



A few Thoughts on the Use of ChatGPT, GPT 3.5, GPT-4 and LLMs in Parliaments: Reflecting on the results of experimenting with LLMs in the parliamentary context

JÖRN VON LUCKE

The Open Government Institute - TOGI, Zeppelin Universität gemeinnützige GmbH, Friedrichshafen, Germany,
joern.vonlucke@zu.de

SANDER FRANK

The Open Government Institute - TOGI, Zeppelin Universität gemeinnützige GmbH, Friedrichshafen, Germany,
sander.frank@zu.de

Starting in November 2022 with the free provision of ChatGPT, large language models (LLM) are now publicly available. This has significantly increased the number of publications which scopes potential changes caused by the application of generative artificial intelligence (AI) in various societal domains. The private use of AI and the economic integration of generative LLMs have increased significantly. However, for parliamentarians and parliamentary professionals, the technology often remains abstract, impacting everyday work only peripherally. Due to the special responsibility of parliaments, governments, and administrations as the organizational instances of society, and through the inherent legitimations by society itself, there is a necessity to examine the implications of the use of generative LLMs within these institutions and traditional structures as well as their influence on political system logic. The paper analyzes the responses that the generative LLMs GPT 3.5 and GPT 4 have provided via ChatGPT, based on the same input command (prompt) over different times. The responses help to assess how LLMs can be used in the parliamentary context, to reflect what dangers exist as well as to respond to the question on how a business model of an AI department in parliament might look like. Furthermore, it shall be explored whether there are fluctuations in the quality of the responses and how these should be evaluated against the backdrop of the need for accurate and precise workflows in parliamentary operations. Ultimately, the paper aims to provide an answer as to whether the application of ChatGPT together with the LLMs GPT-3.5 and GPT-4 could already deliver this necessary quality and consistency for the parliamentary working environment today.

CCS CONCEPTS • Government technology policy • Artificial intelligence • Computing in government

Additional Keywords and Phrases: large language model, ChatGPT, GPT 3.5, GPT-4, parliament

1 INTRODUCTION

To appropriately assess and evaluate the disruptive changes of artificial intelligence (AI) in general, and generative large language models (LLM) in particular, on parliaments, governments and administrations, the research team, like other scholars, has been engaging with questions regarding the potentials, opportunities, risks, and challenges of these technologies [23][24][25][71][72][74]. Through collaborative efforts with scientists, administrative staff, and decision-makers from parliaments, 210 proposals for the use of AI in parliaments were generated and evaluated [74]. It became evident that ParITech [36] and the use of AI in parliaments is not just a technical issue, even if there are several (generative) AI models and AI technologies that are already suitable for use. Moreover, this phenomenon poses challenges in ethical, legal, and administrative realms, coupled with limitations in practical utilization that intersect with fundamental democratic principles of society. Particularly within the parliamentary context, emphasis must be placed on precision, accuracy, and transparency in the results of work processes. Otherwise, the equal

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2024 Copyright held by the owner/author(s).

ACM 2691-199X/2024/05-ART

<http://dx.doi.org/10.1145/3665333>

treatment of all people, as well as the legality of resolutions, official decisions, and choices for the common good, for example in discretionary decisions, could not be guaranteed without distortions.

In light of these considerations, the objective of the paper is to assess the contribution of the generative pretrained transformer model ChatGPT, in addressing the multifaceted questions surrounding the implementation of AI in parliamentary settings. To make it clear from the beginning, LLMs work according to mathematical probabilities, not according to human knowledge and experience. The team therefore approached this task with skepticism, analyzing responses to three prompts that were presented in the same wording over different times, and using via ChatGPT both the free (GPT 3.5) and the premium model (GPT 4). This approach was chosen to determine the added value of using an LLM in everyday applications. A multi-level cluster analysis with a variety of queries would certainly have offered scientific value, but it misses the research interest in the sense that selecting the best answers from a multitude of queries is not always possible in many everyday parliamentary operations. Therefore, the focus was on determining whether significant quality differences in the responses to the posed questions could be observed with just a few queries over a period of six months, and whether this would lead to considerable variability in outcomes in everyday use. Doubts about the responses arises partly from ethical considerations, asserting that legislation must be made by humans, legitimized by democratic majorities and human empathy and the reproduction of biases in the underlying training data by any LLM. It also stems from the fact that generative LLMs could be manipulated, raising ethical questions about the promotion of a technical solutionism [4].

