

Master Degree in Computer Science

Master Degree in Data Science for Economics and Health

Natural Language Processing

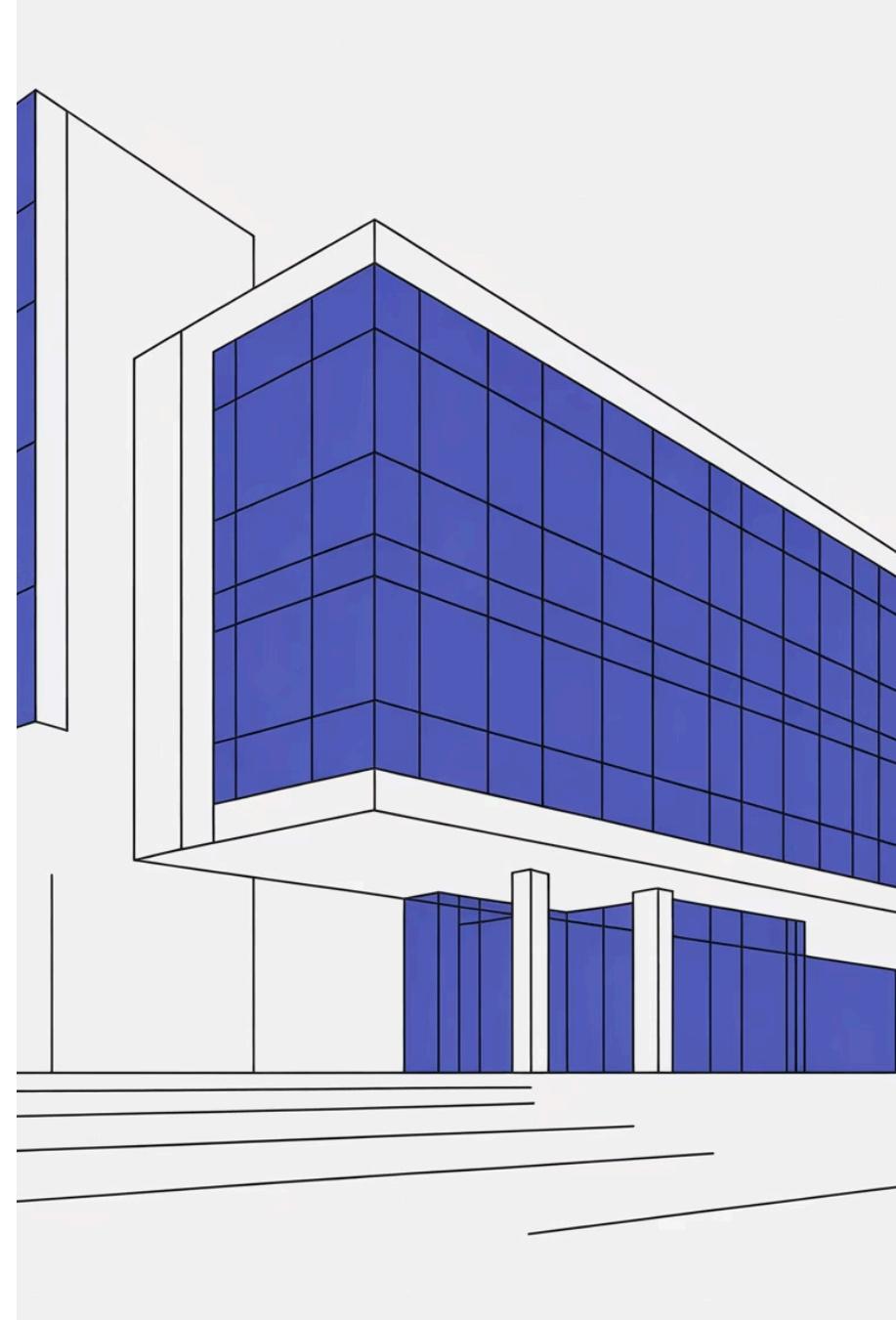
Prof. Alfio Ferrara

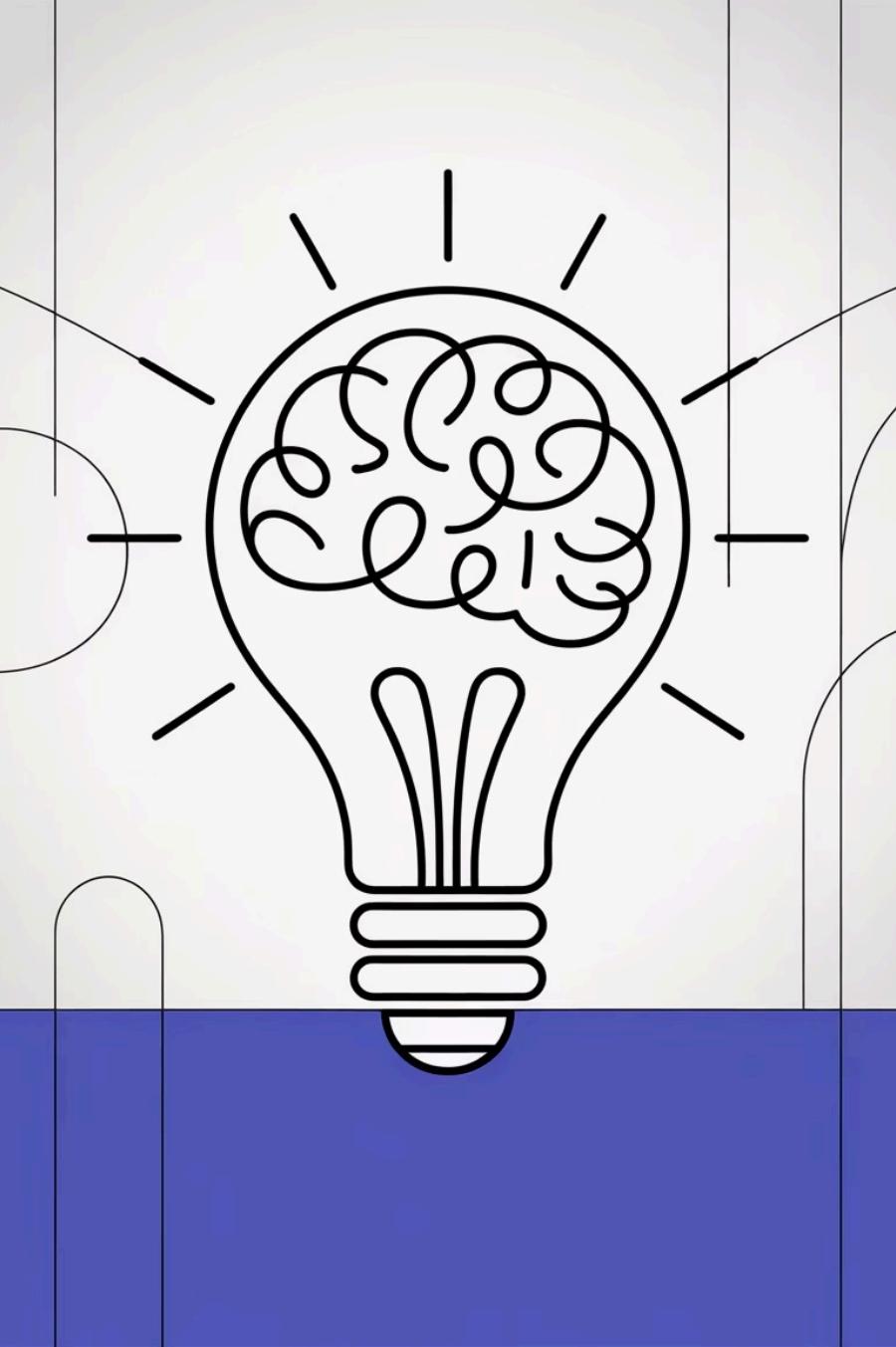
Dott. Sergio Picascia, Dott.ssa Elisabetta Rocchetti

