

State-of-the-Art and Challenges in Timeline Summarization

Katja Markert

Institute of Computational Linguistics
University of Heidelberg
Heidelberg, Germany

Abstract

Timeline Summarization (TLS) creates an overview of long-running events via dated daily summaries for the most important dates and is essential to keep track of a flood of information on, for example, crises data. TLS differs from standard single and multi-document summarization in the importance of date selection, interdependencies between summaries of different dates, the lack of large-scale human training data and a very low compression rate, i.e. very short summaries in comparison to the number of corpus documents. In this talk, I will discuss the impact these properties have on (i) optimization algorithms and objective functions for timeline summarization algorithms (ii) evaluation of timeline summarization outputs and (iii) the dependence of TLS on information retrieval components.

Short Biography

Katja Markert is Chair of Computational Linguistics at Heidelberg University (Germany) since 2016. After her PhD at Freiburg University (Germany) in 1999, she worked as Postdoc and Emmy Noether-Fellow at The University of Edinburgh, UK, as well as Lecturer and Reader at the University of Leeds, UK, during that time also having longer research stays at the University of Hannover as well as the Heidelberg Institute of Theoretical Studies. She works both in more theoretical computational linguistic areas, such as the automatic recognition of anaphora and figurative language, as well as application areas such as sentiment analysis and summarization. She also is co-director of the Leibniz-Science-Campus "Empirical Linguistics and Computational Language Modeling", which concentrates on NLP models for German, in a joint project between Heidelberg University and the Institute for the German Language, Mannheim.