One currently pivotal global challenge for scientific research in the digital age concerns the potential contradiction between (largely) automatized processing of an ever-growing amount of data and the need for validating, verifying and securing results. This two-day conference will illustrate how these essential challenges regarding data provenance, collection, storage, processing and interpretation are tackled in a number of different disciplines such as physics, bioinformatics, materials science and the digital humanities as represented by computational linguistics. In addition to gathering state-of-the-art facts and insights from these different subjects, the conference aims at promoting exchange and reflection from a broader, interdisciplinary perspective. The focus will thereby lie on methodological issues and deliberately refrain from addressing -equally essential- ethical and legal aspects.

Confirmed speakers and preliminary topics:

**April 8**

10:30  Conference opening
Peter Strohschneider (President of the DFG)

10:45  Traceability, Reproducibility, Replicability…
What It Means for Computational Linguistics
Nancy Ide (Vassar College)

11:45  Towards (more) transparent Natural Language Processing technologies: How teaching others about our tools forces us to ask the right questions
Antske Fokkens (VU Amsterdam)

14:00  Models of Provenance
Peter Buneman (University of Edinburgh)

15:00  FAIR data: The European Galaxy Server
Björn Grüning (Freiburg University)

16:30  Can knowledge inform Machine Learning?
Christian Bauckhage (Fraunhofer IAIS/University of Bonn)

**April 9**

09:00  Robust and reliable machine learning
Matthias Hein (University of Tübingen)

10:00  Towards Reproducibility in Machine Learning and AI
Kristian Kersting (TU Darmstadt)

11:30  Traceability in materials design: A use case from molecular simulation
Chandler Becker (National Institute of Standards and Technology)

14:00  Automizing work flows in computational materials design
Jörg Neugebauer (Max-Planck-Institut für Eisenforschung)

15:00  What is a measurement record?
Michael-Andreas Esfeld (Lausanne University)

16:30  Mastering complex data processing procedures: from particle detector measurements via machine learning algorithms to physics results
Martin Erdmann (RWTH Aachen)