VotingAid: A Retrieval-Augmented LLM Approach for Political Decision-Making*

Nicolas Mahn[1,*], Nadeem Abbas[1]

[1]Linnaeus University, Universitetsplatsen 1, 352 52 Växjö, Sweden

Abstract

The increasing prevalence of disinformation in political discourse, exacerbated by generative AI and social media algorithms, has heightened the need for tools that assist voters in making informed decisions. This paper presents *VotingAid*, a Retrieval-Augmented Generation (RAG) system integrated with Large Language Models (LLMs) to analyze and compare political party manifestos. Unlike traditional Voting Advice Applications (VAAs), which rely on predefined expert-curated statements, *VotingAid* allows users to express their political concerns freely, dynamically retrieving and comparing relevant manifesto content. The system employs OpenAI's `text-embedding-3-small` model for semantic search within a ChromaDB-powered database. Our implementation demonstrates that *VotingAid* can accurately retrieve relevant party positions while ensuring transparency by citing source documents. The tool is accessible via a user-friendly interface, enabling voters to engage with political content interactively. Our findings suggest that RAG-based systems offer a promising approach to voter education, with potential applications in combating misinformation and improving democratic engagement. Future improvements will focus on refining response generation, optimizing performance, and incorporating additional data sources for broader political analysis.

Keywords

Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Voting Advice Application (VAA), Political Manifesto Analysis

1. Introduction

"Disinformation is on the rise" [1]. Aided by generative AI, it has become increasingly easy for malicious actors to spread falsehoods. Examples of this phenomenon are abundant. Taiwan's elections in January 2024 demonstrated the extensive use of misinformation by the People's Republic of China to influence the island's political landscape [2]. Similarly, in February 2024, France discovered Russian disinformation campaigns targeting Europe [3]. The impact of such tactics was starkly evident in December 2024, when Romania's presidential election was annulled after the constitutional court ruled that one of the main candidates had "illegally benefited from the abusive exploitation of social-media platform algorithms" [4].

The proliferation of disinformation, fuelled by generative AI and facilitated by social media algorithms, underscores the urgent need for systems capable of countering fake news. In February 2025, Germany held federal elections, which also suffered from Russian disinformation campaigns [5]. To combat this growing threat, it is imperative to design and develop tools and solutions that empower citizens to discern factual information from falsehoods, especially during election campaigns and when evaluating political party manifestos. To that end, this study explores the use of the Retrieval-Augmented Generation (RAG) [6] and Large Language Models (LLM) [7] to analyze and compare manifestos of political parties, educate citizens about party policies and assist them in selecting a party which is right for them.s More specifically, this study aims to answer following two research questions:

RQ1: How can RAG-DBs and LLMs be used to assess and compare the positions of different political parties based on publicly available data?

*IVUS 2025: Information Society and University Studies 2025, May 15, Kaunas, Lithuania
*Corresponding author.
nm222xc@student.lnu.se (N. Mahn); nadeem.abbas@lnu.se (N. Abbas)
https://github.com/NicolasMahn (N. Mahn); https://lnu.se/personal/nadeem.abbas/ (N. Abbas)

RQ2: How can RAG-DBs and LLMs assist voters in making informed decisions?

To answer these questions, this study explores the integrated use of RAG with LLMs to analyze and compare party manifestos. It introduces a tool, namely *VotingAid*, which combines a structured database with advanced information retrieval and generation techniques. The tool enables efficient access to relevant manifesto content, allowing for clear and unbiased comparisons of political positions. By leveraging RAG technologies and advanced prompt engineering techniques, this study seeks to address critical gaps in current methodologies and contribute to the development of tools for combating misinformation. Unlike tools such as Germany's Wahl-O-Mat[1], provided by the German Federal Agency for Civic Education, which relies on a predefined set of propositions created by experts and political professionals, *VotingAid* employs a more dynamic, bottom-up approach. In the Wahl-O-Mat, users respond to preset statements designed to reflect the political discourse of the election period, enabling a "top-down query" approach that matches users' responses with political party stances. However, this predefined structure does not independently capture the actual concerns of individual voters, as the focus remains on preselected issues already present in the party landscape. By contrast, *VotingAid* enables users to articulate their own political preferences, allowing for greater personalization and inclusivity in addressing nuanced or user-specific concerns. This approach aligns with recent advancements in Voting Advice Applications (VAAs), which emphasize the role of RAG-supported LLMs in providing voter-centric recommendations by prioritizing user-defined political concerns over predefined propositions [8].

