



**MDMC**

Master in Data Management  
and Curation

**AREA**  
SCIENCE PARK

**SISSA**

# Applied AI

Matteo Biagetti, Emanuele Panizon, Tommaso Rodani

Laboratory of Data Engineering, Area Science Park

25 November 2025



# Applied Artificial Intelligence Course

---

Who are we?



Matteo Biagetti



Emanuele Panizon



Tommaso Rodani



Maria Pronestì



Ahmed Khalil

# Applied Artificial Intelligence Course

---

- Course Objectives
- Course Structure
- Project 1: SEM images
- Project 2: Sensorium
- Initial Setup

# Applied Artificial Intelligence Course

---

## Course Objective: a FAIR research project

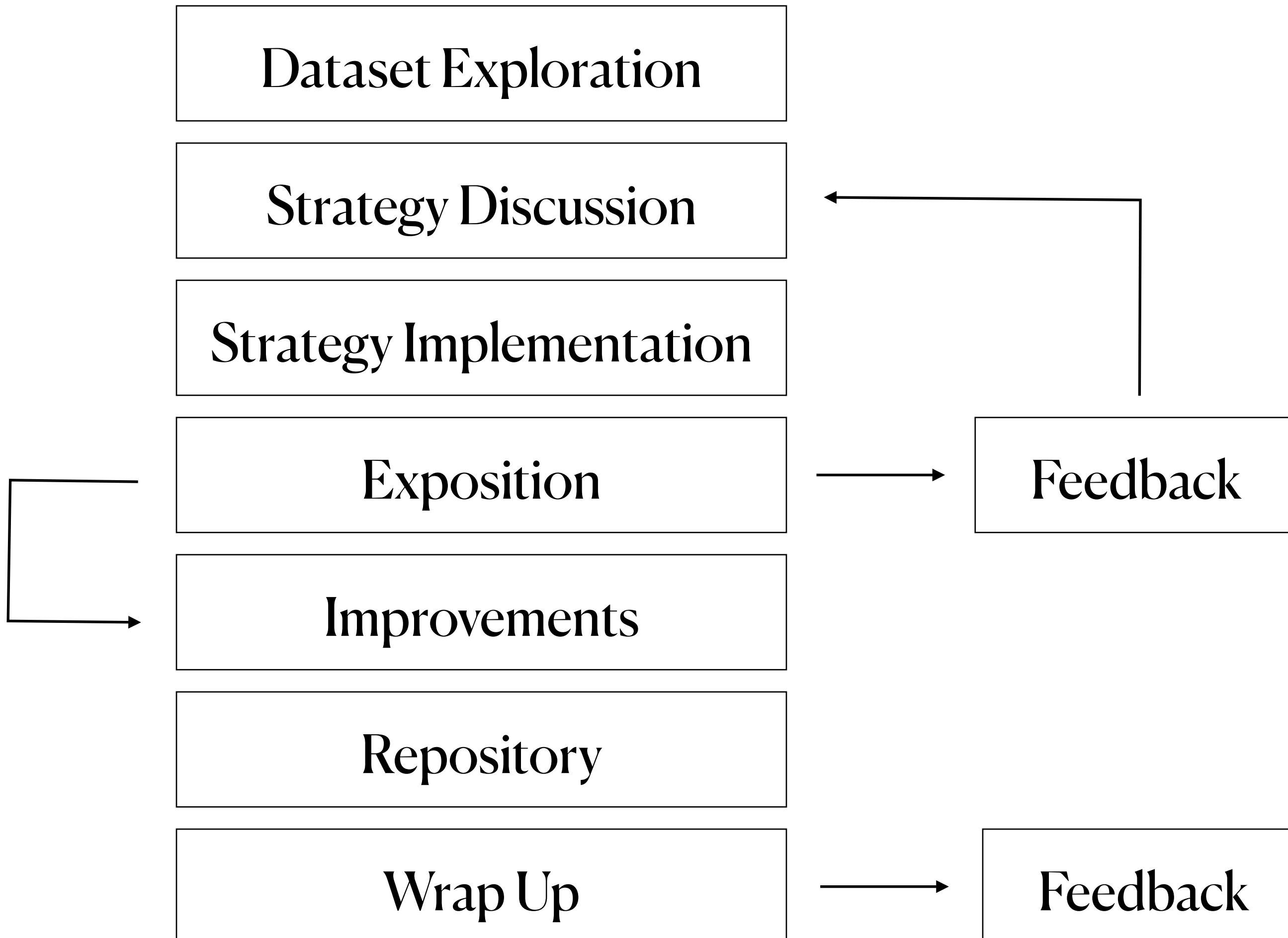
- Problem solving
- Coding, from scratch
- Repository development
- Data pre-processing and post-processing pipeline
- Documentation for reusability
- Open access