Furthermore, due to the relatively pronounced opacity of the technical processes and specific parameters used in AI [55], it remains unclear how comparable or precise the generated content might be. Nonetheless, the team found it fascinating to ask ChatGPT challenging and ethically charged questions and compare the responses analytically. After a literature review (chapter 2), the research questions (chapter 3) were designed. The team was particularly interested in whether the generative LLM addresses these critical questions and whether any relevant considerations are evident in the responses, or if the answers provide relatively unreflective promises about the great potentials and possibilities. By querying the creation of a business canvas for an AI department in parliaments, the team aimed to analyze indirectly whether critical reflection has been incorporated into the results. The negative inquiry about the dangers also appeared as a valid method to understand which dangers the system recognizes and names. The results were achieved through the reflection on the answers generated (chapter 4), a discussion with a SWOT analysis and a brief impact assessment (chapter 5) and a summary of the main findings in the conclusions (chapter 6).

2 LITERATURE REVIEW

The term “artificial intelligence” was introduced in 1956 by mathematicians and computer scientists. They assumed in their study that it was possible to develop a machine that could simulate “every aspect of learning or any other feature of intelligence” [45]. Nowadays, there is neither a single technology nor a collection of niche applications. Rather, numerous technologies are assigned to AI today [13][66]. After the highly regarded paper [68] from 2017, which dealt with sequence transduction models and proposed with the “transformer” a “new simple network architecture”, scientific discussions on the topic of generative artificial intelligence have received considerable attention. Following the release and open-access

availability of ChatGPT, built upon the LLM GPT-3.5 architecture, in November 2022 by OpenAI, there has been a heightened focus on researching the potential applications, deployment strategies, as well as the associated risks and challenges inherent in AI [2]. Popular media, newspapers and magazines [14], as well as government commissions, governments [7][18][21][32][75] and expert reviewers [57], have increasingly addressed these topics, which substantiates the presumption of having reached another peak of the hype surrounding AI according to Gartner [28]. As a cross-sectional technology used in many different societal and economic areas and sectors, AI also touches upon many different legal areas [60]. So far, there is no comprehensive legal framework for the use and application of AI and its products, as well as for the generated content. On December 9, 2023, the EU achieved an agreement between the European Parliament and the European Council on the planned AI legislation. After the approval by the European Parliament in March 2024 and the European Council, it will be directly applicable to all EU member states and is effective 24 months after entry into force [19][20]. The draft contains a series of regulations and requirements for the use of AI, including for the public sector. But there is currently no other international treaty that regulates the numerous issues related to AI on an international level [60]. However, some governments and supranational unions are working on corresponding comprehensive frameworks. There have are numerous soft law initiatives, such as recommendations, strategy papers, declarations of intent, or guidelines [18][10], which address the ethical development and principles of the use of AI or critique the guidelines themselves [31][33][46][49]. The renowned conference in the field of AI, "Conference on Fairness, Accountability, and Transparency" by the "Association for Computing Machinery", has seen a significant increase and a tenfold rise in accepted conference contributions from 2018 to 2022, which deal with ethical and related issues in the context of AI [43]. The question of consciousness in current AI models has also been discussed in this context [52]. Ethical deliberations range from fundamental questions about the technical and societal design of ethical AI ecosystems [64] to the ethical challenges and concerns related to AI usage [1][48][57][63][65], or challenges and perpetuation of discrimination, racism, and sexism due to AI use and corresponding biases in training data. The studies address general societal challenges [8] as well as issues in medicine, government actions, legislation, education, work and development, and human rights [9][15][27][30][39][40] [59][61][69][77]. Various research projects have examined the influence of AI on governments, administrations, and parliamentary systems, with the aim of identifying potential dangers, showing opportunities to reduce bureaucracy, the acceleration of processes, and the enhancement of service orientation towards citizens [3][37][67][74]. Some of these investigations have employed a practice-oriented approach, directing the research towards the parliaments and their employees themselves, utilizing a scientific methodology [24][54][70]. But the security risks in the governmental context have also already been addressed. In 2023, several national security agencies and intelligence services developed an international guideline for secure AI systems development [47]. Issues related to synthetic media and manipulation have also been the subject of governmental and parliamentary deliberations in the context of AI [11]. Furthermore, the State Commissioner for Data Protection and Freedom of Information Baden-Württemberg has also dealt with issues regarding data protection. His team compiled legal foundations for data protection in the context of the use of AI [35]. AI is therefore no niche topic in the governmental context anymore.