The remainder of this paper is organized as follows. Section 2 presents and discusses related work, highlighting previous research on Voting Advice Applications (VAAs) and the integration of Retrieval-Augmented Generation (RAG) with Large Language Models (LLMs). Section 3 describes the research method used for this study. Section 4 introduces the tool developed as a result of this study. Section 5 discusses key challenges and limitations, as well as planned improvements to enhance the system. Finally, Section 6 concludes the paper and outlines directions for future research.

2. Related Work

Recent research highlights the potential of integrating Retrieval-Augmented Generation (RAG) techniques with Large Language Models (LLMs) to address the limitations of conventional VAAs. Gittmann et al. (2024) present the development of *wahlweise*, a RAG-supported LLM-based VAA designed for the 2024 state elections in Germany. Unlike traditional VAAs, *wahlweise* adopts a bottom-up appeal, like *VotingAid* does too. It enables users to freely articulate their political concerns and benchmark them against party policies. This approach shifts the focus from predefined propositions to user-centric issues, promoting democratic engagement and reducing biases associated with top-down query models like *Wahl-O-Mat* [8].

Moreover, Gittmann et al. emphasize the importance of transparency, security, and ethical considerations in VAA design. By anonymizing party references and utilizing rigorous security testing, *wahlweise* ensures political neutrality and minimizes risks associated with biased outputs. The integration of RAG techniques allows the system to generate fact-based responses from party manifestos, addressing the limitations of hallucination and bias often associated with standalone LLMs. This innovative framework exemplifies the potential of RAG-supported LLMs to create impartial and user-driven advisory tools for voters.

Gittmann et al.'s research aligns closely with the objectives of *VotingAid*, which also leverages RAG and LLMs to provide personalized, unbiased insights into political party positions. While *wahlweise* and *VotingAid* share similarities in their use of advanced technologies, *VotingAid* further emphasizes user-defined preferences, allowing for dynamic and flexible comparisons of party manifestos. In contrast, *wahlweise* focuses on providing users with an AI assistant that explains the manifestos. Together, these studies highlight the transformative potential of AI-driven VAAs in modernizing democratic engagement and addressing the challenges of misinformation and bias in electoral processes.

[1]https://www.deutschland.de/en/topic/politics/wahl-o-mat-bundestag-election-election-manifesto-parties

An important part of this project is the extraction of relevant information from both the initial user query and from the manifestos of political parties. A paper by Polak and Morgan [9] introduces a method called "ChatExtract," which aims to extract relevant data from research papers. The authors implement a series of prompts using multi-step prompt engineering to improve the accuracy of data extraction. This paper is particularly useful for developing multi-step instructions for *VotingAid*. Specifically, their flowcharts provide a useful template workflow [9, Figures 1-3, especially Figure 2]. The authors demonstrate very high performance in their approach, achieving 90.8% precision and 87.7% recall for extracting bulk modulus values [9, pg. 9]. Additionally, they effectively minimize hallucinations by incorporating redundancy through follow-up prompts [9, pg. 3].

3. Method

This study follows the Design Science Research Methodology (DSRM) [10], structured in phases to guide the development of *VotingAid*.

3.1. Problem Identification and Motivation

The widespread dissemination of disinformation and the prevalence of biased political narratives create challenges for voters seeking reliable information. Many existing VAAs are based on a predefined set of statements, which restricts their ability to capture the nuanced concerns of individual users. *VotingAid* addresses these limitations by allowing users to freely articulate their political interests, ensuring a more voter-driven, bottom-up approach that provides personalized recommendations.

3.2. Objectives of a Solution

The solution aims to retrieve relevant manifesto sections for user queries, ensuring that responses are based solely on party manifestos. Transparency is prioritized by providing users with references to source documents. The system must also be able to dynamically adjust to different user-defined concerns, allowing for a more flexible approach to political comparisons. Additionally, the tool should present its results in a way that is both comprehensible and accessible, fostering informed decision-making among users.