Department of Computer Science, Università degli Studi di Milano

Room 7012 via Celoria 18, 20133 Milano, Italia

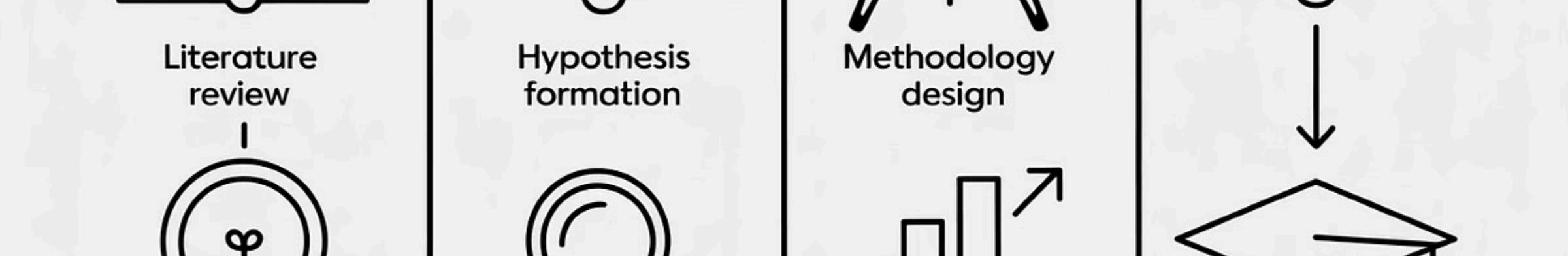
alfio.ferrara@unimi.it





Ideas for Final Projects

Explore cutting-edge research opportunities in Natural Language Processing through innovative project ideas spanning multiple thematic clusters. Each project offers unique challenges and methodological approaches to advance our understanding of language models and their applications.



Literature review

Hypothesis formation

Methodology design

Instructions

The final project consists in the preparation of a **short study** on one of the topics of the course, identifying a precise research question and measurable objectives. The project will propose a methodology for solving the research question and provide an experimental verification of the results obtained according to results evaluation metrics.

The emphasis is **not on obtaining high performance** but rather on the **critical discussion of the results obtained** in order to understand the potential effectiveness of the proposed methodology.