Other studies have a more technical focus and attempt to investigate what happens inside the black box of AI. For this purpose, generative models are often used, and the results are examined, for example to find

a possible explanation for the question of how sources and selection processes of answers from generative models can be identified and described [26]. The question of the determinants of the quality of outcomes in generative AI also plays an important role in this context [62][76]. Other studies pursue more of a socio-technical context, illuminating, for example, the creative capabilities of generative models [16][44] and their impact on the economy and the creative industry [22].

In the context of research on generative AI, large language models (LLM), the software ChatGPT and the underlying models GPT-X from the company OpenAI play a particularly significant role and are often the subject of studies. These currently delve into topics such as the effects on scientific education and research [12][56], influence and research prospects in human resource management [6], impacts on journalism and the media landscape [53], implications for management theories and approaches [38], societal and economic development [34][42], changes in tourism and hospitality [17], and policy implications [29][34]. A survey was conducted in spring 2023 to determine whether generative AI would lead to significant changes to the 210 proposals developed for the use of AI in parliaments [73]. In March 2024, the first manual how to deal with generative AI in Congress was published [58]. [5] view the generative LLM GPT-4 as a new generation of LLMs that possess greater general intelligence than previous AI models. They argue that this new generation can tackle complex tasks and should be understood as an incomplete but early version of general artificial intelligence (AGI) [5].

In summary, it can be stated that research on AI has been conducted for several years, but especially due to the model proposals by [68] and the availability of the GPT-3.5 and GPT-4 LLM, it has experienced a real hype. The research topics are quite broad and diverse, as AI affects many societal and scientific areas. Research on ethical questions and challenges, driven by political expectations and societal fears, is currently often the focus of many scientific discussions. These are of particular importance in the context of the present work, as it focuses on parliamentary issues that concern sensitive areas. Due to the wide range of scientific research topics related to AI and the predominantly opaque mechanisms of the black box AI, universally valid and generalizing statements and findings are difficult to make.

3 RESEARCH APPROACH

The aim of the first experiments of the team was to attempt in practice how generative AI can be used as a constructive tool in a parliament and in the context of parliamentary issues. Additionally, the team wanted to use appropriate questions to determine the quality of the generated answers and thus of the LLM. The starting point should be challenging prompts that also include legal and ethical dimensions, i.e. that do not provide simple answers and might discredit the LLM itself. The team wanted in-depth answers with an ethical standard, which, however, could still be easily compared with each other due to the probabilistic functioning of the LLM. The aim was to gain an impression of the generated content, its quality, and its limitations, for further practical use. This leads to the research question RQ1: "To what extent can GPT-X be used in the parliamentary context, taking into account ethical and legal constraints?" The prompt: "Propose 10 fields of application for ChatGPT in a parliament" seems to be the most appropriate for this purpose because it is suited for everyday use.

The answers to the prompt were also intended to outline areas of application for AI in parliaments and show which recommendations for action could already be generated for a parliament based on existing

LLMs and their underlying knowledge shaped by the ideas and contributions of the training data and texts. Simple, easy-to-understand suggestions were expected. This leads to research question RQ2: “How can we use GPT-X to obtain direct or indirect insights into how parliamentary proceedings can be improved and optimized through AI applications?”

The texts are generated freshly each time, based on the similar prompt and on the given statistical probabilities of the system itself. Generating the same texts again would have to be a major coincidence, but it could happen, especially when the LLM has limitations, a small vocabulary and can only generate a manageable number of responses. However, when carrying out these initial tests, it was not yet clear that the LLMs used could improve significantly over the course of a few months. The focus of the used LLM is obviously on a broad and general application. Looking at the results from RQ1+RQ2, there are certainly evaluative statements from which quality assessments of the system can be derived. These statements should therefore be examined more closely with regard to their content and validity. This leads to research question RQ3: “How valuable are the responses and assessments of GPT-X regarding the opportunities and concerns related to the use of AI in the day-to-day operations of parliaments?”

Considering the use of such LLMs in a national parliament, it would be surprising and inappropriate if the quality of the answers depended on the “daily form” of the language model. Continuous improvement of the language models would be desirable, but a constant oscillation between bad, mediocre and good would be unacceptable, especially in a parliamentary setting. The question must therefore be asked whether the quality is improving, whether it varies within a reasonable variance or whether it depends on chance and the associated “daily form”. This leads to research question RQ4: What differences, redundancies, and analysis categories arise when comparing responses generated by GPT-X with the same prompt at different times?

