

# Design and Realization of the EXCITEMENT Open Platform for Textual Entailment

**Günter Neumann<sup>†</sup> and Sebastian Padó<sup>‡</sup>**

<sup>†</sup>DFKI GmbH, Saarbrücken, Germany

<sup>‡</sup>IMS, Universität Stuttgart, Germany

Textual Entailment (TE) is a binary relation between two natural language text which holds if the truth of a first text implies the truth of the second one, or at least makes it very likely. Good methods to recognize TE have the potential to impact many NLP tasks, where the ability to draw conclusions from textual expressed facts is a key challenge. The area of TE has seen the development of a range of algorithms, methods, and technologies over the last decade.

Unfortunately, research on TE (like semantics research more generally), is fragmented into studies focussing on various aspects of semantics such as world knowledge, lexical and syntactic relations, and so on. This fragmentation has problematic practical consequences. Notably, interoperability among existing RTE systems is poor, and reuse of resources and algorithms is mostly infeasible. This also makes systematic evaluations very difficult to carry out. Finally, TE presents a wide array of approaches to potential end users with little guidance on which to pick.

Our contribution to this situation is a novel architecture and platform, the EXCITEMENT Open Platform (EOP), which was developed to enable and encourage the consolidation of methods and resources in the TE area. Starting out from and generalizing over three existing systems (BIUTEE, EDITS, and TIE), our architecture decomposes RTE into components with strongly typed interfaces. The specifications cover (a) a modular linguistic analysis pipeline and (b) a decomposition of the "core" RTE methods into top-level algorithms and subcomponents. We identify four major subcomponent types, including different kinds of knowledge bases. The architecture was developed with a focus on generality, supporting all major approaches to RTE, as well as encouraging language independence.

The practical implementation of this architecture forms the EXCITEMENT open platform (EOP). It is a suite of textual entailment algorithms and components which contains the three systems named above, including linguistic-analysis pipelines for three languages (English, German, and Italian), and comprises a number of linguistic resources. By addressing the problems outlined above, the platform provides a comprehensive and flexible basis for research and experimentation in Textual Entailment. We discuss the current scope and functionality of the platform, which is available as free open source software, and outline existing and future use cases.