Automated Artwork Metadata Generation using Multimodal Large Language Models

Politecnico di Torino

FONDAZIONE links PASSION FOR INNOVATION

PROJECT PRESENTATION

This project focuses on developing a **Multimodal Large Language Model** (MLLM) system to automatically generate metadata for new paintings.

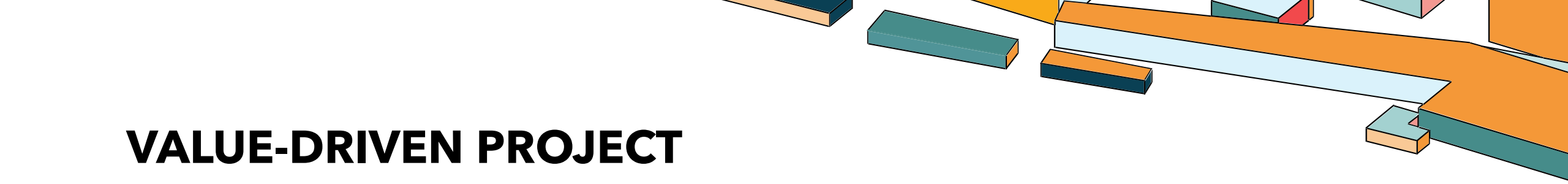
Image

What is the artist of this painting?

Question

AI
(MLLM)

Leonardo Da Vinci

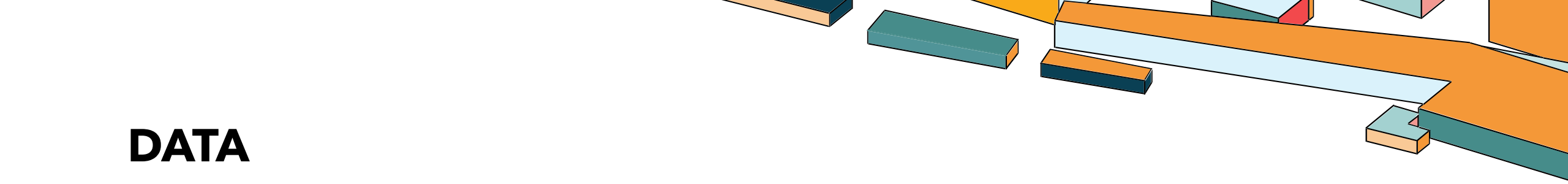


VALUE-DRIVEN PROJECT

This project enhances **innovation, accessibility, and cultural preservation** in the **art sector** by developing an AI-driven tool that automatically generates metadata for new paintings. The system improves efficiency in **art management**, supports accurate documentation, and **enables broader access to artistic content**—creating value through automation, creativity, and digital inclusion.

REEVALUATE

The project is carried out under the supervision of **LINKS Foundation**, within the framework of the European project "**REEVALUATE**."



DATA

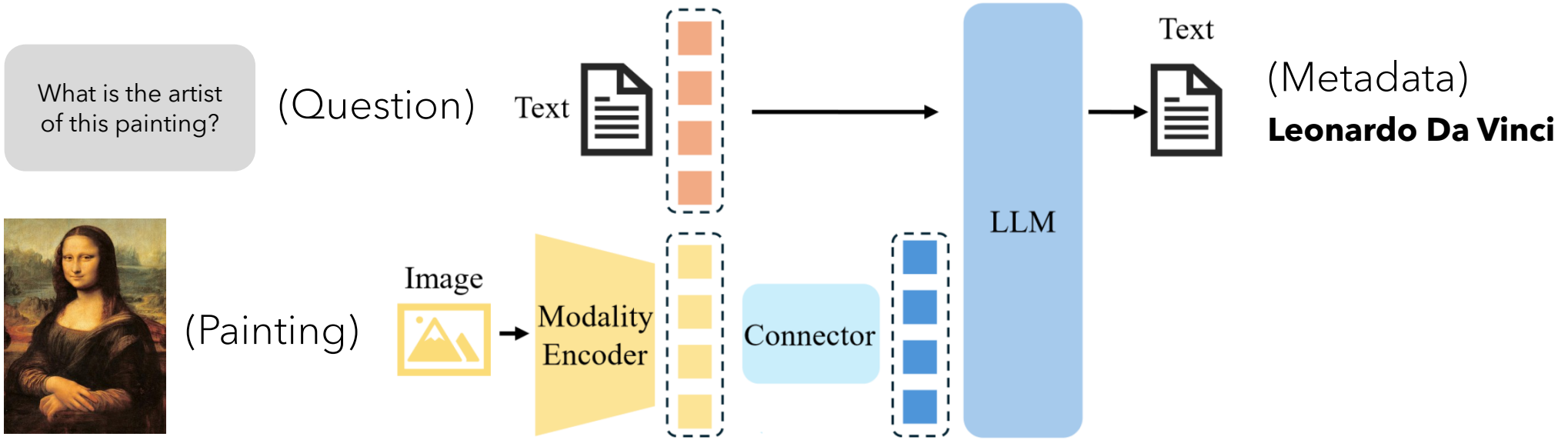
- The dataset comprises **images** and **textual metadata** derived from WikiData dataset
  - **Folder** with images of paintings
  - **CSV File** with textual metadata for each painting in the folder

