Subject: Internship Proposal

ID	PTI_EN_Ficara Annamaria_13/03/2026 17.56.14
Data	13/03/2026 17.56.14

Project Supervisor

Surname	Ficara
Name	Annamaria
Department	Department of Cognitive Sciences, Psychology, Education, and Cultural Studies
Laboratory	Via Concezione 6/8, A206
E-mail	aficara@unime.it
Phone number	

Project Co-Supervisor

Surname	
Name	
Job Position	
Department	

Laboratory	
E-mail	
Phone number	

Project details

Title	Automatic or Semi-Automatic Reconstruction of Criminal Networks from Judicial Documents

Detailed description: Judicial documents such as pre-trial detention orders contain detailed descriptions of criminal investigations, including information about suspects, intercepted communications, and meetings observed by investigators during surveillance activities. These documents represent a valuable source of information for the reconstruction of interaction structures within criminal organizations.

In many studies on criminal network analysis, such networks are reconstructed manually by carefully reading judicial documents and identifying the individuals involved in the investigation together with the interactions described in the text. In this process, the individuals mentioned in the documents are represented as nodes of a graph, while the interactions between them define the edges of the network.

Interactions described in judicial documents may include different types of relationships between individuals. For example, the documents may report intercepted telephone communications between suspects or physical meetings observed during surveillance activities. These interactions can be represented as edges connecting the individuals involved, allowing the construction of different types of criminal networks derived from the same document.

The aim of this internship is to design and implement a software prototype for the automatic or semi-automatic extraction of criminal networks from judicial documents written in Italian. The objective is to transform the information contained in unstructured legal texts into structured graph representations that can be used for criminal network analysis.

The internship activities will include:
- analysis of the structure and linguistic characteristics of judicial documents;
- text preprocessing and sentence identification;
- identification of individuals involved in the investigation using natural language processing techniques;
- extraction of interactions described in the text, such as telephone communications and

meetings between suspects;
- design of data structures suitable for representing the corresponding criminal networks;
- implementation of a prototype capable of generating the networks automatically or semi-automatically;
- visualization and preliminary network analysis of the generated graphs.
The final goal of the internship is to develop a prototype capable of transforming judicial textual information into graph representations of criminal networks that can support subsequent analytical or investigative studies.
The expected outcome of the internship includes a working software prototype, documented code for text processing and network generation, and a final report describing the methodology, implementation, and experimental results.

Duration (month – max 12)	6
Duration (hours)	75
Open positions	2

Internship Skills

Technical requirements: - Python and related development environments - Natural Language Processing tools and libraries for Italian text processing - Network analysis libraries (e.g. NetworkX) - Data handling libraries (e.g. pandas and NumPy)	
Other skills	- Scientific writing and documentation - Problem-solving attitude - Ability to read and understand technical documentation and research papers