Preface

For many decades, NLP has suffered from low software engineering standards causing a limited degree of re-usability of code and interoperability of different modules within larger NLP systems. While this did not really hamper success in limited task areas (such as implementing a parser), it caused serious problems for the emerging field of language technology where the focus is on building complex integrated software systems, e.g., for information extraction or machine translation. This lack of integration has led to duplicated software development, work-arounds for programs written in different (versions of) programming languages, and ad-hoc tweaking of interfaces between modules developed at different sites.

In recent years, the Unstructured Information Management Architecture (UIMA) framework has been proposed as a middleware platform which offers integration by design through common type systems and standardized communication methods for components analysing streams of unstructured information, such as natural language. The UIMA framework offers a solid processing infrastructure that allows developers to concentrate on the implementation of the actual analytics components. An increasing number of members of the NLP community thus have adopted UIMA as a platform facilitating the creation of reusable NLP components that can be assembled to address different NLP tasks depending on their order, combination and configuration.

This workshop aims at bringing together members of the NLP community – users, developers or providers of either UIMA components or UIMA-related tools in order to explore and discuss the opportunities and challenges in using UIMA as a platform for modern, well-engineered NLP.

This volume now contains the proceedings of the 3rd UIMA workshop to be held under the auspices of the German Language Technology and Computational Linguistics Society (Gesellschaft für Sprachverarbeitung und Computerlinguistik - GSCL) in Darmstadt, September 23, 2013. From 11 submissions, the programme committee selected 7 full papers and 2 short papers. The organizers of the workshop wish to thank all people involved in this meeting - submitters of papers, reviewers, GSCL staff and representatives - for their great support, rapid and reliable responses, and willingness to act on very sharp time lines. We appreciate their enthusiasm and cooperation.

September 2013

Peter Kluegl, Richard Eckart de Castilho, Katrin Tomanek (Eds.)