3.3. Design and Development

The *VotingAid* system consists of several key components (see Figure 1). A database stores embeddings of party manifestos, which are processed using OpenAI's `text-embedding-3-small` model. A Retrieval-Augmented Generation (RAG) pipeline powered by ChromaDB enables efficient retrieval and comparison of relevant manifesto content. The frontend, built with Dash, provides an intuitive interface where users can input their political preferences and receive detailed comparisons. The system also employs a multi-step prompt engineering strategy to refine responses, ensuring that outputs remain accurate and informative.

Arrows in the diagram denote both direct data transfers, such as embedding queries, database storage, and retrieval, as well as logical or conceptual processes, including the analysis of relationships between user opinions and party positions.

3.4. Demonstration

Users interact with *VotingAid* by entering their political concerns into the system. The input is processed by embedding it into a numerical representation, which is then compared against the stored manifesto embeddings. The system retrieves the most relevant manifesto sections based on similarity scores and generates responses using the GPT-4o mini model. The results are displayed to the user along with citations, ensuring transparency and allowing users to verify the source of the information presented.

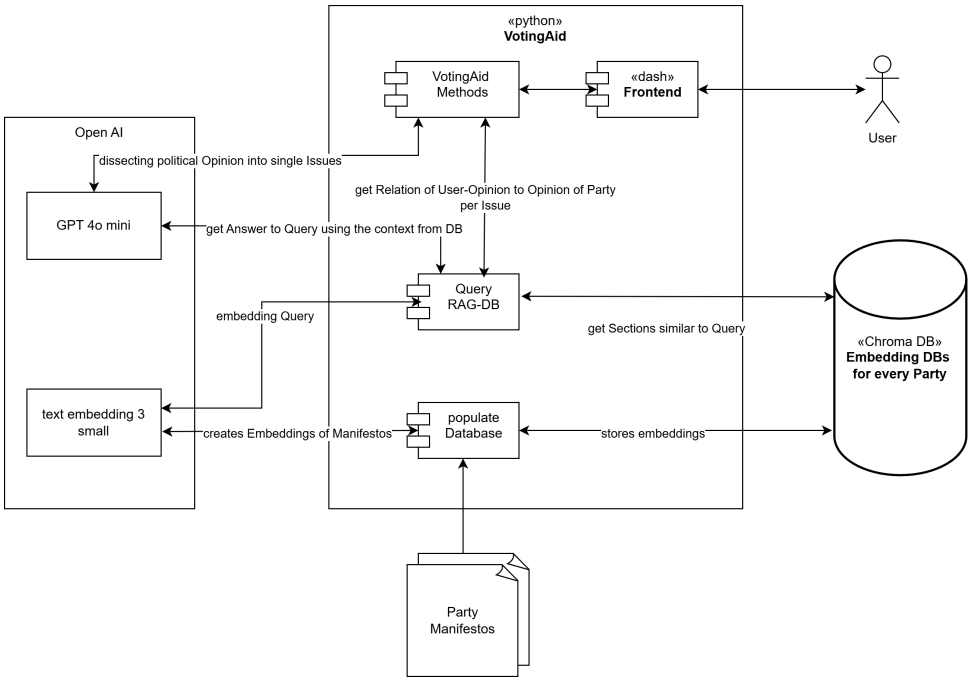

Figure 1: System Diagram of *VotingAid*

4. Results: Tool Presentation

The implementation of *VotingAid* demonstrates its ability to analyze user-defined political preferences and compare them against party manifestos. The tool is publicly accessible at https://votingaid.nicolas-mahn.de/. Users interact with the system by inputting a topic and a specific stance, after which the tool retrieves relevant sections from party manifestos and presents them in a structured format. The system ensures transparency by linking each result back to the original source documents.

An example of the system's output is presented in Figure 2. In this instance, a user entered the topic "Erneuerbare Energien" (Renewable Energy) and expressed the view that "Erneuerbare Energien sollen vom Staat gefördert werden" (Renewable energy should be supported by the state). The image displays the two parties whose manifestos the system identified as the most relevant to the user's stance: *Bündnis 90/Die Grünen* (The Greens) and *Die Linke* (The Left), both of which showed a 100% alignment with the user's position. Party manifestos assessed as less aligned with the user's stance are listed further down.

If the user provides a more detailed prompt, the system will return a more detailed response. With the "Erweitere die Analyse um ein Thema" button, users can also add another topic to expand their query, allowing them to compare multiple policy areas simultaneously.

