# Dagstuhl Summer School: 3rd Summer Datathon on Linguistic Linked Open Data

# (SD-LLOD-19, <http://datathon2019.linguistic-lod.org>)

**The 3rd Summer Datathon on Linguistic Linked Open Data (SD-LLOD-19)** will be held from May 12 to 17, 2019 in Schloss Dagstuhl – Leibniz Center for Informatics, Wadern, Germany (see). SD-LLOD-19 is organized in conjunction with and held before the 2nd Conference on Language, Data and Knowledge (LDK-2019), May 20-22, Leipzig, Germany (<http://2019.ldk-conf.org>).

The SD-LLOD datathon has the main goal of giving people from industry and academia practical knowledge in the field of Linked Data and its application to natural language data and natural language annotations, from areas as diverse as knowledge engineering, lexicography, the language sciences, natural language processing and computational philology.

We provide tutorials and supervision by leading experts on Linked Open Data and its application to language resources, human language technology and the language sciences. The final aim is to enable participants to migrate their own linguistic data and publish them as Linked Data on the Web.

The SD-LLOD datathon is unique in its specialization worldwide and continues a series of international hackathons and summer schools organized since 2012. This edition is supported by several international projects addressing different aspects of natural language processing, lexicography and digital humanities: The Research Group “[Linked Open Dictionaries (LiODi)](http://www.acoli.informatik.uni-frankfurt.de/liodi/)”, funded by the German Federal Ministry of Education and Research (BMBF), the Horizon 2020 Research and Innovation Action “[Prêt-à-LLOD. Ready-to-use Multilingual Linked Language Data for Knowledge Services across Sectors](http://www.pret-a-llod.eu/)” and the Horizon 2020 Research and Innovation Action “[ELEXIS. European Lexicographic Infrastructure](https://elex.is)”.

During the datathon, participants will:

* Generate and publish their own Linguistic Linked (Open) Data from some existing data sources or existing tools.
* Apply Linked Data principles and Semantic Web technologies (Ontologies, RDF, Linked Data) to language resources and human language technology.
* Use the principal models for representing Linguistic Linked Open Data, in particular [Ontolex-Lemon](https://www.w3.org/2016/05/ontolex/), [Web Annotation](https://www.w3.org/blog/news/archives/6156) and knowledge representation vocabularies such as SKOS and OWL.
* Gather experiences with terminology resources developed for or used in the Linguistic Linked Open Data context, such as [lexvo](http://www.lexvo.org/), [lexinfo](http://lexinfo.net/), [OLiA](http://purl.org/olia/) and [GOLD](http://linguistics-ontology.org).
* Learn about multilingual dataset linking against knowledge bases from the Web of Data, such as [DBpedia](http://dbpedia.org/) or [BabelNet](http://babelnet.org/).
* Learn about benefits and applications of linguistic linked data for specific use cases.

The program of the summer datathon will contain three types of sessions:

1. **Seminars** to show novel aspects and discuss selected topics.
2. **Practical sessions** to introduce the basic foundations of each topic, methods, and technologies and where participants will perform different tasks using the methods and technologies presented.
3. **Hacking sessions,** where participants will follow the whole process of generating and publishing Linguistic Linked Data with some existing data set.

Attendants are invited to propose a miniproject related to the topic and to bring to the datathon some dataset of linguistic data produced by their organizations  to work with during the hacking sessions and transform into linked data. Participants who cannot provide their own linguistic dataset can join another’s miniproject.

Participants should bring their own laptops to follow and participate in the hacking sessions, they will be provided with digital copies of all necessary material used during the course and will have assistance for installing all the required software.

## Invited Speakers

[Gerard de Melo](http://gerard.demelo.org/), Deep Data Lab, Dept. of Computer Science, Rutgers University, New Jersey

[Richard Eckart de Castilho](https://www.informatik.tu-darmstadt.de/ukp/ukp_home/staff_ukp/detailseite_mitarbeiter_1_42176.en.jsp), Ubiquitous Knowledge Processing (UKP) Lab, Dept. of Computer Science, Technische Universität Darmstadt, Germany

## How to Apply

We invite scholars, students and practitioners from any field of lexicography, the language sciences, natural language processing, knowledge engineering, and computational philology to apply for participation. Further details under <https://datathon2019.linguistic-lod.org>.

Participants are encouraged to submit a short abstract of their ideas for the datathon (description of possible resources to be converted, linked or reused during the datathon, ideas for use cases, etc.). A selection of them will be presented during the event and proposed as miniprojects to be developed during the datathon. Please note that the number of participants is strictly limited, so we encourage to apply as soon as possible.

## Important dates

|  |  |
| --- | --- |
| Registration opens: | January 25, 2019 |
| Registration closes: | April 4, 2019 |
| Notification: | April 12, 2019 |
| Datathon: | May 12 to 17, 2019 |

## Questions/Applications

[datathon@linguistic-lod.org](mailto:datathon@linguistic-lod.org)

## Organizers

[Christian Chiarcos](http://acoli.informatik.uni-frankfurt.de/en.html) Goethe Universität Frankfurt, Germany

[Jorge Gracia](http://jogracia.url.ph/web/) University of Zaragoza, Spain

[John P. McCrae](https://john.mccr.ae/) National University of Ireland, Galway, Ireland