# Natural Language Processing for Social Media

## Dott. Federica Iurescia, Dott. Rachele Sprugnoli

***COURSE AIMS AND INTENDED LEARNING OUTCOMES***

This course aims at introducing the students to current research on Natural Language Processing (NLP) tools and methods for processing social media texts, highlighting the challenges to deal with when analysing those texts in contrast with traditional documents (e.g. news).

At the end of the course, students will be able to collect social media data, perform the linguistic pre-processing of social media texts, annotate such texts both manually and automatically by using a set of NLP tools.

***COURSE CONTENT***

The topics covered by this course will be the following:

- introduction to social media analysis;

- social media data collection and annotation;

- linguistic pre-processing of social media texts;

- applications of social media text analysis with a specific focus on Sentiment Analysis and Hate Speech Detection;

- ethical issues in social media analysis.

***READING LIST***

Consultation material will be provided during the course.

The following readings are not mandatory, and serve as references for many topics of the course:

* Jurafsky, Daniel, and James H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Second edition. Drafts of third, updated edition freely accessible at [https://web.stanford.edu/∼jurafsky/slp3/](https://web.stanford.edu/%E2%88%BCjurafsky/slp3/). Englewood Cliffs, NJ, USA: Prentice Hall, 2008.
* Farzindar, Atefeh, and Inkpen, Diana. *Natural language processing for social media*. Third edition. Morgan & Claypool Publishers (Synthesis lectures on human language technologies, Number 46), 2020.

***TEACHING METHOD***

Lectures (in English) with use and demonstration of NLP tools in class.

***ASSESSMENT METHODS AND CRITERIA***

The exam consists of two parts:

1. discussion of a project prepared by the student during the course;
2. questions about the theoretical and practical contents of the course.

The two parts will be considered of equivalent weight when determining the final grade.

***NOTES AND PREREQUISITES***

This course, being an introduction to the discipline, does not require any prior specific expertise; familiarity with basic notions of Computational Linguistics and Natural Language Processing will help.

*Office hours for students*

On appointment, by sending an e-mail to federica.iurescia@unicatt.it, or to rachele.sprugnoli@unipr.it.

Place for in-person meeting: CIRCSE Research Center, Franciscanum building, second floor, room n. 219