NLP session 02: Preparation for next session

- Read Jurafsky & Martin's Chapter 17: "Sequence Labeling for Parts of Speech and Named Entities". While it is useful to read the section, we will not discuss Conditional Random Fields (Section 17.5) in this class. So feel free to skim or skip this part if you're short on time.

- Work through sections 5-9 from Chapter 1 of the free interactive course "Advanced NLP with spaCy": https://course.spacy.io/en/

Recommended

- If we haven't done so in class: Do (at least) the outstanding exercises 4 and 5 from class (Are the recipes vegan or vegetarian? and What are the 5 most frequently mentioned ingredients?).

Optional

There are four topics you could introduce next week. This would count toward your in-class participation grade. If you want to present one of them, announce this on this week's forum on Aula Global. In this way others will know that the topic is already taken.

1. Give an introduction to *Part of Speech Tagging*, based on Jurafsky & Martin's Chapter 17 (from the beginning of the chapter to the end of 17.2)

2. Give an introduction to *Named Entity Tagging*, based on Jurafsky & Martin's Chapter 17 (subsection 17.3, focusing on explaining BIO, IO, and BIOES tagging, and evaluation; subsection 17.6)

3. Introduce HMM Part-of-Speech Tagging, based on Jurafsky & Martin's Chapter 17 (subsection 17.4). This is a more complex topic and you will need more than a few minutes to explain it. Accordingly, it counts for your entire in-class participation grade.

4. If we haven't done so in class: Explain your solution to the outstanding exercises 4 and 5 from class (Are the recipes vegan or vegetarian? and What are the 5 most frequently mentioned ingredients?)