UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PHD PROGRAM IN INFORMATION AND COMMUNICATION TECHNOLOGY FOR HEALTH

Teaching Module Announcement

Title: **Healthcare Text Analytics in the AI Era**

Lecturer/Organizer:

Prof. Vincenzo Moscato, PhD

University of Naples Federico II
Email: vincenzo.moscato@unina.it

***Bio**: Prof. Vincenzo Moscato is a Ful Professor of Database and Big Data Engineering at Department of Electrical Engineering and Information Technologies of University of Naples Federico II, where he received the Ph.D in Computer Science and Engineering. His current research interests lie in the area of data-driven AI applications and Big Multimedia Data analytics.*

External Lecturers:

Valerio La Gatta, PhD

Northwestern University
Email: valerio.lagatta@northwestern.edu

***Bio**: Dr. Valerio La Gatta is a postdoctoral researcher at the Northwestern Security and AI Lab (NSAIL). He received is his MSc degree in Computer Engineering and PhD degree from the University of Naples Federico II. His work encompasses addressing disinformation and harmful content on social media integrating AI and Social Network Analysis methodologies.*

Marco Postiglione, PhD

Northwestern University
Email: marco.postiglione@northwestern.edu

***Bio**: Dr. Marco Postiglione is a postdoctoral researcher at the Northwestern Security and AI Lab (NSAIL). He received is his MSc degree in Computer Engineering and PhD degree from the University of Naples Federico II. His work encompasses healthcare data analytics, deepfake detection, security of AI assets. In 2022, he placed at the first position in the "DISTEMIST: DISease TExt Mining Shared Task" at BioASQ.*

Credits: **6**

Overview

Artificial Intelligence (AI) and Natural Language Processing (NLP) are transforming healthcare by enabling efficient analysis of medical texts, clinical notes, and research literature. This course provides a comprehensive introduction to the use of AI-driven NLP techniques in healthcare, covering foundational concepts, traditional and advanced methods (e.g. Large Language Models, Retrieval Augmented Generation, Agentic AI), preprocessing challenges, and real-world applications.

Students will gain hands-on experience with NLP tools and frameworks, learning how to extract insights from healthcare text data while addressing critical issues such as data privacy, regulatory compliance, and clinical workflow integration. By the end of the course, students will be equipped with the knowledge and skills to design, evaluate, and deploy AI-powered NLP solutions for healthcare applications.

Schedule

Lecture	Date	Time	Topics	Lecturer
1	29 September 2025	10:30 – 12:30	Introduction to the Course	Vincenzo Moscato
2	30 September 2025	15:30 – 17:30	Foundations of Healthcare NLP	Valerio La Gatta
3	3 October 2025	10:30 – 14:30	Traditional NLP Methods for Healthcare	Valerio La Gatta
4	8 October 2025	10:30 – 14:30	Text Preprocessing for Healthcare	Valerio La Gatta
5	10 October 2025	10:30 – 14:30	Advanced AI Methods in Healthcare NLP	Marco Postiglione
6	15 October 2025	10:30 – 14:30	Real-World Applications of AI in Healthcare	Marco Postiglione
7 (first part)	22 October 2025	14:00 – 16:00	Evaluation and Deployment of AI Models in Healthcare	Marco Postiglione
7 (second part)	22 October 2025	16:00 – 17:00	Presentation of project works	Vincenzo Moscato

Room: The course will take place at the ITEM Lab of CINI (Consorzio Interuniversitario Nazionale per l'Informatica), Complesso Universitario di Monte Sant'Angelo, via Cinthia, Napoli.

Content details

Lessons marked with ⍰ include hands-on Python activities. Students will receive code and resources to support their learning.

Lesson 1 – Introduction to the Course: This session provides an overview of how AI and NLP are revolutionizing healthcare through text analysis. Students will explore real-world case studies showcasing AI-driven innovations in clinical decision-making, medical research, and patient care. The session also outlines key course topics and learning objectives.

Part 1 – *Foundations of Healthcare NLP*

Lecturer: Valerio La Gatta

Lesson 2 – Foundations of Healthcare NLP ⍰: Students will build a strong foundation in healthcare text data, exploring various sources such as electronic health records (EHRs), clinical notes, and medical research papers. The session introduces fundamental NLP techniques—including tokenization, lemmatization, and part-of-speech

tagging—while addressing critical privacy and compliance considerations such as HIPAA regulations, Protected Health Information (PHI) handling, and data governance.

Lesson 3 – Traditional NLP Methods for Healthcare : This lecture covers established NLP techniques tailored for healthcare applications. Students will learn how to design rule-based systems, apply regular expressions to medical text, extract key clinical information, and use topic modeling to analyze medical literature. Traditional classification methods for medical text processing will also be explored.

Lesson 4 – Text Preprocessing for Healthcare : This session focuses on specialized preprocessing techniques necessary for handling complex medical texts. Topics include medical terminology normalization, handling domain-specific abbreviations, and processing negations in clinical narratives. Students will implement best practices for cleaning and structuring healthcare text data.

Part 2 – *Advanced AI Methods and Applications in Healthcare*

Lecturer: Marco Postiglione

Lesson 5 – Advanced AI Methods in Healthcare NLP : Students will explore state-of-the-art AI techniques in healthcare text analysis. This session covers medical word embeddings, transfer learning strategies, transformer architectures, and domain-specific large language models (LLMs) designed for healthcare applications. Practical implementations and case studies will highlight how these advanced methods improve clinical text analysis.

Lesson 6 – Real-World Applications of AI in Healthcare : This session delves into real-world AI applications that could enhance clinical workflows and decision-making. Students will examine AI-powered clinical decision support systems, retrieval-augmented generation for medical queries, literature mining methods, and other impactful solutions deployed in healthcare settings.

Lesson 7 – Evaluation and Deployment of AI Models in Healthcare : Students will learn best practices for evaluating and deploying AI models in real-world healthcare environments. Topics include model performance assessment, interpretability techniques, clinical workflow integration, and regulatory compliance considerations. The session emphasizes ensuring AI solutions meet the rigorous standards required for clinical adoption.

To register, participants are requested to complete the form provided at the link below:

https://forms.gle/SiPRs4AV5GCeYpun6

By September 28th, 2025, participants are requested to join the MS Teams group that will be communicated via email. Once accepted in the Teams group, students must fill the "students_info.xlsx" file with their information. The course is in presence. However, students pursuing their PhD period abroad (for research purposes) have the option to request remote attendance for classes via MS Teams. The development of project works during the lesson, which will be presented in the last lesson, is required.

For information: Prof. Vincenzo Moscato (DIETI, UniNA) – vincenzo.moscato@unina.it