To answer the four research questions, an experimental setting consisting of three prompts was set up. A brainstorming session on possible applications was used to experimentally find out which suggestions are generated. These impulses then need to be analyzed. With a provocative prompt aimed at the system’s own weaknesses, the handling of negative and damning arguments should be analyzed, too. Additionally, this approach aimed to determine whether the model, when posed with a negative question, addresses the advantages, and tries to downplay the risks and concerns with less convincing arguments. Thirdly, by applying Osterwalder’s business canvas model [50][51] to an AI department of a parliament, the breadth and depth of argumentation that a LLM already possesses in the context of AI in parliaments will be demonstrated.

In this context, value must also be placed on the ethical reasoning of the systems. It is important to check whether ethical issues only play a role when asked directly. Or are they also considered in a positively posed question about the design of an AI department, which offers at second glance numerous challenging trapdoors? Practical considerations were taken into account, recognizing that various issues that touch on ethical issues may arise in day-to-day operations. The use in daily operations would be excluded if ethical considerations are not fundamentally included, but only when specifically asked for. This approach is intended to evaluate content, argumentation, consistency, breadth, depth, transparency and ethical

reflection. With these requirements in mind, the team decided to have the responses analyzed and evaluated by a scientist with specialist knowledge. In any event of doubts regarding the responses, an ethics officer and a gender officer might have been included. Ensuring sensitivity to ethical concerns was therefore a given during the development of these findings and is also generally suggested for the use of AI, for example through the consolidation of experts. The selection of the right prompt is crucial for obtaining answers to these questions. The prompt command can be self-developed, developed in the team or with the help of another prompt ("Please develop 10 prompts that best highlight the broad application possibilities of AI in parliaments"). Once the most suitable prompt is available, it can generate novel artefacts with any text-generating AI. For further analysis, the text contributions were saved in the form of electronic documents, processed and subsequently analyzed, from which findings can then be derived for research questions RQ1, RQ2, RQ3 and RQ4.

The starting point for all considerations was a brainstorming session conducted in German in January 2023 (ChatGPT 3.5) on possible areas of application for ChatGPT in parliaments: "Schlage 10 Einsatzfelder für ChatGPT in einem Parlament vor" (Propose 10 fields of application for ChatGPT in a parliament). When the same prompt was repeated in English five months and a few versions later, much higher-quality results were generated. Based on this quite surprising realization, a further study with three prompts (brainstorming, provocation, business canvas) was designed in English. Furthermore, this observation has led to the impression that the quality of the responses at this time is strongly dependent on the language used, resulting in significant differences due to the choice of language. On three dates, in June 2023 (GPT 3.5), in September 2023 (GPT 4.0) and in December 2023 (GPT 4.0), these three prompts were deliberately carried out experimentally with different ChatGPT versions. The prompts (Table 1) were intentionally posed in English, as the current LLMs including GPT 3.5 and GPT 4.0 have a particularly large foundation in this language [41].

Table 1: Prompting Overview Regarding the Experiments

Date	LLM (via ChatGPT)	Prompt
January 2023	GPT 3.5 – Dec 2022 Version	1 [DE]
June 2023	GPT 3.5 – May 24 2023 Version	1, 2, 3 [EN]
September 2023	GPT 4.0 – Aug 03 2023 Version	1, 2, 3 [EN]
December 2023	GPT 4.0 – Dec 2023 Version	1, 2, 3 [EN]

4 RESULTS FOR THE RESEARCH QUESTIONS

4.1 Results for Research Question RQ1

The analysis of the brainstorming results from January, June, September, and December 2023 shows that they are of varying quality and that the underlying transformers have set different priorities with regard to the content and order of the answers. When asking the same prompts a few months later, the team was surprised about the improvements regarding response behavior in the meantime. It was not expected that addressing questions of parliamentary use were at the forefront of the model's development. There may be good reasons why the respective LLM came up with these answers at that particular time. For external experts, these cannot be understood through the answers. Compiling the results in the form of simple tables initially leaves a well-considered, wide-ranging impression of the proposals and arguments.

However, it is only through a comparative analysis possible to determine which arguments were used and which were not addressed in the responses. The divergence in generated answers (Table 2) highlights the variations that have emerged.

