

KEY NOTE SPEECH

DAN TUFIS

Data and Language-centric AI

Abstract

In our globalized world language barriers affect cross-lingual communication and the free flow of knowledge and information. Language technologies, especially multilingual technologies, can help overcome these omnipresent and ubiquitous language barriers, significantly improving trade, administration, politics, communication and cross-cultural understanding. Language models are successfully used in multi-task settings and as such, the interest in building and using large language models was boosted in several research areas besides NLP.

However, creation of large language models, requires huge quantities of training data, which are not always easy to find, especially for "smaller languages". The EU, for which the multilingualism is at the heart of European principles, engaged in a challenging process to bring all European language on the same technological level by 2030.

The talk will outline the community efforts on the way to Digital Language Equality (the agenda and the roadmap) and steps towards establishing a primary platform and market place for Language Technology in Europe.