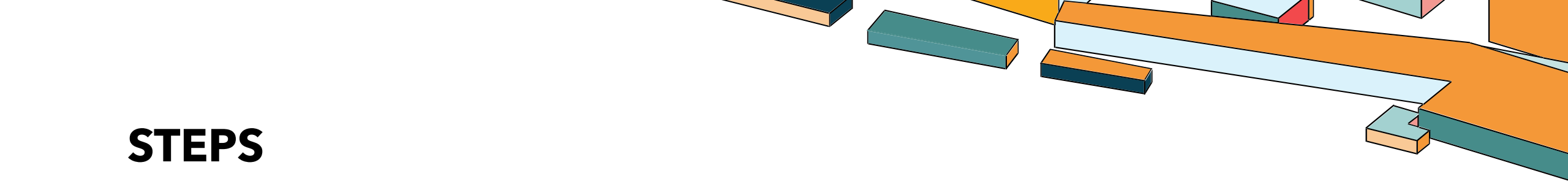
- 3 split:
  - **Train**: Main portion of the dataset used to train the model.
  - **Validation**: Contains samples with the same distribution as the training set (same metadata categories), used to monitor performance and prevent overfitting.
  - **Hard Validation**: Includes entries from metadata categories not present in the training set, designed to evaluate the model's ability to generalize to unseen categories.

TASK

Visual Question Answering: providing accurate answers to questions about a given image by combining visual understanding with natural language reasoning.

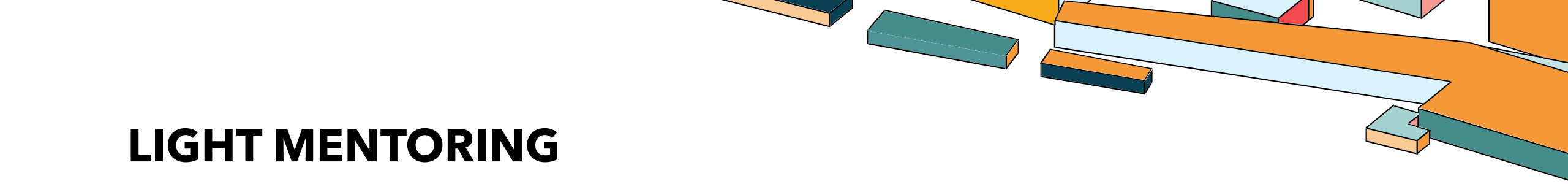
The main goal of this project is to fine-tune a **Multimodal Large Language Model (MLLM)** on the art dataset using **Parameter-Efficient Fine-Tuning (PEFT)** techniques to improve its performance on domain-specific visual questions.





STEPS

- Data Exploration
- Zero-shot evaluation of the MLLM on the VQA task
- Training (adaptation) of the MLLM for the VQA task.
- Evaluation on the validation splits

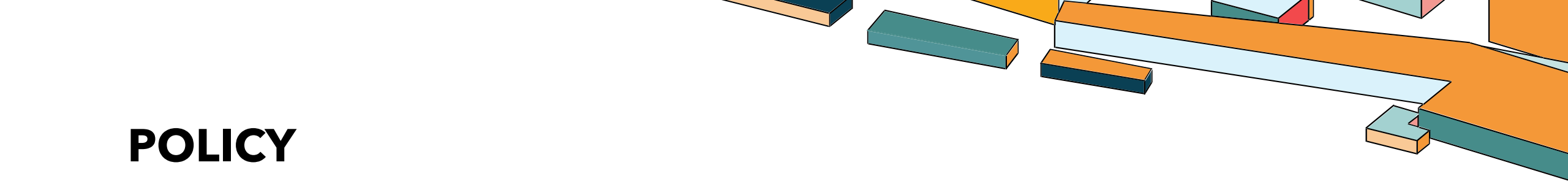


LIGHT MENTORING

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

- **Weekly one-hour calls** with students for the whole duration of the semester.

- Feel free to reach out via Slack or email at any time for any questions or doubts.



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.

CONTACTS

Luca Catalano
luca.catalano@linksfoundation.com

Federico D’Asaro
federico.dasaro@linksfoundation.com