Summer School:

“Large Language Models for Digital Humanities Research”

8.-11. September 2025

Introductory Workshop (Day 1+2):

“Introduction to LLMs and Prompt Engineering”, given by Christopher Pollin (Digital Humanities Craft OG)

Abstract:

This workshop provides a hands-on introduction to key concepts of generative AI, including tokenization, embeddings, and context windows in LLMs. Participants will systematically learn fundamental prompting techniques and practice creating effective prompts in interactive exercises. The workshop also addresses common challenges such as bias and hallucinations, and how to identify and mitigate them effectively.

4 Parallel Tracks (Day 3+4):

1) **„Large Language Models for Qualitative Data Analysis“**, given by Sarah Oberbichler (Leibniz Institute of European History, Mainz)

Abstract:

Recent AI developments have opened new perspectives for qualitative research. While corpus-based discourse or content analysis traditionally treated texts as a "bag of words" and identified linguistic patterns based on frequency distributions, Large Language Models (LLMs) offer a significantly more profound ability to process language, semantics, and context. This workshop will introduce how LLMs can be integrated into qualitative research by providing insights into their capabilities and limitations, alongside hands-on training in prompt engineering, model implementation, and critical bias evaluation—all demonstrated through practical analysis of earthquake newspaper reporting and their argumentative units.

2) **„Deep Learning in the Computational Literary Studies“**, given by Janis Pagel (University of Cologne)

Abstract:

"Computational Literary Studies (CLS) is an emerging field within the Digital Humanities that applies computational methods to literary texts and tries to answer literary research questions with computational means. In this course, we will look at different methods involving deep neural networks and how they are used in CLS. A special focus will be on large language models (LLMs), current state-of-the-art for using LLMs in CLS, as well as potential problems, risks and obstacles when using LLMs to answer literary research questions."

3) **„LLM-Supported Modeling, Operationalization, and Exploration for Digital Editions“**, given by Christopher Pollin (Digital Humanities Craft OG)

Abstract:

This workshop introduces practical methods for applying LLMs in the context of digital editions. Through concrete examples and exercises in prompting and programming, participants will learn how LLMs can be used for modeling, operationalization, and exploratory analysis of text and data collections. The workshop also covers specific programming techniques for edition-related tasks such as frontend development, data enrichment, and data transformation.

4) **“Using LLMs for Psycholinguistic Research”**, given by Hanna Woloszyn, Job Schepens and Benjamin Gagl (University of Cologne)

Abstract:

Are you curious about how artificial intelligence is reshaping the study of language and the mind? This two-day workshop is designed especially for students and researchers who are interested in the intersection of language, cognition, and technology.

We'll explore how large language models (LLMs) can be used as tools for psycholinguistic research. Through accessible sessions, you'll learn:

- How these models are influencing debates in linguistic and cognitive theory

- Real-world examples of LLMs used in behavioral, neuronal, and linguistic data analysis

- Applying these models to your research questions—no coding skills required!

This hands-on and discussion-based workshop includes a session by Job Schepens, focusing on text analysis to evaluate LLM output (i.e., how lexically rich the output is), a session by Hanna Woloszyn on how to extract how large language models represent meaning and how this can be compared with how humans store semantic information. Benjamin Gagl will further provide insights into how we can use the central training task of LLMs (i.e., next-word prediction) to investigate human predictive language capacities.

Whether you're a linguist, psychologist, philosopher, or cognitive scientist, you'll leave with new tools and ideas for your work.