



Bilkent EEE



Bilkent EEE Distinguished Seminar Series

Bilkent University - Department of Electrical and Electronics Engineering



InterText: Modelling Text as a Living Object in Cross-Document Context

Iryna Gurevych

Technical University of Darmstadt, Germany

February 28, 2023 – 17:30 (on Zoom)

<https://bit.ly/BilEEESem230228>



Digital texts are cheap to produce, fast to update, easy to interlink, and there are a lot of them. The ability to aggregate and critically assess information from connected, evolving texts is at the core of most intellectual work – from education to business and policy-making. Yet, humans are not very good at handling large amounts of text. And while modern language models do a good job at finding documents, extracting information from them and generating natural-sounding language, the progress in helping humans read, connect, and make sense of interrelated texts has been very much limited.

Funded by the European Research Council, the InterText project brings natural language processing (NLP) forward by developing a general framework for modelling and analyzing fine-grained relationships between texts – intertextual relationships. This crucial milestone for AI would allow tracing the origin and evolution of texts and ideas and enable a new generation of AI applications for text work and critical reading. Using scientific peer review as a prototypical model of collaborative knowledge construction anchored in text, this talk will present the foundations of our intertextual approach to NLP, from data modelling and representation learning to task design, practical applications and intricacies of data collection. We will discuss the limitations of the state of the art, report on our latest findings and outline the open challenges on the path towards general-purpose AI for fine-grained cross-document analysis of texts.

Bio: Iryna Gurevych (PhD 2003, U. Duisburg-Essen, Germany) is Professor of Computer Science and Director of the Ubiquitous Knowledge Processing (UKP) Lab at the Technical University (TU) of Darmstadt in Germany. Her main research interests are in machine learning for large-scale language understanding and text semantics. Iryna's work has received numerous awards. Examples are the ACL fellow award 2020 and the first-ever Hessian LOEWE Distinguished Chair award (2,5 mil. Euro) in 2021. Iryna is the Co-Director of the NLP program within ELLIS, a network of excellence in machine learning. She is currently the president of the Association for Computational Linguistics (ACL). In 2022, she has been awarded an ERC Advanced Grant.

