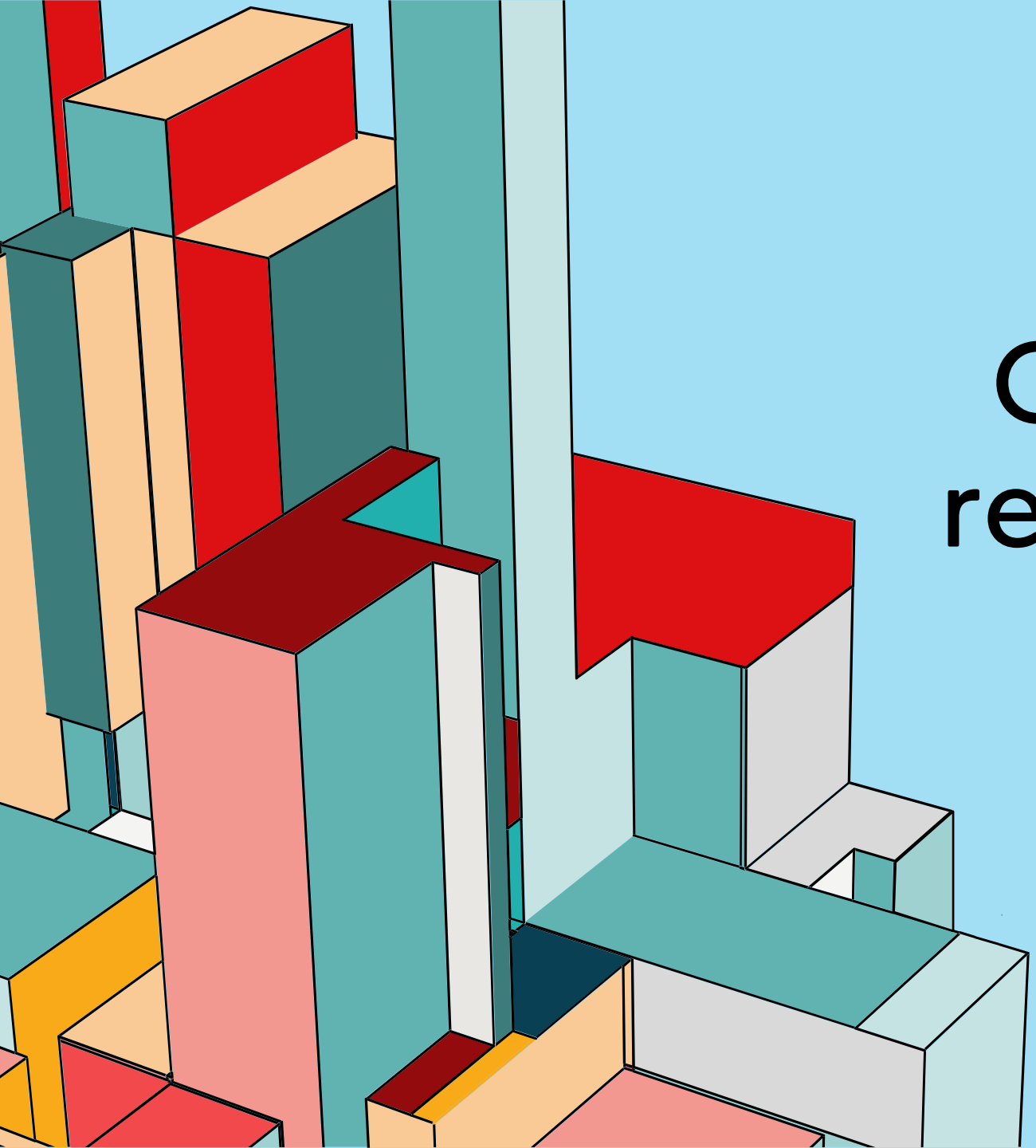


Classification and retrieval system for gas pipe repairs

Politecnico di Torino

FONDAZIONE links

PASSION FOR INNOVATION

PROJECT PRESENTATION

This project focuses on developing a **Retrieval-Augmented Generation (RAG) system** that helps determine whether a specific patch can be applied to damaged gas pipes.