01

Documentation

Short article of 4-8 pages using provided templates

02

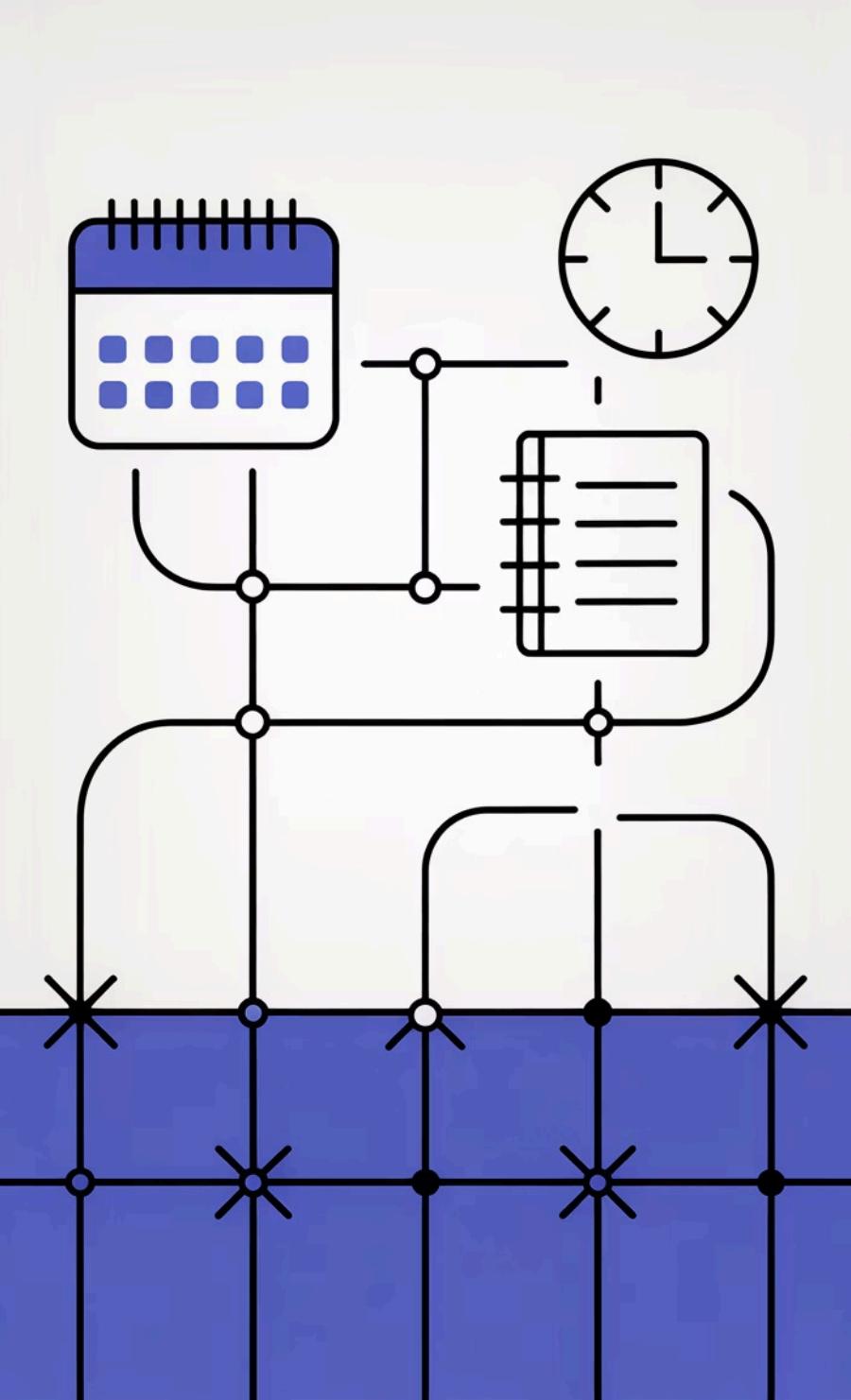
Code Repository

GitHub repository with reproducible experimental results

03

Presentation

10 minutes presentation in English with slides



Procedure

Exam dates are just for the registration of the final grade. The project discussion will be set by appointment, according to the following procedure:

- 1
- 2
- 3

Subscribe to Available Date

Register for any available exam date in the system

Contact Professor

Reach out when project is finished and ready for discussion

Setup Appointment

Schedule and discuss your completed work

- Required Information:** When contacting Prof. Ferrara, provide: (1) Your subscribed exam date, (2) PDF version of your report, (3) GitHub repository link

If you are **interested in doing your final master thesis** on these topics, the final project may be a preliminary work in view of the thesis. Discuss the contents with Prof. Ferrara during the project discussion.

Structure of the Paper

1. Introduction

Provides an overview of the project and a short discussion on the pertinent literature

2. Research Question & Methodology

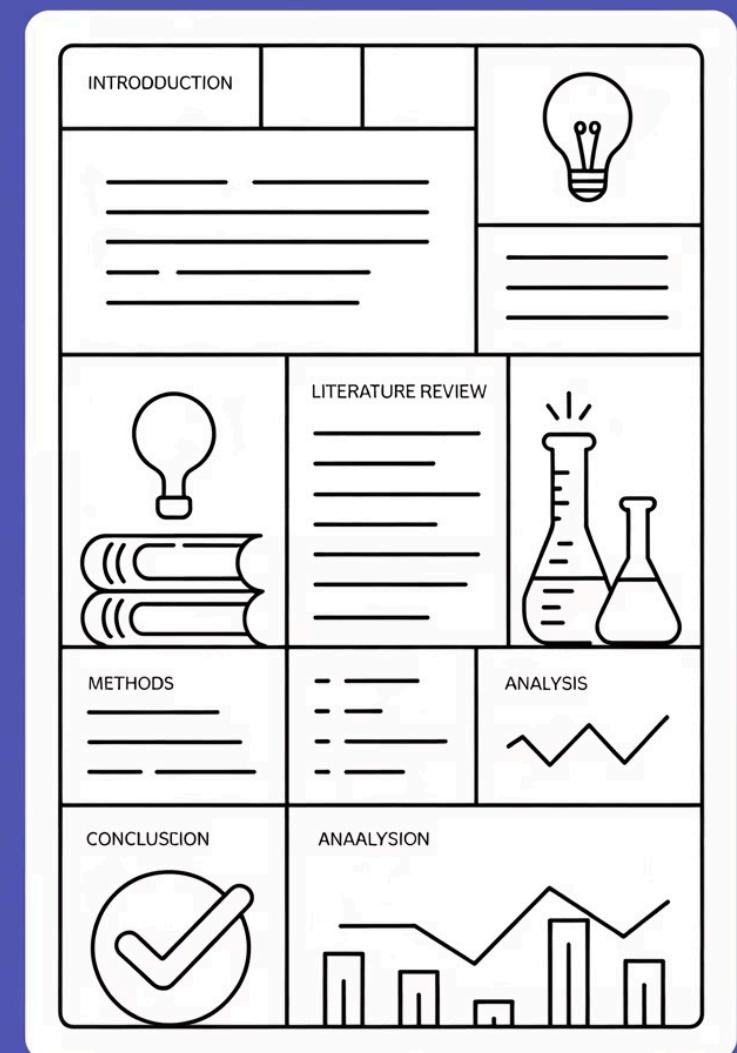
Clear statement of goals, overview of proposed approach, and formal problem definition