# Applied Artificial Intelligence Course

---

**Course Objective:** a FAIR research project

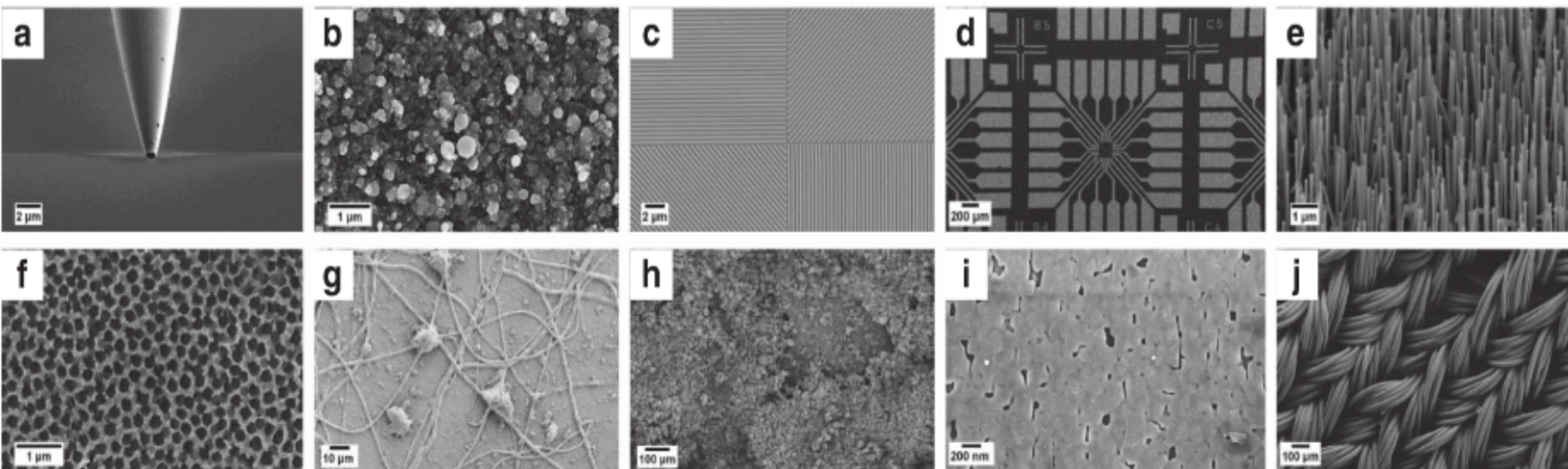
Structure:



# Applied Artificial Intelligence Course

## Project 1: Scanning Electron Microscopy (SEM) Images

DOI: <https://b2share.eudat.eu/records/e344a8afef08463a855ada08aadb352>



- (a) Tips (b) Particles (c) Patterned surfaces (d) MEMS devices and electrodes
- (e) Nanowires (f) Porous sponge (g) Biological (h) Powder (i) Films and coated surfaces (j) Fibres.

Public dataset of processed images + raw experimental data

# Applied Artificial Intelligence Course

---

## Project 1: Scanning Electron Microscopy (SEM) Images

### Dataset Structure:

**Public Data:** 25,537 JPEG images distributed across 10 categories, provided as .tar archives for each category (e.g., Biological.tar)

**Raw Backup:** a collection of .tiff images acquired directly from the SEM microscope. All files are unorganized and lack inherent structure or labeling conventions.

