Subject: Internship Project Proposal

Code Name	PTA_EN_133526/2025 _08/01/2026 13.23.05
Date	08/01/2026 13.23.05
Target	Bachelor's Degree - Data Analysis

Host Institution

Host Institution	Cooperativa Servizi Sociali
Protocol	133526/2025
Protocol Date	03/10/2025
Country	Italy
City	Messina
Address	Via Risorgimento 93, Messina (ME), 98123
Website	https://www.coopservizisociali.it
Employees Number	>50
Contant Person	Carmelo Galipò
Phone Number	+393381614242
Email	info@coopservizisociali.it

Project Supervisor

Name and Surname	Carmelo Galipò
Phone Number	3381614242
Email	info@coopservizisociali.it

Internship Project Details

Title	AI-assistant for automatic drafting of medical home-visit reports

Detailed Description: The project aims to design and implement a pipeline that transforms free-text “dictations” of home care visits into a structured draft of ADI visit reports, aligned with predefined templates (e.g., intake form, periodic visit, short clinical-care report). The healthcare context is Integrated Home Care (ADI) in Sicily, where documentation is largely paper-based and poorly structured, creating a high bureaucratic burden for professionals; the micro-assistant supports digitalization and reduces manual form-filling effort using speech recognition outputs and large language models (LLMs).

The student will:

Analyze ADI document models and select 1–2 target templates to be generated.

Build a synthetic dataset of 60–80 “dictations” and corresponding structured records (JSON/CSV) with demographic fields, diagnoses, problems, interventions and follow-up.

Implement a Python NLP pipeline to extract key sections (reason for visit, brief anamnesis, vital signs, interventions, critical issues) and auto-complete ADI forms, with basic quality checks (mandatory fields and reminders for missing information).

Define a data structure compatible with hypothetical integration into an Electronic Health Record (EHR), including patient identifiers and minimal coding of diagnoses/problems, to facilitate future integration in broader e-health systems.

Tools and technologies: Python, standard NLP libraries (e.g., spaCy/transformer-based models), JSON/CSV handling, and optionally simple rule-based components for information extraction; synthetic data only will be used, in full compliance with privacy and data protection constraints.

Topics	Data science / Natural Language Processing for e-health (Integrated Home Care – ADI)
Reimbursement of Expenses (YES/NO)	NO
Refund Amount	0
Availability for Travel (YES/NO)	NO
Kind of employment	Part time
Duration in months (max 12)	3
Duration in hours	150
Internship Date Start	15/02/2026
Internship Date End	15/07/2026
Number of Open Position(s)	2

Internship Skills

Required Skills: Basic knowledge of Python programming and data manipulation (JSON/CSV, pandas). Introductory experience in Natural Language Processing (tokenization, entities/keywords extraction, simple pattern rules). Familiarity with machine learning workflows and version control (e.g., Git).

Interest in e-health, clinical documentation or healthcare information systems.	
Other Skills	Ability to work independently while interacting regularly with the supervisor for design decisions. Precision and attention to detail in handling clinical-like information, even when synthetic. Good written communication skills in English for documenting methods. Problem-solving attitude and willingness to iterate on prototypes based on feedback.