Guest lecture

Wednesday, 11 December 2019, 6:15 p.m.

Seminar für Iberoromanistik, room E005, Maiengasse 51, Universität Basel

Anđelka Zečević
(University of Belgrade)

Computational Linguistics (Natural Language Processing) and its application: an introduction

The domain of Computational Linguistics welcomes both language-lovers and computer scientists. It gets more and more importance and applicability due to the large amount of text we generate in our modern daily lives or inherit from the past. This talk will address a few interesting **research questions** the community has been dealing with:

- How can we represent the semantics of the words computationally?
- Which algorithms hide behind recognition of names of places or dates?
- How do we train intelligent models to automatically classify the text at hand as topically relevant or not?

We will talk about **tools** developed to cope with the complex morphological and syntactic nature of the languages: sentence splitters, tokenizers, part-of-speech taggers, and parsers. For the **real-user experience**, we will go through the development of the platform for the visualization of sustainable development goals, the recent project of the **United Nations Development Program**. We will see how we can single out titles and sentences from the voluntary national reports, extract the relevant computational features, and prepare the representations suitable to assign one of the 17 Sustainable Development Goals to each sentence.

Anđelka Zečević is a teaching assistant at the Faculty of Mathematics, University of Belgrade. She is doing her Ph.D. research in the domain of abstractive document summarization in collaboration with the Microsoft Development Center Serbia. The projects she participated in are quite diverse ranging from the corpus creation of Judeo-Spanish texts to the processing of the legislation in the domain of pharmacology.

Contact: sandra.schlumpf@unibas.ch
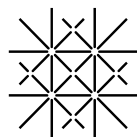
Universität
Basel

Departement
Sprach- und Literaturwissenschaften