## Analysing Scientific Publications Using AI-based Large Language Models (LLMs)

Scientific publications analysis, a special kind of Document Analysis (DA), is the quantitative and qualitative analysis of the content of a publication with the sole purpose of making more sense of the written content, generating more insights beyond the abstract and making the publication document more understandable to readers.

The main goal of the thesis project is to develop the expertise and tools needed to extract, summarise and analyse scientific publications written in LaTeX.

The proposed thesis project research methodology shows that the candidate will review LaTeX source scientific publications and apply his/her programming skills to segment the documents into manageable chunks that the LLM of an AI tool (e.g., ChatGPT, LangChain) can understand. Using a suitable LLM, she/he will *generate, analyse and compile the summaries* into a ''new'' article for various audiences.
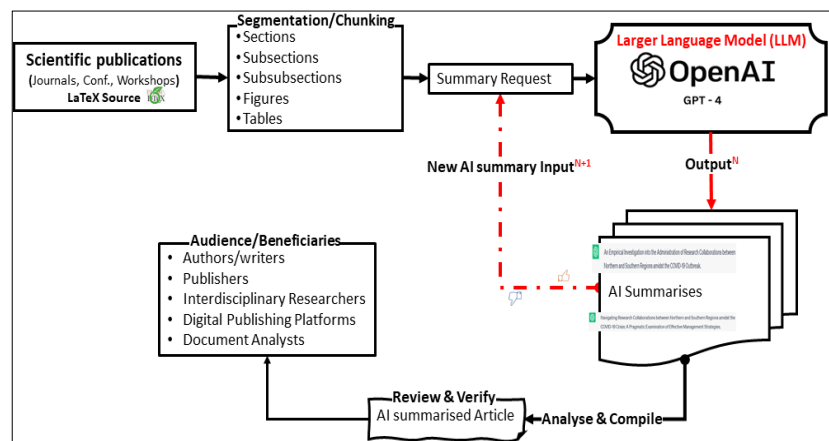
Master's Thesis Project Research Methodology

The candidate is welcome to join and get support from the DBIS/Fraunhofer FIT AI(LLM) working group to help you understand the latest LLMs R&D trends. In collaboration with your advisor, you will review and verify the outputs of the AI summaries. You are encouraged to publish your findings, tools, code, and the systems you used. Your advisor and supervisor will support you in documenting the lessons you have learnt, the research challenges you encountered, and the future research directions you plan to undertake.

Skills you need or are willing to learn to succeed:	**Opportunities and benefits:**
1. Python or a suitable programming language. 2. Experience in using GitHub repositories. 3. Knowledge of APIs 4. How to use and write documents in LaTeX. 5. Text mining, NLP, and Neural networks. 6. Good German and English reading and writing skills. 7. Ability to quickly adapt to working in a large multiculturally academic environment.	1. Get support in learning practical skills to prepare you for the ''world of work''. 2. Learn to write and co-publish scientific papers with expert senior researchers and professors. 3. Opportunities to travel to present your research at international conferences and workshops. 4. Opportunity and support to take your research to the next level (PhD).

Opportunities RWTH your thesis project: In a world dominated by AI, the ***demand*** for computer scientists and software engineers with AI-based document analyst knowledge and expertise is limitless. For example, big companies like IBM, SAP, Dexpro.de, Deutsche Bank, and Google's Document AI Solutions, to mention a few, use AI to analyse various documents and workflows.

Interested, eager to start and have fun?

Contact Advisor: Dr. Sulayman K Sowe (sowe@dbis.rwth-aachen.de)
Supervisor: Prof. Stefan Decker
For more information, please visit:
- Information about Thesis Process

Scan me!