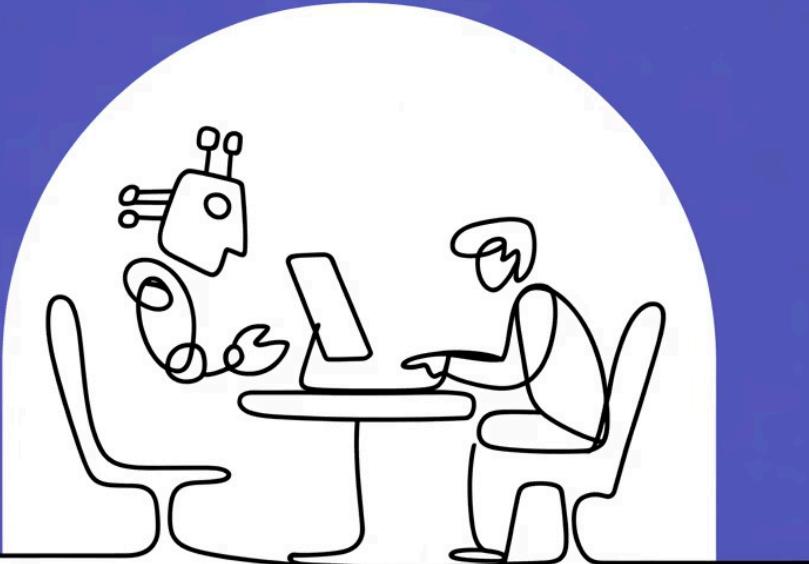
3. Experimental Results

Dataset overview, evaluation metrics, experimental methodology, and results presentation

4. Concluding Remarks

Critical discussion of results and ideas for future work





AI Usage Disclaimer

Parts of this projects have been developed with the assistance of OpenAI's **ChatGPT (GPT-5)**. The AI was used to support the **development of project ideas, the structuring of methodological workflows, the drafting of descriptive texts, and the identification of relevant datasets and references.**

All content produced with AI assistance has been **carefully reviewed, edited, and validated** by me. I take full responsibility for the final content and its accuracy, relevance, and academic integrity.

Using AI (for students)

Generative AI tools (such as ChatGPT, Claude, Mistral, or similar models) **may be used in this project**, both as an object of investigation and as a tool to support the development process.



Encouraged Uses

Explore how models function, interact creatively, leverage for inspiration, ideation, drafting, or experimentation



Important Limitation

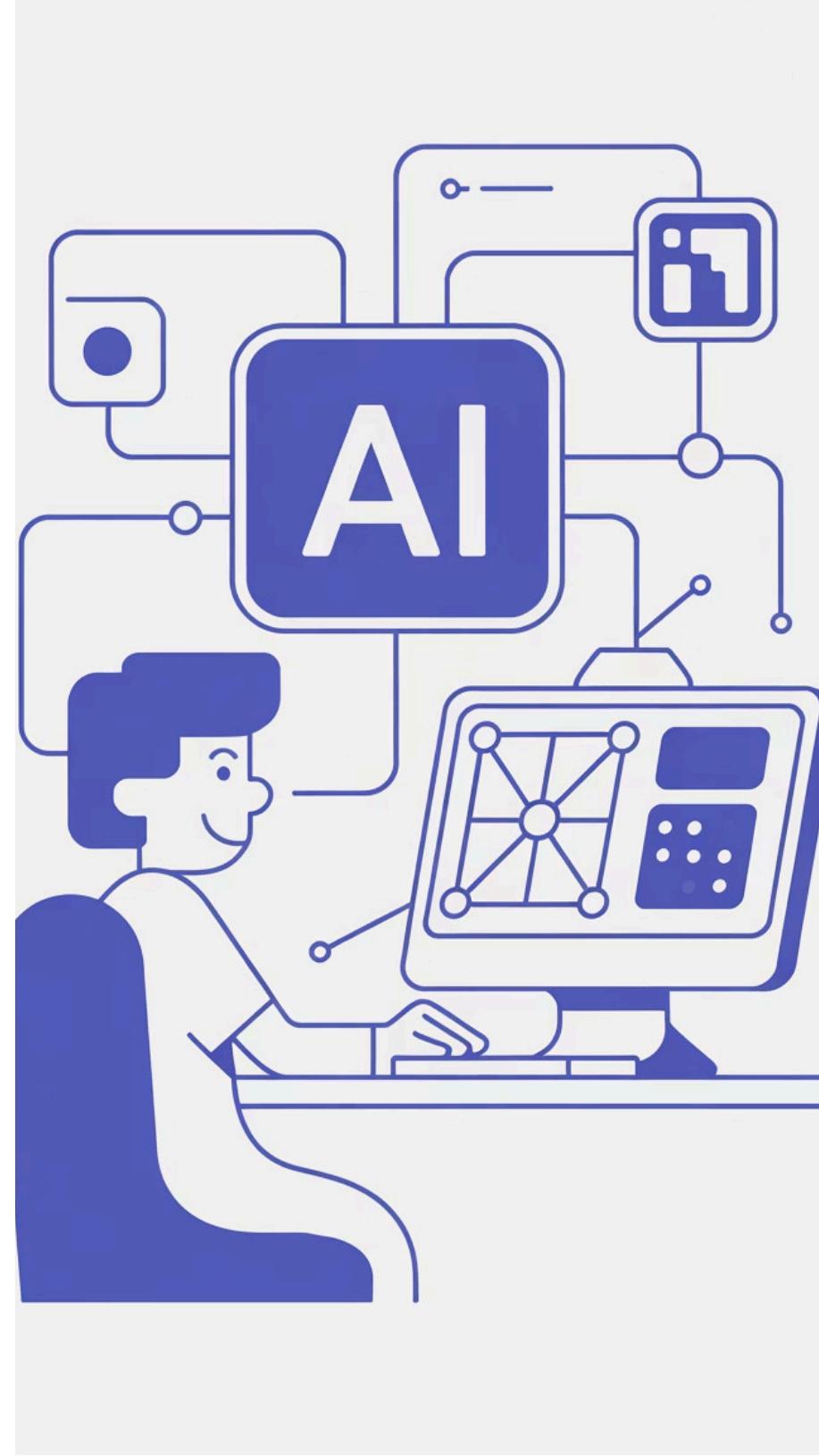
AI should not substitute original work. Student responsibility for structure, reasoning, and understanding remains

