

## Analysing Policy Documents Using AI-based Large Language Models (LLMs)

Policy documents analysis, a special kind of Document analysis (DA), can be defined as the quantitative and qualitative analysis of the content of a policy document with the sole purpose of making more sense of the written content, generating more insights beyond the executive summary (if present), and making the policy more understandable to a broader audience.

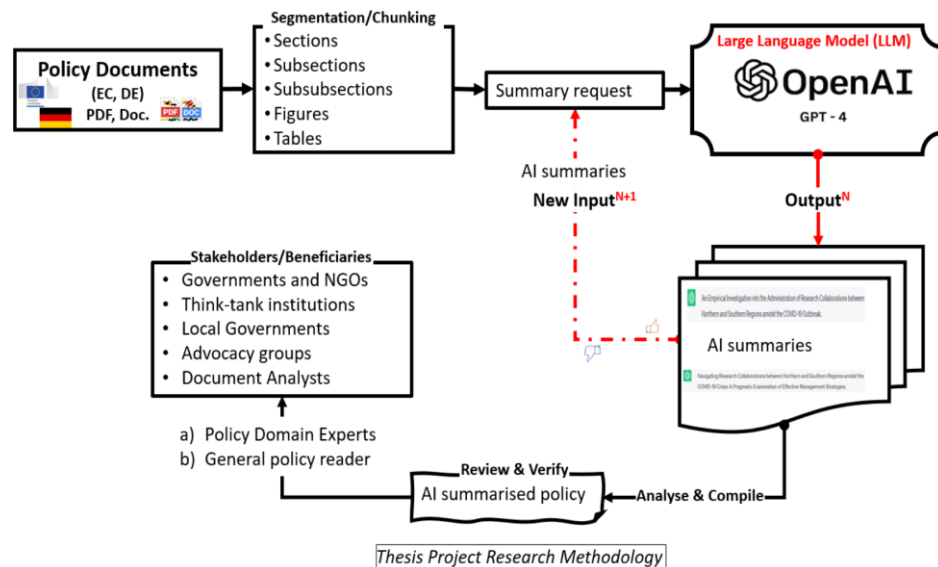
*The main goal of the thesis project is to develop the expertise and tools needed to summarise policy documents (in pdf or doc format) for various stakeholders.*

Thanks to recent advances in AI and LLMs that use Natural Language Processing and artificial neural networks, it is now possible to generate humanlike "intelligent" summaries that are incomparable to what current tools and services offer.

The proposed thesis project research methodology shows

that the candidate will review published two or three European Commission or German Government policy documents as a case study. The candidate will apply his/her programming skills to divide the policy document into manageable chunks that the LLM of an AI tool (e.g., ChatGPT, LangChain, etc.) can understand. She/he will use a suitable LLM to generate summaries for various stakeholders - with different expertise and interest in the summarised policy.

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:
<ol style="list-style-type: none"> <li>1. Python or a suitable programming language.</li> <li>2. Experience in using GitHub repositories.</li> <li>3. Knowledge of APIs</li> <li>4. Text mining, NLP, and Neural networks.</li> <li>5. Good German and English reading and writing skills.</li> <li>6. Ability to quickly adapt to working in a large multiculturally academic environment.</li> </ol>	<ol style="list-style-type: none"> <li>1. Get support in learning practical skills to prepare you for the "world of work".</li> <li>2. Learn to write and co-publish scientific papers with expert senior researchers and professors.</li> <li>3. Opportunities to travel to present your research at international conferences and workshops.</li> <li>4. Opportunity and support to take your research to the next level (PhD).</li> </ol>

In a world dominated by AI, the **demand** for 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](mailto:sowe@dbis.rwth-aachen.de))

Supervisor: Prof. Stefan Decker

For more information, please visit: [Information about Thesis Process](#)



Scan me!