Both results are accompanied by links to relevant sections of the party manifestos, allowing users to verify the information themselves. The blue, bracketed text represents the chapter titles of the most relevant sections in the party manifestos.

To reassure users of the results' objectivity, they can have individual positions reassessed using the "Nochmal analysieren" button.

The interface is designed to be accessible and easy to navigate, ensuring that users can quickly compare political party positions in an unbiased manner. This example illustrates how *VotingAid* provides a transparent, data-driven approach to political decision-making.

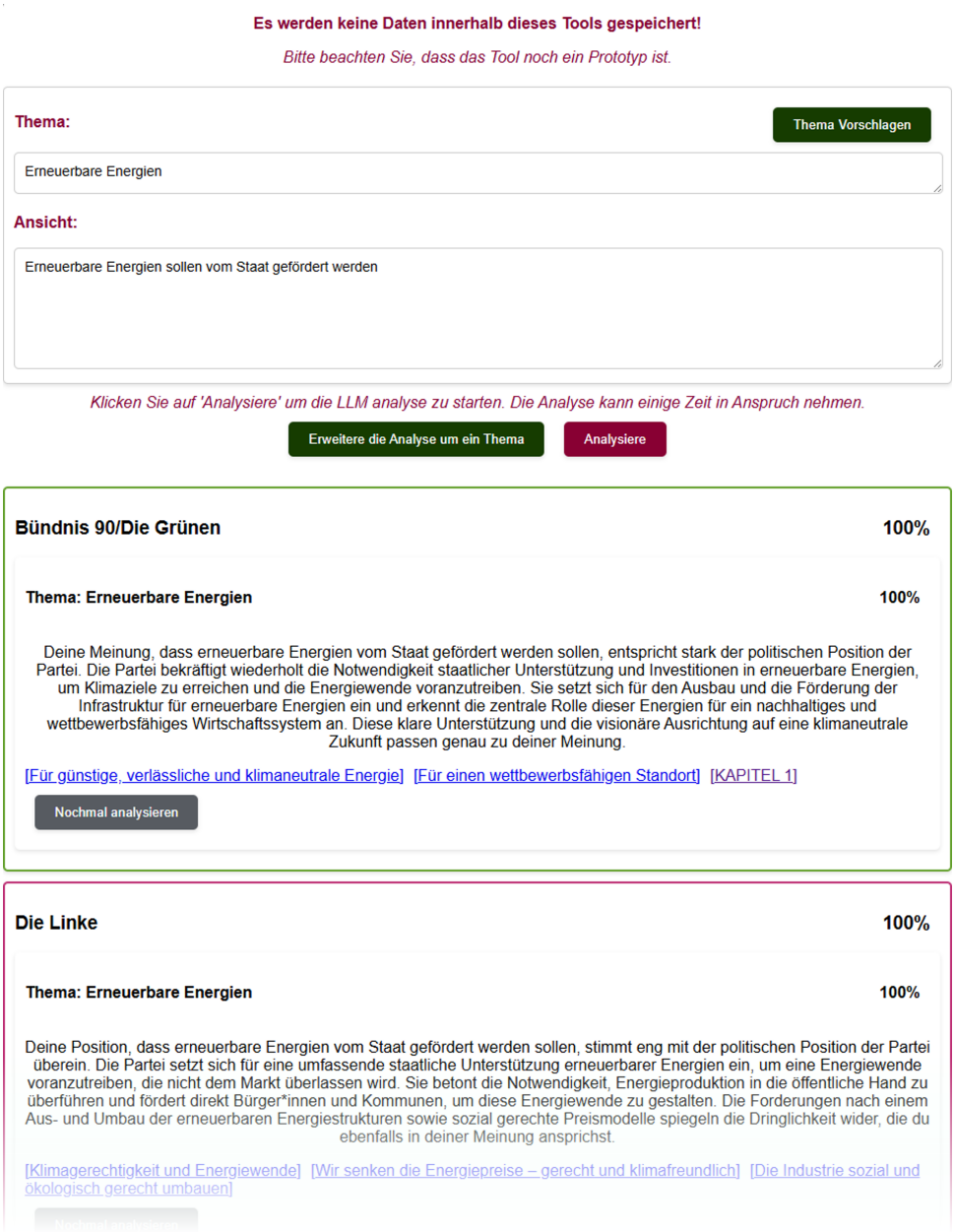

Figure 2: Example Results of *VotingAid*

5. Discussion

5.1. Challenges and Limitations

Despite its advantages, *VotingAid* faces several challenges. One key issue is the system's ability to handle nuanced political relationships, such as complex economic policies or environmental regulations that involve multiple interacting factors. Additionally, optimizing performance remains an ongoing task, as retrieving and analyzing manifesto data in real-time can be computationally intensive. Another limitation is the reliance on party manifestos, which may not fully reflect real-world policy implementation or party actions beyond their stated positions.

5.2. Planned Improvements

Several improvements are planned for future iterations of *VotingAid*. The system will incorporate an option for users to prioritize issues, allowing them to weigh specific topics more heavily in the analysis. Enhancing response readability is another priority, as better formatting and structured explanations will improve user comprehension. Expanding contextual retrieval beyond manifestos to include third-party policy analyses and fact-checking reports will further enhance the reliability of the system.

6. Conclusion and Future Work

VotingAid represents a novel approach to assisting voters in making informed decisions by leveraging Retrieval-Augmented Generation techniques. The system addresses key limitations of traditional VAAs by allowing users to define their own concerns rather than relying on predefined statements. While the current implementation demonstrates the feasibility of this approach, future work will focus on user evaluations to assess usability and effectiveness. Further development will also include an enhancing of the underlying data by integrating multiple sources beyond party manifestos to ensure clarity and accuracy in political comparisons.

Declaration on Generative AI