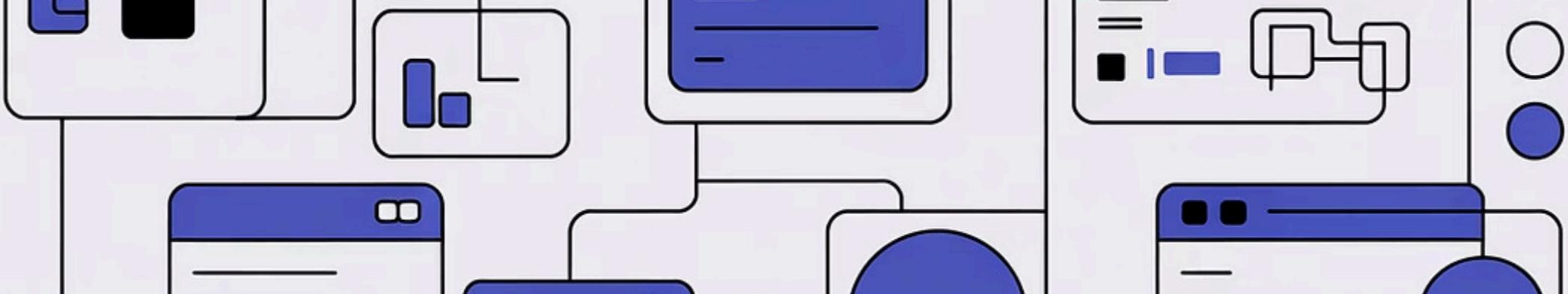


Mandatory Disclaimer

Specify which models used, for what purposes, and to what extent outputs were modified or verified

The project will be assessed on output and the **student's ability to explain and justify all choices made**. A final interview will evaluate depth of understanding.





Instructions on Coding

All project code should be written with **clarity, modularity, and reusability** in mind. The implementation should **not consist of a single large notebook**, but rather follow a structured and maintainable design.

Python Modules & OOP

Organize logic into Python modules and packages using object-oriented programming principles

1

Separation of Concerns

Separate data loading, preprocessing, model interface, evaluation, and visualization

2

Jupyter for Demo

Use notebooks primarily for demonstration, experimentation, and visualization

3

Clean Repository

Ensure clean, reproducible, and extensible codebase for replication and future development

4

Project Ideas

The following are ideas for projects. For each idea, a short description, example of datasets that can be used, and bibliographic references are provided.

Choose Existing Ideas

Students may **choose one of the following** as their project theme, with complete descriptions, methodologies, and references provided.

Projects are organized in thematic clusters. The methodological notes, datasets and references are intended as starting guidelines. Students are encouraged to find their own data and references when needed.

Propose Your Own

Students can **propose their own idea**, structuring the proposal as those presented in this document. Send project description to Prof. Ferrara.



