

Bachelor's/Master's Thesis: Incorporating NLP-Based Explanations in XAI Dashboards

Supervisor: Philipp Hansen, M.Sc. (philipp.hansen@uni-passau.de)

Start date: At the next possible date

Motivation and Goals

Dashboards are graphical user interfaces that allow for a concise and comprehensible presentation of data to users (Becker et al., 2023). In the realm of Explainable Artificial Intelligence (XAI), they typically enable users to assess a machine learning model's output and, e. g., recognize its flaws (Becker et al., 2023). Yet, visualizations can be quite difficult to comprehend and may even presuppose visualization literacy on behalf of the users (Hohman et al., 2019). If the transparent interaction of users with the dashboard is hampered, they might not be able to extract relevant information (Ruoff & Gnewuch, 2021). Natural Language Processing (NLP) can foster human-machine communication by enabling Artificial Intelligence (AI) systems to interact with their users through human language (Becker et al., 2023).

Following the Design Science Research methodology (Hever et al., 2004), the goal of the thesis is to develop an NLP-based explanation feature for an XAI dashboard. To this end, a bachelor's thesis is expected to comprise a small-scale evaluation of the feature, e.g., with fellow students, while a master's thesis is expected to comprise a full experiment, either as a within-subject or between-subject design.

Required Skills

- Strong interest in (X)AI and NLP
- Good English skills
- Ideally, prior coding experience

Starting Literature (Topic)

Becker, M., Vishwesh, V., Birnstill, P., Schwall, F., Wu, S., & Beyerer, J. (2023). RIXA - Explaining Artificial Intelligence in Natural Language. *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, 875–884. <https://doi.org/10.1109/ICDMW60847.2023.00118>

Hohman, F., Srinivasan, A., & Drucker, S. M. (2019). TeleGam: Combining Visualization and Verbalization for Interpretable Machine Learning. *2019 IEEE Visualization Conference (VIS)*, 151–155. <https://doi.org/10.1109/VISUAL.2019.8933695>

Ruoff, M., & Gnewuch, U. (2021). Designing Conversational Dashboards for Effective Use in Crisis Response. *ECIS 2021 Research-in-Progress Papers*. 45. https://aisel.aisnet.org/ecis2021_rip/45

Starting Literature (Method)

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8. <https://doi.org/10.1016/j.jebo.2011.08.009>

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>

- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to Design Science Research. In J. vom Brocke, A. Hevner, & A. Maedche (Eds.), *Design Science Research. Cases* (pp. 1–13). *Springer International Publishing*. https://doi.org/10.1007/978-3-030-46781-4_1