"There is a tube with a yellow valve and a fault at the bottom"

Example of a textual description

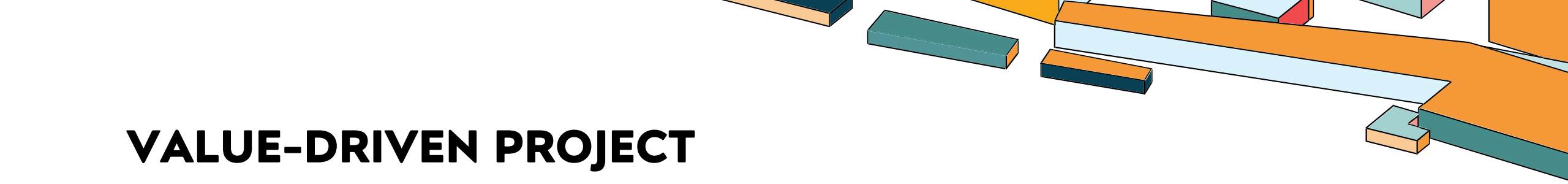
AI

APPLICABLE

Accompanied by a short description

NOT APPLICABLE

Accompanied by a short description



VALUE-DRIVEN PROJECT

This project improves **safety, efficiency and economic savings** in gas pipe repairs by providing a **decision-support tool** to quickly and accurately assess patch applicability

The project is supervised by **LINKS Foundation**, which aims to drive innovation within companies.

DATA

- The dataset comprises **real-world data** derived from historical gas pipeline repairs, along with **synthetic data**, leading to a total of **1,000 records**
  - 80% of the data are synthetic

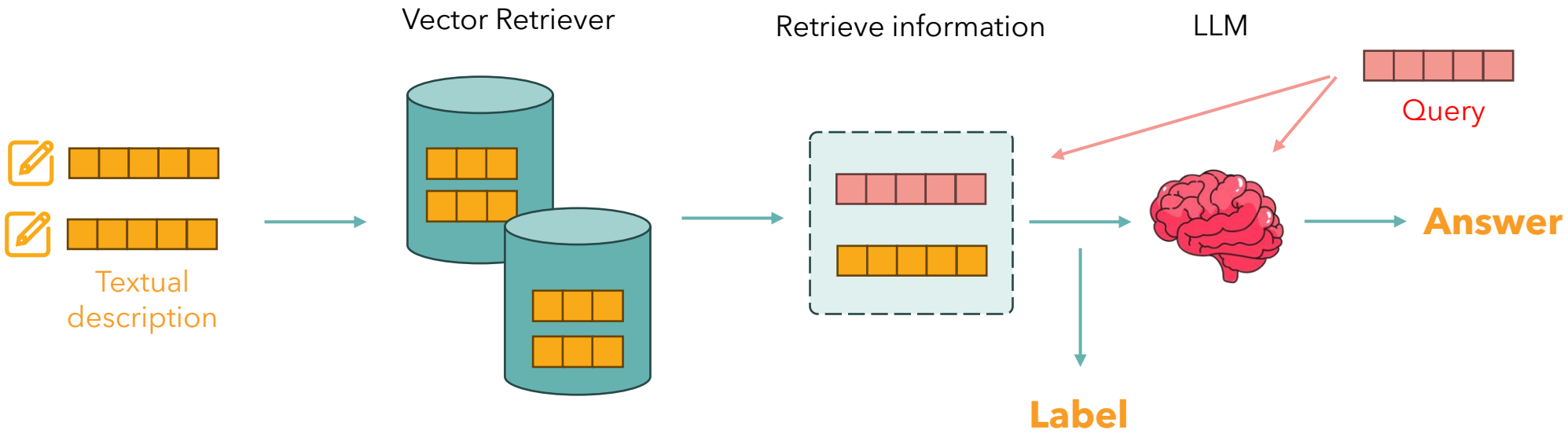
- It is a **CSV file**, containing **textual descriptions** obtained through an LLM model that summarize images of pipe faults, **labels**, and **25 supplementary features** that provide additional contextual information

