Preface

This volume presents the proceedings of the 4th edition of the *Conference on CMC and Social Media Corpora for the Humanities* (*cmc-corpora2016*) which was held on September 27–28 at the University of Ljubljana, Slovenia. The conference series (http://cmc-corpora.org/) is dedicated to the collection, organization, annotation, processing, analysis and sharing of data and corpora from computer-mediated communication (CMC) and social media genres for research purposes. The genres of interest to the cmc-corpora conference community include e-mail, chats, forums, newsgroups, blogs, news comments, wiki discussions, SMS and mobile messaging applications (WhatsApp, etc.), interactions on social network sites (Facebook, Twitter etc.), on YouTube and in multimodal online environments. The conference brings together research questions from linguistics, philology, communication sciences, media and social sciences with methods, tools and infrastructures from the fields of corpus and computational linguistics, natural language processing, text technology and digital humanities. The focus of the conferences is on

- language-centered research using computational methods and tools for the empirical analysis of CMC and social media phenomena,

- approaches towards automatic processing and annotation of CMC and social media data,

- corpus-linguistic research on collecting, processing, representing and providing CMC and social media corpora on the basis of standards in the field of digital humanities.

Previous conferences have been held in Dortmund/Germany (2013 and 2014) and in Rennes/France (2015).

Besides keynote talks by two invited speakers, *Dawn Knight* from Cardiff University (UK) and *Petra Kralj Novak* from the Jožef Stefan Institute (Slovenia), the 4th cmc-corpora conference featured 17 papers, 4 posters and 1 student paper written by 40 authors and co-authors from 24 research institutions in 11 countries, addressing key issues and current trends in the research field on data from 8 different languages.

We thank all colleagues who have contributed to the conference and to this volume with their papers, talks and posters, and as members of the scientific committee. We hope that the results of the conference will mark another step towards a lively exchange of approaches, expertise, resources, tools and best practices between researchers and existing networks in the field and pave the ground for future standards in building and using CMC and social media corpora for research in the humanities.

Darja Fišer, University of Ljubljana, Slovenia
Michael Beißwenger, University of Duisburg-Essen, Germany

Co-chairs of the Scientific Committee

Ljubljana and Essen, September 2016