATIONS FOR FURTHER INVESTIGATION

ADVANCED AI TECHNIQUES FOR DATA EXTRACTION

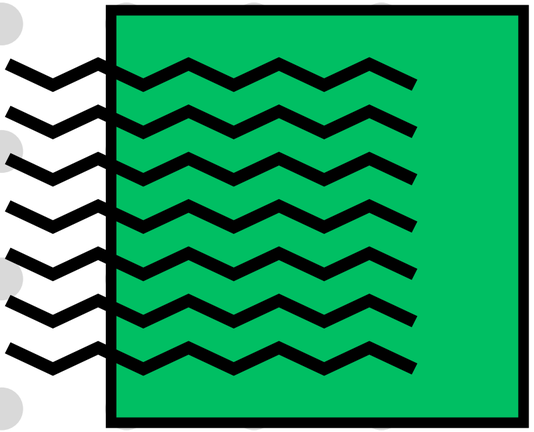
Enhance Natural Language Processing(NLP) models to handle multilingual & unstructured data

INTERACTIVE DASHBOARD FOR VISUALIZATION

Enable dynamic filtering & real-time comparisons of pollution KPIs

AI-POWERED REGULATORY COMPLIANCE ASSISTANT

Build a virtual assistant that automatically checks whether a vehicle model complies with different regulations



THANK YOU