P3. The Negotiation Arena

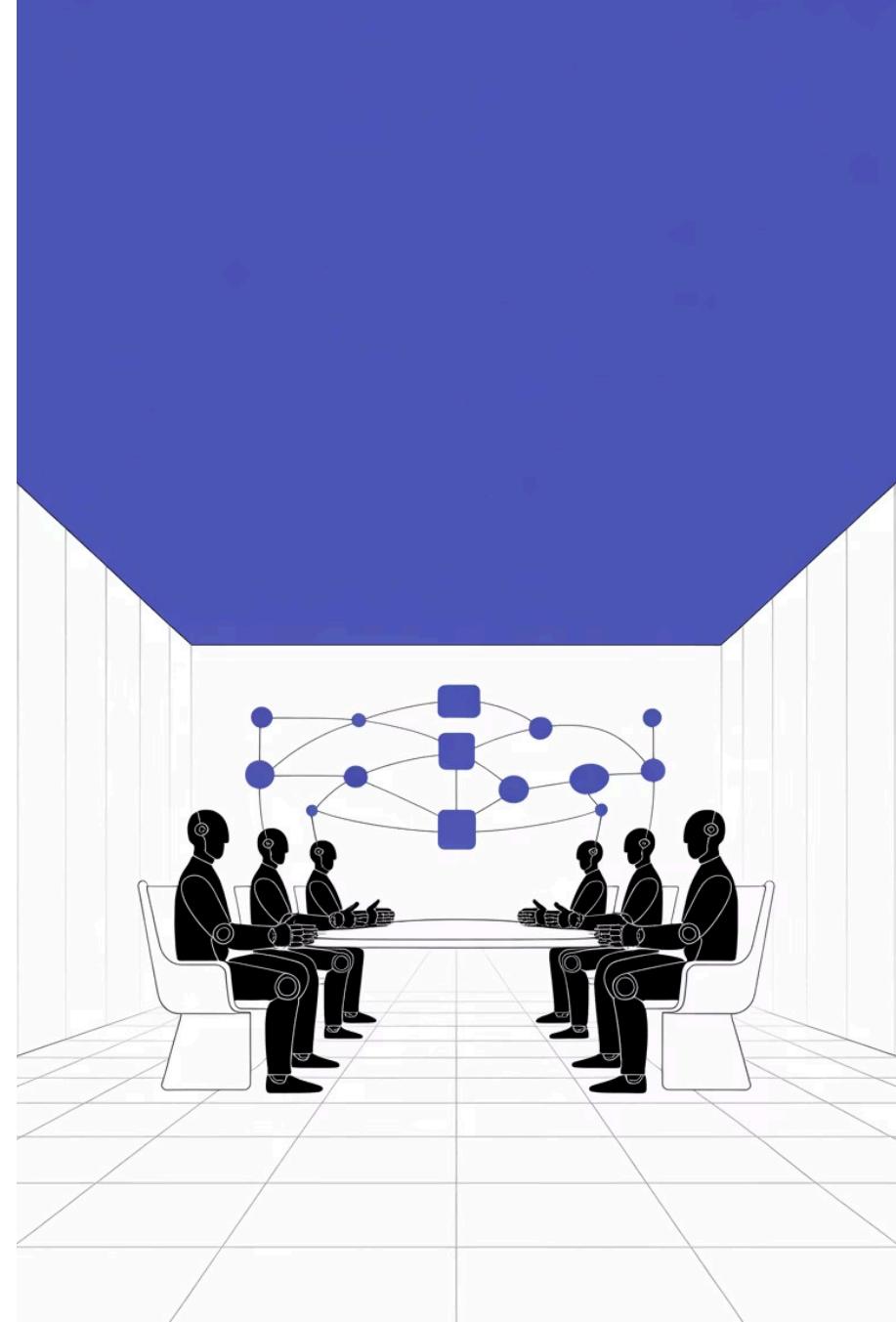
This project investigates how **Large Language Models (LLMs)** behave as autonomous agents engaged in negotiation, cooperation, or strategic dialogue. Two or more models are placed in simulated scenarios where they must **reach an agreement, trade resources, or align on decisions** despite having distinct goals or incomplete information.

