Contribution ID: **9**   Type: **not specified**

Extracting Measurements from (legacy) publications

Scientific Publishing has built the basis of knowledge exchange since the inception of the modern scientific method. Papers of last centuries contain uncountable experimental and theoretical findings. When exploring new materials or their facets, it becomes instrumental to extract these information from a myriad of papers. In this work, we present first attempts to extract viable physics information from existing publications using large language models. While the extraction of clearly defined terminology is straight forward, we showcase results with more vague information. We will compare our findings using different input formats and language models such as Mixtral 8x7B. We hope to start a conversation in the community how to bridge the paper-to-data-table gap in our community.

Primary authors: STEINBACH, Peter (HZDR); NIEHOFF, Timo; GOTTSCHALL, Tino

Presenter: STEINBACH, Peter (HZDR)

Track Classification: Prompt Engineering