

# About

Language, Data and Knowledge (LDK) aims at bringing 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.

## Invited Speakers

We are happy to announce **Diana Maynard** (University of Sheffield), **Ruben Verborgh** (Ghent University), and **Ruth Wodak** (Lancaster University/University of Vienna), as keynote speakers for LDK 2023.

## Workshops and tutorials

### 12.09.2023

<b>8:30 – 9:00</b>	Registration
	W3C Community day
<b>9:00 – 13:30</b>	Disinformation and Toxic Content Analysis ProLingKNOWER
<b>14:00-18:00</b>	W3C Community day Tutorial: Perspectivized Multimodal Datasets
<b>18:30</b>	Research Meets Industry Meetup

### 13.09.2023

<b>8:30 – 9:00</b>	Registration
	DisLiDaS
	4LR
<b>9:00 – 13:00</b>	TermTrends
	Tutorial: DBpedia KG
	DisLiDaS
<b>14:00 – 18:00</b>	DeepLearning and BATS
	SALLD-3
	Tutorial: LODification Wikibase
<b>18:30</b>	Wine reception



LANGUAGE, DATA and  
KNOWLEDGE 2023

# LDK 2023

## 4th Conference on Language, Data and Knowledge

*Location:* Vienna, Austria  
*Dates:* 12 - 15 September



This article/publication is based upon work from COST Action Nexus Linguarum, supported by COST (European Cooperation in Science and Technology).

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.



# Programme

## DAY 1 (14.09.2023)

- 8:30 – 9:00** Registration
- 9:00 – 9:30** Welcome session
- 9:30 – 10:30** Keynote by Diana Maynard: "Towards an Early Warning System for Online and Offline Violence"
- 10:30-11:00** Coffee break

### SESSION 1: Lexicons in Language, Data and Knowledge

- 11:00 – 11:30** *Refinement of the Classification of Translations – Extension of the vartrans Module in OntoLex-Lemon*, by Frances Gillis-Webber
- 11:30 – 12:00** *Leveraging DBnary Data to Enrich Information of Multiword Terms in Wiktionary*, by Gilles Sérasset, Thierry Declerck, Lenka Bajčetić
- 12:00 – 12:30** *Lexico-Semantic Mapping of a Historical Dictionary: An Automated Approach with DBpedia*, by Sabine Tittel
- 12:30 – 14:00** Lunck Break

### SESSION 2: Digital Humanities and Under-Resourced Languages

- 14:00 – 14:30** *Linking the Computational Historical Semantics corpus to the LiLa Knowledge Base of Interoperable Linguistic Resources for Latin*, by Giulia Pedonese, Flavio Massimiliano Cecchini, Marco Carlo Passarotti
- 14:30 – 15:00** *Graph Databases for Diachronic Language Data Modelling*, by Barbara McGillivray, Pierluigi Cassotti, Davide Di Pierro, Paola Marongiu, Anas Fahad Khan, Stefano Ferilli, Pierpaolo Basile
- 15:00 – 15:30** *Contextual Profiling of Charged Terms in Historical Newspapers*, by Ryan Brate, Marieke van Erp, Antal van den Bosch
- 15:30 – 16:00** *The Cardamom Workbench for Historical and Under-Resourced Languages*, by Adrian Doyle, Theodorus Fransen, Bernardo Stearns, John Philip McCrae, Oksana Dereza, Priya Rani
- 16:00 – 16:30** Coffee break

### SESSION 3: Sentiment and Natural Language Inference

- 16:30 – 17:00** *Sentiment Inference for Gender Profiling*, by Manfred Klenner
- 17:00 – 17:30** *Multimodal Offensive Meme Classification with Natural Language Inference*, by Shardul Rajendra Suryawanshi, Mihael Arcan, Suzanne Little, Paul Buitelaar
- 17:30 – 18:30** Poster session
- 18:45** Social event
- Conference dinner

## DAY 2 (15.09.2023)

- 9:00 – 10:00** Keynote by Ruben Verborgh: "Decentralizing data and knowledge via Solid pods"

### SESSION 4: Transformer-Generated Programs and Pre-Trained Language Models

- 10:00 – 10:15** *Decoding Strategies for Code Conciseness and Efficiency in Transformer-Generated Programs*, by Brendan Kondracki
- 10:15 – 10:30** *MEAN: Metaphoric Erroneous Analogies dataset for PTLMs metaphor knowledge probing*, by Lucia Pitarch, Jordi Bernad, Jorge Gracia
- 10:30 – 11:00** Coffee break

### SESSION 5: Corpora and annotation

- 11:00 – 11:30** *An Empirical Analysis of Task Relations in the Multi-Task Annotation of an Arabizi Corpus*, by Elisa Gugliotta, Marco Dinarelli
- 11:30 – 12:00** *Crowdsourcing OLIA Annotation Models the Indirect Way*, by Christian Chiarcos
- 12:00 – 12:30** *Towards ELTeC-LLOD: European Literary Text Collection Linguistic Linked Open Data*, by Ranka Stanković, Christian Chiarcos, Miloš Utvić, Olivera Kitanović
- 12:30 – 14:00** Lunch break

### SESSION 6: Human-Machine Annotation and Question Answering in Linked Data

- 14:00 – 14:30** *Human-Machine Collaborative Annotation: A Case Study with GPT-3*, by Ole Magnus Holter, Basil Eli
- 14:30 – 15:00** *A framework for the automatic induction of ontology lexica for Question Answering over linked data*, by Mohammad Fazleh Elahi, Basil Eli, Philipp Cimiano
- 15:00 – 16:00** Keynote by Ruth Wodak: "Austrian Identity Politics – Past and Present: A Discourse-historical Analysis"
- 16:00 – 16:30** Coffee break

### SESSION 7: Language Data – Use Cases and Applications

- 16:30 – 16:45** *Unifying Emotion Analysis Datasets using Valence Arousal Dominance (VAD)*, by Mo El-Haj, Ryutaro Takanami
- 16:45 – 17:00** *Challenges and Solutions in Transliterating 19th Century Romanian Texts from the Transitional to the Latin Script*, by Marc Eduard Frincu, Ioana Simina Frincu, Marius Penteliuc
- 17:00 – 17:15** *A variationist analysis of two French attitude expressions: je pense and je crois*, by Delin Deng
- 17:15 – 17:30** *Making Non-Normalized Content Retrievable – A Tagging Pipeline for a Corpus of Expert-Layperson Texts*, by Christian Lang, Ngoc Duyen Tanja Tu, Laura Zeidler
- 17:30 – 18:00** Awards & Closing



## Conference venue

The main conference on 14 and 15 September will take place in the main building of the University of Vienna, Universitätsring 1, 1010 Vienna.

The workshop days on 12-13 September will take place at the Centre for Translation Studies, University of Vienna, Gymnasiumstraße 50, 1090 Vienna

More information scanning the following QR:



<http://2023.ldk-conf.org>