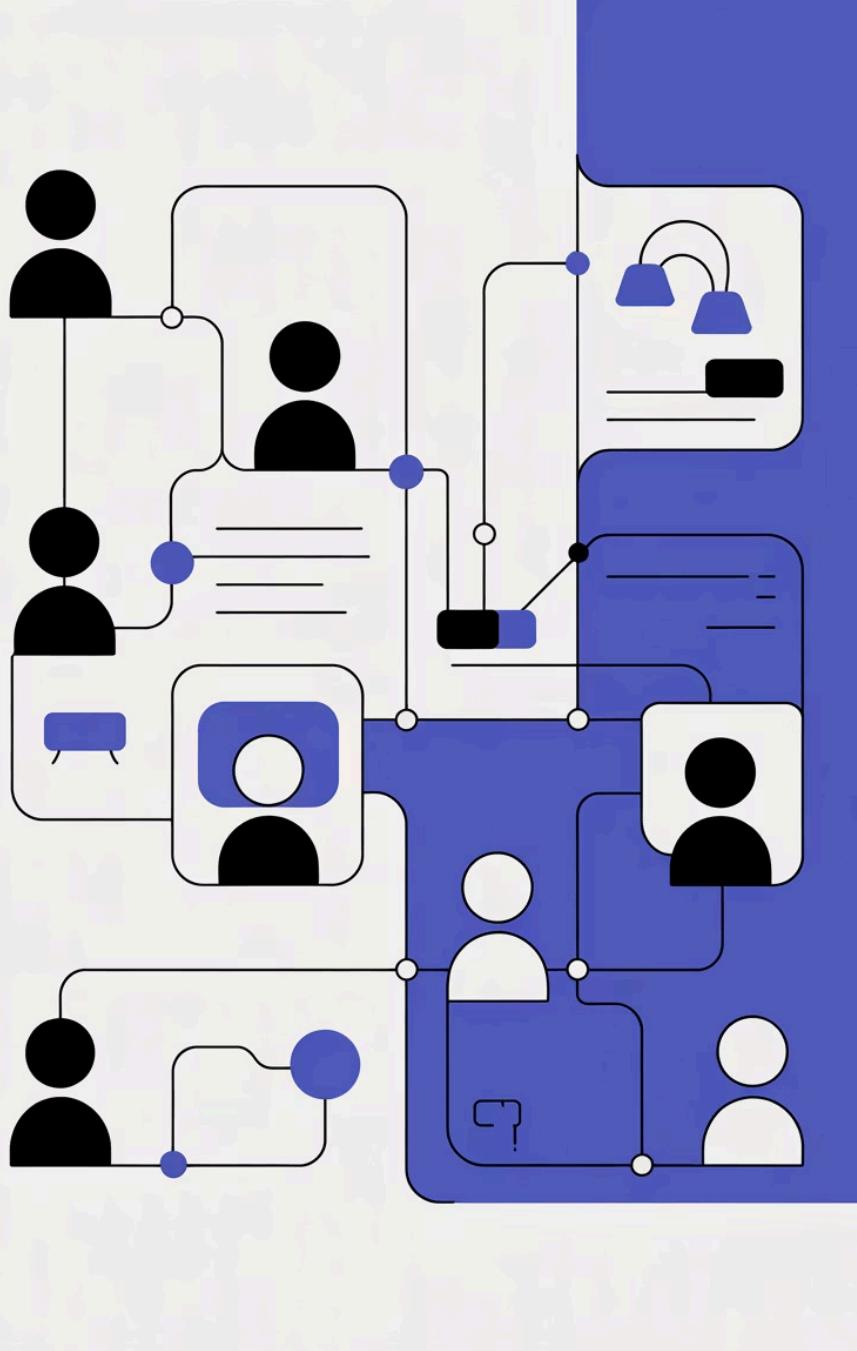
Core Pipeline

Agents instantiated with distinct goals engage in multi-round conversations. Simulations log exchanges, evaluated quantitatively and qualitatively.

Expected Outcomes

Uncover cooperative or adversarial behaviors, identify pragmatic features correlated with negotiation success or failure.





Methodology

1 Scenario Design

Define negotiation settings: resource division, task scheduling, or preference alignment. Each agent receives private information or asymmetric incentives.

2 Agent Configuration

Instantiate LLM agents with distinct personas or objectives. Include optional adjudicator model or human evaluator.

3 Dialogue Simulation

Implement iterative conversation rounds until agreement or impasse. Test cooperative, competitive, and mixed modes.

4 Analysis & Metrics

Measure agreement rate, convergence rounds, utility scores, language complexity. Analyze persuasion tactics and emotional tone.

Dataset & References

Dataset: No fixed dataset required; negotiation scenarios can be **synthetically generated** or adapted from existing dialogue datasets.

References

- Lewis, M., Yarats, D., Dauphin, Y., Parikh, D., & Batra, D. (2017). Deal or No Deal? End-to-End Learning of Negotiation Dialogues. *EMNLP*, 2443-2453.
- Akin, S., et al. (2025). Socialized Learning and Emergent Behaviors in Multi-Agent Systems based on Multimodal Large Language Models. *arXiv preprint arXiv:2510.18515*.
- Gupta, P., et al. (2025). The Role of Social Learning and Collective Norm Formation in Fostering Cooperation in LLM Multi-Agent Systems. *arXiv preprint arXiv:2510.14401*.

