Language, Data and Knowledge
(LDK 2021) conference

From September 1-4 this year, the third edition of the Language, Data and Knowledge (LDK 2021) conference will take place as a hybrid event in Zaragoza, Spain. This conference will build on previous successful editions in Leipzig, Germany in 2019 and in Galway, Ireland in 2017. It aims to bring together researchers from a variety of disciplines working on linguistic data and data science. This includes applications in data science, natural language processing and machine learning as well as digital humanities and commercial applications of these technologies. This year the conference is supported by the NexusLinguarum COST Action for a "European network for web-centred linguistics data science."

Language data is of vital importance to a large number of methodologies in machine learning, natural language processing and Semantic Web research, and these applications crucially depend on linguistic and semantic annotation and the existence of high-quality language resources. As such the conference is concerned with the acquisition, provenance, representation, maintenance, usability, quality as well as legal, organizational and infrastructure aspects of linguistic data. In addition, the application of knowledge graphs and their exploitation in use cases in industry, such as biomedical, fintech and legaltech applications, as well as use cases in humanities, and social sciences is a particular focus of the conference.

This year, we are delighted to welcome three distinguished invited speakers: Mathieu Lafourcade of the University of Montpellier will talk about his work on Games With A Purpose (GWAP) and how these can be used in the development of lexical and semantic networks. Sara Tonelli from the Fondazione Bruno Kessler will give a talk entitled "A Smell is Worth a Thousand Words: Olfactory Information Extraction and Semantic Processing in a Multilingual Perspective." Finally, Mikel Forcada of the University of Alicante will present the development of the Apertium dictionaries in a talk entitled "Free/open-source machine translation for the low-resource languages of Spain." Besides this we have an exciting programme with 21 accepted papers and 17 posters to be presented over the two main conference days.

There is also a packed programme of workshops and tutorials both before and after the main conference, with September 1 hosting

LANGUAGE, DATA and KNOWLEDGE 2021

LDK 2021
3rd Conference
on Language, Data and Knowledge
1-4 September 2021
Zaragoza, Spain – hybrid
http://2021.ldk-conf.org/

tutorials on the Linked Data in Latin (LiLa) project and the DBpedia project. In addition, there is the "1st Workshop on Sentiment Analysis & Linguistic Linked Data", "4th shared task for Translation Inference Across Dictionaries", "2nd International Workshop on Artificial Intelligence for Historical Image and Visual Cultural Artefacts Enrichment" and a "Multisensory Data & Knowledge Workshop", all taking place on that day. After the main conference on September 4, there are meetings of World Wide Web Consortium (W3C) community groups, in particular the Linked Data for Language Technologies (LD4LT) group will meet to discuss the harmonization of linguistic annotation, while the Ontology-Lexicon (OntoLex) group will consider its ongoing work.

Due to the ongoing pandemic, we have made this conference a hybrid event so that all participants can attend. We aim to provide live streams of the talks so that remote participants can be fully involved. In addition, there is no registration fee for either remote or in-person participation due to the support of NexusLinguarum. We hope to see you all either in person or online on September 1-4 for an exciting conference!

John P. McCrae and Thierry Declerck

LDK 2021 Chairs

John P. McCrae is an assistant professor at the Data Science Institute, National University of Ireland Galway, where he leads the research group on linguistic data. He is the coordinator of Prêt-à-LLOD project, leads the work on linked data in the European Lexicographic Infrastructure (ELEXIS) project, vice-chair of NexusLinguarum COST Action and an Irish Research Council Laureature.

Thierry Declerck is a Senior Consultant at the German Research Center for Artificial Intelligence (DFKI GmbH) since 1996, working in the field of multilingual language technologies and currently responsible for its contribution to Prêt-à-LLOD project. He also contributes for the Austrian Center for Digital Humanities to ELEXIS and serves as the Scientific Communication Manager of NexusLinguarum.