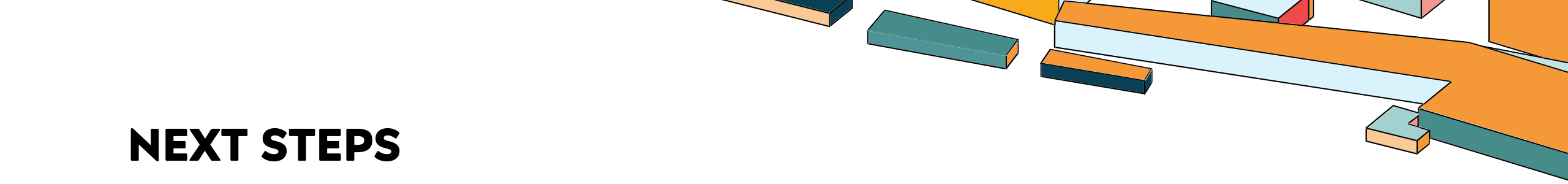
	ID DATABASE	CODICE RIPARAZIONE	Riparazione effettuata per conto di	La tubazione, nella zona della lesione, appare in cattive condizioni di conservazione?	MATERIALE TUBAZIONE DA RIPARARE	DIMENSIONE TUBAZIONE DA RIPARARE	CORRISPONDENZA ALLE SPECIFICHE TECNICHE D'IMPIEGO DEL KIT	TIPOLOGIA LESIONE DA RIPARARE	COLLOCAZIONE/ESPOSIZIONE SOLARE TUBAZIONE	NOME FILE IMMAGINE DELLA RIPARAZIONE	NOME FILE IMMAGINE DELLA LESIONE	Descrizione intervento
0	ID000001	INT000001	GAS01	FALSO	Raccordi zincati	diametro compreso tra 40 e 80 mm o uguale a 80 mm	Le dimensione e la tipologia della lesione cor...	Gomito	Allacci utenza	repair_000001	damage_000001	The polyethylene gas pipe features a distincti...

TASK

Classification and retrieval task: label indicating whether a gas patch is applicable or not for each textual description of an intervention and explain why

- Load Database with Past Failures' Embeddings
- Retrieve Most Similar Embedding to Query
- Label and Explain Using LLM