During the preparation of this work, the author used GPT-4o for Grammar and spelling checking. After using these tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] The Economist, Disinformation is on the rise. how does it work?, 2024. URL: https://www.economist.com/science-and-technology/2024/05/01/disinformation-is-on-the-rise-how-does-it-work, accessed: 2024-12-23.

[2] The Economist, China is flooding taiwan with disinformation, 2023. URL: https://www.economist.com/asia/2023/09/26/china-is-flooding-taiwan-with-disinformation, accessed: 2024-12-23.

[3] The Economist, France uncovers a vast russian disinformation campaign in europe, 2024. URL: https://www.economist.com/europe/2024/02/12/france-uncovers-a-vast-russian-disinformation-campaign-in-europe, accessed: 2024-12-23.

[4] The Economist, Why romania cancelled a pro-russian presidential candidate, 2024. URL: https://www.economist.com/europe/2024/12/12/why-romania-cancelled-a-pro-russian-presidential-candidate, accessed: 2024-12-23.

[5] ZDF, Desinformation: Wie der kreml gezielt falschmeldungen verbreitet, 2024. URL: https://www.zdf.de/nachrichten/politik/ausland/desinformation-sda-kreml-russland-100.html, accessed: 2024-12-23.

[6] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, D. Kiela, Retrieval-augmented generation for knowledge-intensive nlp tasks, in: Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20, Curran Associates Inc., Red Hook, NY, USA, 2020.

[7] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers), 2019, pp. 4171–4186.

[8] S. I. A. G. D. J. P. K. Yannick Gittmann, Martin Schiele, Voting advice applications: Implementation of rag-supported llms (2023). URL: https://www.researchgate.net/publication/382298218_Voting_Advice_Applications_Implementation_of_RAG-supported_LLMs.

[9] M. P. Polak, D. Morgan, Extracting accurate materials data from research papers with conversational language models and prompt engineering, Nature Communications 15 (2024) 1569. URL: https://doi.org/10.1038/s41467-024-45914-8. doi:10.1038/s41467-024-45914-8.

[10] K. Peffers, T. Tuunanen, M. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manage. Inf. Syst. 24 (2007) 45–77. URL: https://doi.org/10.2753/MIS0742-1222240302. doi:10.2753/MIS0742-1222240302.

Appendix: Source Code and Repository

The source code for *VotingAid* is available on GitHub. The main implementation can be accessed at https://github.com/NicolasMahn/VotingAid_Frontend, while the database management system is available at https://github.com/NicolasMahn/VotingAid_DB_Manager. These repositories contain the full implementation details, including setup instructions and deployment configurations, but they lack the underlying data (the manifestos).