Co-organization and Funding

The conference is co-organized by the Institute for Documentology and Scholarly Editing, the Academy Junior professorship for Digital Humanities at the University of Rostock, the Know Center and the Centre for Information Modelling at the University of Graz. It is funded by the University of Rostock and supported by the NEISS project.

Universität Rostock

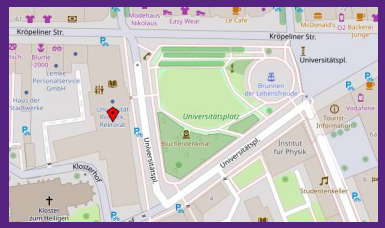
Universität Rostock Traditio et Innovatio

Venues

Talks:
HS 218, Universitätshauptgebäude
Universitätsplatz 1, 18055 Rostock

Breaks:
SR 017, Universitätshauptgebäude
Universitätsplatz 1, 18055 Rostock

Dinner:
Braugasthaus Zum alten Fritz
Warnowufer 65, 18057 Rostock



Organisation

Roman Bleier
Bernhard Geiger
Fabian Kaßner
Ulrike Henny-Krahmer
Marc Lemke
Gerlinde Schneider
Martina Scholger

Mail: ml-dse@i-d-e.de

https://www.i-d-e.de/ml-dse

Universität Rostock Traditio et Innovatio

Machine Learning and Data Mining for Digital Scholarly Editions

University of Rostock
9-10 June 2022

#mldse

Juniorprofessur für Digital Humanities
Institut für Germanistik
Universität Rostock

In several areas of the Digital Humanities, Data Mining and Machine Learning techniques are increasingly applied and discussed, for example for the processing and extraction of information from digital images that represent humanistic sources, or for the analysis of full texts that are relevant for the Humanities and have already been extracted from images or are born digital.

As a subfield of Digital Humanities, Digital Scholarly Editing is no exception to this trend. Data Mining and Machine Learning methods have been used for several tasks in the Digital Scholarly Editing workflow. However, compared to other Digital Humanities subfields such as Computational Linguistics or Computational Literary Studies, so far, these methods have not reached the same widespread use and are not yet discussed as intensely and fundamentally in Digital Scholarly Editing.

The main purpose of this conference is to foster the discussion on Machine Learning and Data Mining techniques in the area of Digital Scholarly Editing, by addressing the following questions, among others:

- Where can Machine Learning and Data Mining be usefully and meaningfully applied in a Digital Scholarly Editing workflow?
- How are Machine Learning and Data Mining already used for the creation of Digital Scholarly Editions and what are potential use cases for the future?
- Do editions pose special challenges to the application of Machine Learning and Data Mining that need to be overcome?
- How does the use of Machine Learning and Data Mining change the way editors work and the way editions are created? Does it change the role of the editor? How does it change the methods of editing?

Thursday, 9 June

1.30 p.m. Arrival

1.50 p.m. Welcoming

2-3.30 p.m. Moderation: Martina Scholger

Tobias Hodel (University of Bern):

Machine Learning Approaches to Making and Using Digital Editions: Premodern Documents, Text, and Named Entity Recognition

Richard Hadden, Nina Rastinger, Matthias Schlögl (Austrian Academy of Sciences):

Data Mining and Data Sowing: Automated Methods for Building a Digital Scholarly Edition of Historical Newspapers

3.30-4 p.m. Coffee break

4-5.30 p.m. Moderation: Ulrike Henny-Krahmer

Sandra Bläß, Marie Flüh, Julia Nantke (Hamburg University), David Maus (University Library Hamburg):

Quality Management for Machine Generated Data in Digital Scholarly Editions – Possibilities and Challenges

Daniel Stökl Ben Ezra (École Pratique des Hautes Études), Hayim Lapin (University of Maryland):

From HTR to Digital Critical Scholarly Edition: Reflexions on the Use of Machine Learning, Computational and Digital Humanities in the Sofer Mahir Project

5.30-6 p.m. Break

6 p.m. Keynote: Roger Labahn (University of Rostock): Machine Learning & Digital Humanities – a personal perspective

Dinner together

Friday, 10 June

9-10.30 a.m. Moderation: Marc Lemke

Werner Scheltjens (University of Bamberg):

Extracting Preindustrial Logistics Patterns from a Digital Edition: Reflections on a Workflow for Data Mining the Schenkenschans Customs Registers (1630-1810)

Manuel Portela (University of Coimbra):

Text Mining the Book of Disquiet for Critical and Creative Explorations

10.30-11 a.m. Coffee break

11-12.30 p.m. Moderation: Martina Scholger

Ulrike Henny-Krahmer (University of Rostock), Frederike Neuber (Berlin-Brandenburg Academy of Sciences and Humanities):

Topic Modeling in Digital Scholarly Editions

Marc Lemke, Konrad Sperfeld, Jochen Zöllner (University of Rostock):

Introducing NTEE: An Easy to Use Tool to Enrich TEI Files with Entities Based on State of the Art Neural Networks

12.30-13 p.m. Closing discussion

13 p.m. Lunch together / Goodbye