Preface

Welcome to Galway for the first international conference on Language, Data and Knowledge (LDK 2017). It is hosted by the Insight Centre for Data Analytics at the National University of Ireland, Galway.

This the first of a new biennial conference series on Language, Data and Knowledge (LDK) which aims to bring together researchers from across disciplines concerned with the acquisition, curation and use of language data in the context of data science and knowledge-based applications. With the advent of the Web and digital technologies, an ever increasing amount of language data is now available across application areas and industry sectors, including social media, digital archives, company records, etc. The efficient and meaningful exploitation of this data in scientific and commercial innovation is at the core of data science research, employing NLP and machine learning methods as well as semantic technologies based on knowledge graphs.

Language data is of increasing importance to machine learning-based approaches in NLP, Linked Data and Semantic Web research and applications that depend on linguistic and semantic annotation with lexical, terminological and ontological resources, manual alignment across language or other human-assigned labels. The acquisition, provenance, representation, maintenance, usability, quality as well as legal, organizational and infrastructure aspects of language data are therefore rapidly becoming major areas of research that are at the focus of the conference.

Knowledge graphs are an active field of research concerned with the extraction, integration, maintenance and use of semantic representations of language data in combination with semantically or otherwise structured data, numerical data and multimodal data among others. Knowledge graph research builds on the exploitation and extension of lexical, terminological and ontological resources, information and knowledge extraction, entity linking, ontology learning, ontology alignment, semantic text similarity, Linked Data and other Semantic Web technologies. The construction and use of knowledge graphs from language data, possibly and ideally in the context of other types of data, is a further specific focus of the conference.

A further focus of the conference is the combined use and exploitation of language data and knowledge graphs in data science-based approaches to use cases in industry, including biomedical applications, as well as use cases in humanities and social sciences.

The LDK conferences have been initiated by a consortium of researchers from the Insight Centre for Data Analytics, InfAI (University Leipzig) and Wolfgang Goethe University and a Scientific Committee of leading researchers in Natural Language Processing, Linked Data and Semantic Web, Language Resources and Digital Humanities. LDK is endorsed by several international organisations: DBpedia, ACL SIGANN, Global Wordnet Association, CLARIN and Big Data Value Association (BDVA).

There were a total of 68 submissions reviewed, of which 15 were accepted as full papers (one was subsequently withdrawn) and 19 as short papers (an acceptance rate of 50%).

Reviewing was single blind (due to the difficulty in making language resources anonymous). There were at least 3 reviews per paper, with an average of 3.7. Papers came from all over the world.

A successful conference is the result of many people's efforts and contributions. The papers included in the program are the results of our participants' contributions as well as the insightful reviews from the program committee members. Besides the oral and poster paper presentations, the program is enriched by the keynote and invited speakers: Dr. Chris Welty, Senior Research Scientist at Google in New York, and an Endowed Professor of Cognitive Systems at the VU University, Amsterdam; Kathleen R. McKeown, Henry and Gertrude Rothschild Professor of Computer Science and Director of the Institute for Data Sciences and Engineering at Columbia University; Antal van den Bosch, director of the Meertens Institute in Amsterdam and Professor of Language and Speech Technology at the Centre for Language Studies at Radboud University, Nijmegen, the Netherlands; and Graham Isaac of the Irish Department at NUI Galway. On behalf of the program committee, we express our thanks to them all.

We would like to thank the scientific advisory committee for their guidance, and the local organizing committee led by Dr John P. McCrae and Dr Paul Buitelaar, for their great work in ensuring the success of this conference.

Finally, we wish that you will all enjoy the conference presentations, discussions, and exchanges between old and new friends at the beautiful harbor city of Galway.

April 2017   Jorge Gracia
                                                                                        Francis Bond