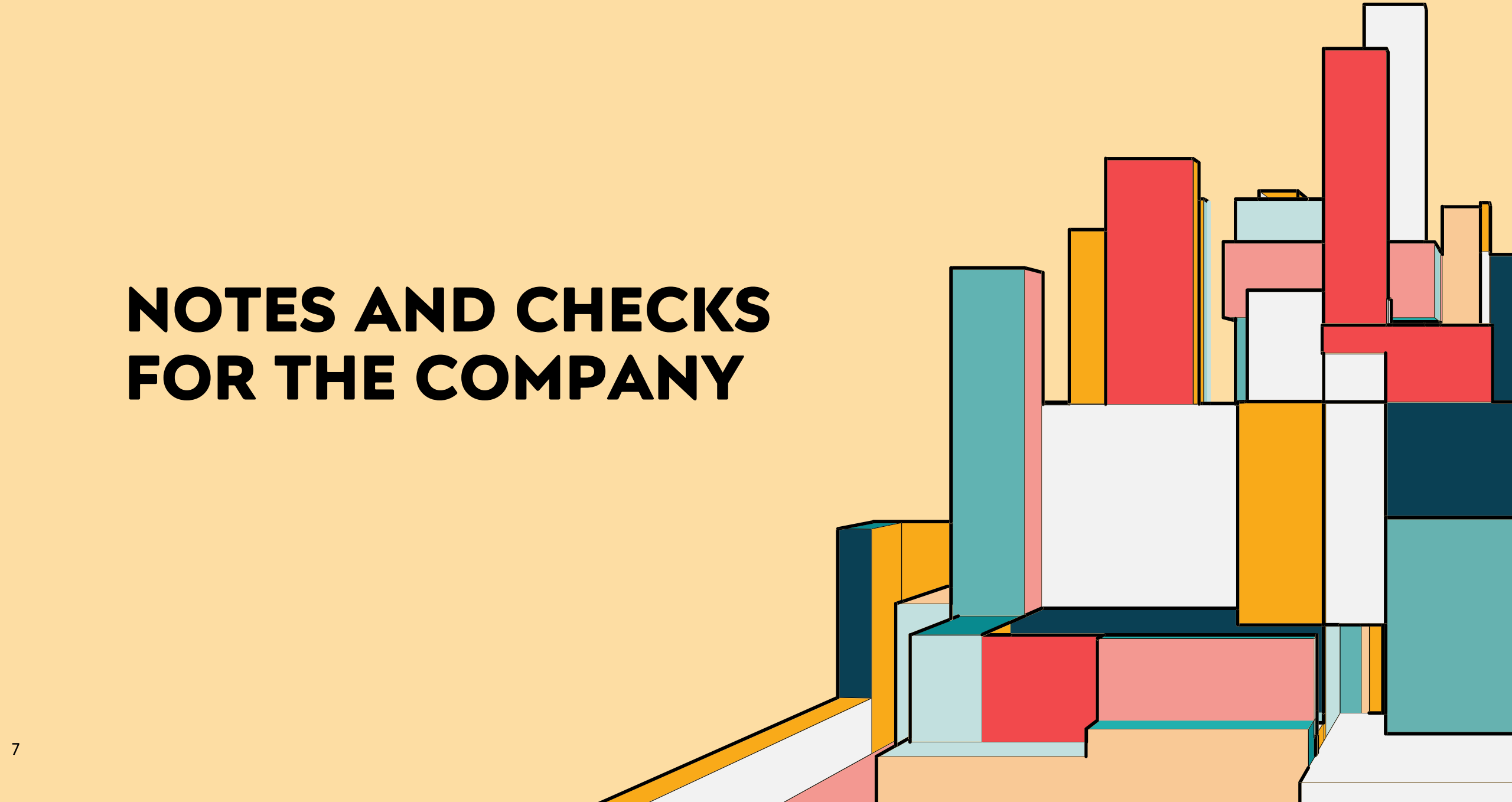


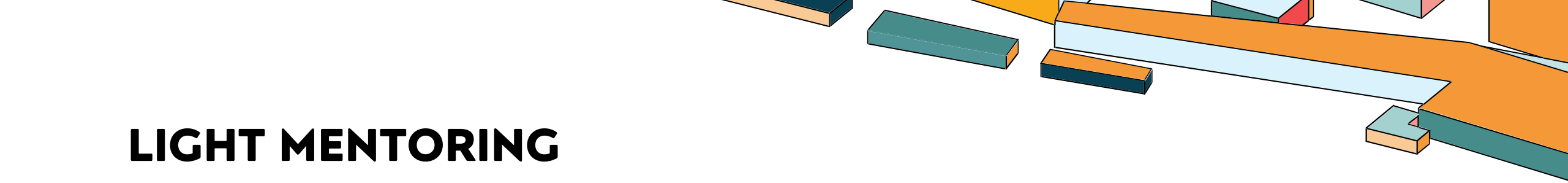


NEXT STEPS

Dataset Exploration	
RAG System Implementation	Use a pre-trained encoder and LLM Fine-tune a textual encoder to generate more informative embeddings
Evaluation Metrics	Accuracy, F1 score, and Recall@K

NOTES AND CHECKS FOR THE COMPANY

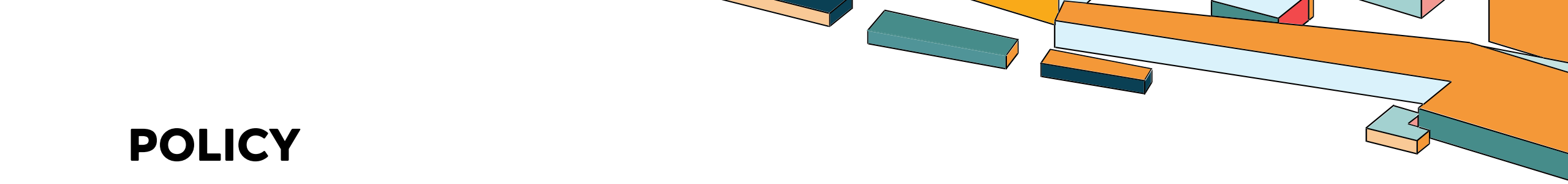




LIGHT MENTORING

- **First (longer) meeting** as an introduction to the existing pipeline and methodology, context and data presentation.

- **30 minutes biweekly of calls** with students for the whole duration of the semester



POLICY

- Both project descriptions and implementations will be part of a repository group published on GitHub
- The repositories will be public unless requests from the organization that will be discussed
- Ideally, the projects should be conceived open from the design

CONTACTS

Luca Catalano

luca.catalano@linksfoundation.com