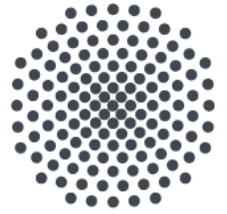# Advanced Seminar (Hauptseminar)

Visualization for Natural Language, Text, and Documents

Text can be seen as a visual representation of natural language if printed or displayed on a screen. While text representations are not considered data visualizations, text is an integral part of many visualization approaches. This seminar explores how textual information and data visualization can complement each other in various scenarios and tasks.

The beneficial combination of both modalities has a long history, from early illustrations in printed books to modern data-driven journalism. Recent developments include notebook-style editors and IDEs incorporating both for data science applications. Small visual views can

Figure 1: DocuCompass, Han et al., 2018 [1]

be integrated into textual documents; texts can be automatically enriched with data visualization, and vice versa, linking both modalities and making them interactively explorable.

Figure 2: Analyzing LLM behaviour, Satkunarajan

Besides approaches directly integrating text and visualization, a large body of research supports text analysis and text mining tasks, mainly to facilitate the comprehension of the results and their quality. But even in such directed workflows, it is only possible to think of one modality together with the other since visualization always needs to link back to the textual information to make visualized facts interpretable.
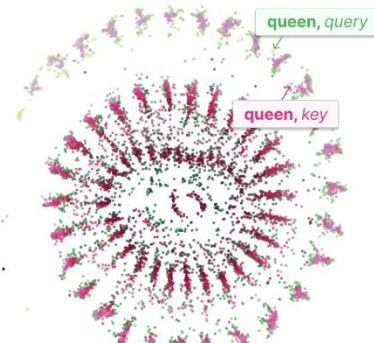
Other methods aim to explain text processing approaches, from simple analyses to modern large language (LLMs, e.g., GPT [3]) or multi-modal models. Goals vary from understanding the inner workings of text processing methods to detecting biased outcomes. Visualization typically plays a crucial role in such approaches for explainable artificial intelligence (XAI).

Figure 3: Global Transformer Attention, Yeh et al., 2023 [2]

In summary, methods involving a combination of data visualization and digital text have seen an unprecedented increase in relevance for industry, commerce, journalism, personal use, and academia. The seminar will provide students with the state of the art in this flourishing field.

The seminar participants will present and discuss the role of text representation and visualization in the abovementioned areas. A subtopic will be assigned to each attendee which they have to research related work for. Based on articles provided by the seminar organizers and the researched ones, all participants will give a presentation explaining their topic to the other seminar participants. In addition, all participants have to write an article summarizing and highlighting important details of their respective topic.

Language

English

Dates

The seminar will take place in the summer semester 2024 at the Computer Science building in room 0.118 on Mondays at 9:45 am starting on April 8th. There will be no preparation meeting, subtopics will be presented during the first seminar class.

Contact

Dr. Steffen Koch, Prof. Dr. Michael Sedlmair

http://www.vis.uni-stuttgart.de

References:

[1] Han, Qi, et al. "Visual quality guidance for document exploration with focus+ context techniques." *IEEE Transactions on Visualization and Computer Graphics* 26.8 (2019): 2715-2731.

[2] Yeh, Catherine, et al. "AttentionViz: A Global View of Transformer Attention." *arXiv preprint arXiv:2305.03210* (2023).

[3] https://openai.com/