# Applied Artificial Intelligence Course

---

## **Project 1:** Scanning Electron Microscopy (SEM) Images

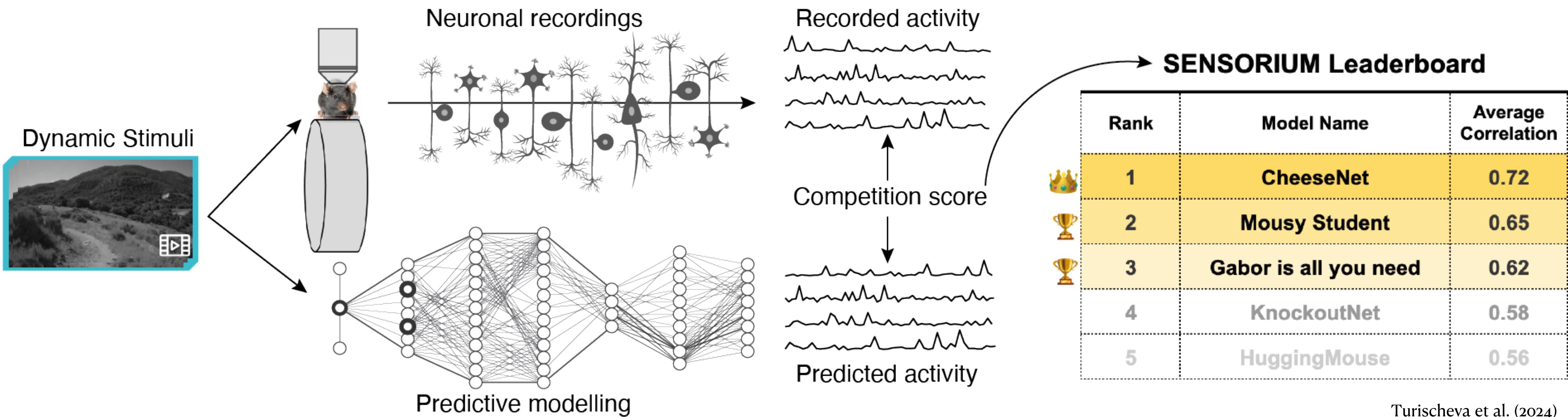
Project Goals:

**Primary goal:** start with the “biological.tar” archive to extract the JPEG images and establish a valid mapping between the two data formats.

**Secondary goal:** extend the data provenance reconstruction to the entire published dataset

# Applied Artificial Intelligence Course

## Project 2: the SENSORIUM dataset



# Applied Artificial Intelligence Course

---

## Project 2: the SENSORIUM dataset

### Dataset Structure:

**Data:** includes the variables that were recorded during the experiment.  
The experimental variables are saved as a collection of numpy arrays.

- videos
- responses
- behavior (pupil dilation, running speed)
- pupil center

**Meta:** includes meta data of the experiment

- neurons coordinates
- statistics of experimental variables

# Applied Artificial Intelligence Course

---

## Project 2: the SENSORIUM dataset

Project A Goals:

**Primary goal:** label videos by types (naturalistic, gaussian, waves) and eventually subtypes

**Secondary goal:** build visualisation tool for exploration of the whole dataset  
(together with group B)

# Applied Artificial Intelligence Course

---

## Project 2: the SENSORIUM dataset

Project B Goals:

**Primary goal:** combine videos and neural responses with all relevant metadata in a data structure that allows an efficient query for preliminary exploration

**Secondary goal:** build visualisation tool for exploration of the whole dataset  
(together with group B)

# Applied Artificial Intelligence Course

---

## Initial Setup

- Teams aggregation
- Access to data at /orfeo/cephfs/scratch/mdmc/appliedAI
- Access to GPU resources - instructions

Papers:	<a href="https://www.nature.com/articles/sdata2018172">https://www.nature.com/articles/sdata2018172</a>	SEM
	<a href="https://www.sensorium-competition.net/">https://www.sensorium-competition.net/</a>	Sensorium

# Applied Artificial Intelligence Course

---

## Teams

### Group 1

SIMONETTI Riccardo  
CHALUVADI Sandeep Kumar  
LAMBERTI Roberta  
ZUPPA Jacopo

### Group 2

NYANDU Kagarabi Emmanuel  
MECCHINA Andrea  
FLO Ana  
NANA Njantang Epse Osazuwa Ruth  
LONGATO Enrico

### Group 3

PALACIOS FLORES Luis Fernando  
NKANA NGAN Valentin Eric Brice  
TALUKDAR Smritirekha  
TSOPTIO Fougang Lesly



**Finanziato  
dall'Unione europea**  
NextGenerationEU



**Ministero  
dell'Università  
e della Ricerca**



**Italiadomani**  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



**E-ARGO**



**beanTech®**  
IT moves your business

**CERIC**

# Applied AI

This Pilot training activity has been funded by the European Union – NextGenerationEU within the PNRR projects funded pursuant to Article 11, paragraph 1, of Notice 594/2024:

- “NFFA-DI cod. IR0000015, Missione 4, “Istruzione e Ricerca” – Componente 2, “Dalla ricerca all’impresa” – Linea di investimento 3.1, “Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione” – Azione 3.1.1, “Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti” (CUP B53C22004310006).
- “EFC cod. SSU2024-00002, Missione 4 “Istruzione e ricerca” - Componente 1, “Potenziamento dell’offerta dei servizi all’istruzione: dagli asili nido all’università” - Investimento 3.4 “Didattica e competenze universitarie avanzate” - Sub-Investimento “Rafforzamento delle scuole universitarie superiori” (CUP: G97G24000100007).