At a first glance, the headlines generated by GPT-3.5 and GPT-4 with the same prompt seem to fit. Some of the prompt results are similar, sometimes using synonymous words with slight variations. In terms of content, there are clear and correct answers to the simple question, so there are no outliers, incorrect results, or hallucinations. The comparative analysis of the three answers has shown that the generated content with headings and in-depth explanations differ from each other and open a much broader range of answers and arguments. The response behavior stands and falls with the quality of the underlying LLM and the training data used. Surprisingly, one identical answer was received once in the pre-training (in December 2023, 3 months apart). This means that the possibility of a system generating exactly the same answer twice or multiple times cannot be ruled out. If the LLM has a limited linguistic treasure, repetitions may occur. For use in parliaments, text-generating AI must attach importance to the correctness and accuracy of the answers regarding the questions asked.

Table 2: Results of the First Prompt

Translated English Version of the Original German Version 10.01.2023 (ChatGPT Dec 22 Version 2022), ChatGPT-Translation: 23.06.2023	
Prompt: Propose 10 fields of application for ChatGPT in a parliament	
1. Creation of draft laws:	ChatGPT can be used to automatically create draft bills based on existing laws and regulations and the requirements and needs of Parliament.
2. Taking minutes of plenary sessions:	ChatGPT can be used to automatically create minutes of plenary sessions based on audio recordings and transmissions of the sessions.
3. Translation of documents:	ChatGPT can be used to automatically translate documents that are in another language into the desired
4. Generation of responses to requests:	ChatGPT can be used to automatically generate responses to requests from citizens, journalists and other stakeholders.
5. Generation of presentations:	ChatGPT can be used to automatically generate presentations for plenary sessions, committee meetings and other occasions.
6. Generation of analyses and reports:	ChatGPT can be used to automatically generate analyses and reports on various policy issues.
7. Creation of speeches and addresses:	ChatGPT can be used to automatically generate speeches and addresses for various occasions and occasions.
8. Creation of letters and emails:	ChatGPT can be used to automatically create letters and emails for various purposes, such as responses to enquiries or invitations to events.
9. Generation of Q&A sessions:	ChatGPT can be used to automatically generate questions and answers for Q&A sessions with MPs or other representatives of Parliament.
10. Generation of session agendas and agendas:	ChatGPT can be used to automatically generate agendas and agendas for meetings of Parliament and its committees.
Original English Version 23.06.2023 (ChatGPT May 24 Version 2023), T: 23.06.2023	
Prompt: Propose 10 fields of application for ChatGPT in a parliament	
1. Legislative research and analysis:	ChatGPT can assist parliamentarians by providing real-time access to legislative information, analyzing bills, and summarizing key points for better decision-making.
2. Constituent engagement:	ChatGPT can engage with constituents through various digital platforms, answering their questions, addressing concerns, and providing information about parliamentary proceedings.

Original English Version 23.06.2023

(ChatGPT May 24 Version 2023), T: 23.06.2023

3. **Speechwriting assistance:** ChatGPT can help parliamentarians in crafting speeches and statements by suggesting content, providing background information, and offering language refinement.
 4. **Policy drafting and analysis:** ChatGPT can aid in drafting policy proposals by offering insights, conducting research on similar policies, and analyzing potential impacts.
 5. **Committee support:** ChatGPT can assist in committee work by organizing and summarizing submissions, analyzing data, and generating reports based on discussions and hearings.
 6. **Multilingual communication:** ChatGPT can facilitate communication with non-native speakers or constituents who speak different languages by providing real-time translation and interpretation services.
 7. **Fact-checking and misinformation detection:** ChatGPT can help identify false or misleading information in parliamentary debates, speeches, and documents, thereby promoting evidence-based decision-making.
 8. **Procedural guidance:** ChatGPT can offer guidance on parliamentary rules and procedures, ensuring that parliamentarians follow proper protocols during debates, voting, and other proceedings.
 9. **Public opinion analysis:** ChatGPT can analyze social media trends, public sentiment, and online discussions to provide insights into public opinion on various legislative issues.
 10. **Training and knowledge dissemination:** ChatGPT can serve as a virtual tutor, providing training modules, educational resources, and interactive sessions to parliamentarians and staff on parliamentary procedures, public speaking, and policy topics.
-

Original English Version 18.09.2023

(GPT4 Premium Aug 3 Version 2023), T: 18.09.2023

Prompt: Propose 10 fields of application for ChatGPT in a parliament

1. **Document Summarization:** ChatGPT could be used to summarize lengthy reports, bills, or legislative documents, making it easier for parliamentarians to quickly grasp key points.
 2. **Research Assistance:** Members of parliament often need to cite statistics, historical events, or international precedents during debates. ChatGPT can quickly provide this information, helping MPs to back up their arguments with factual data.
 3. **Translation and Interpretation:** In parliaments where multiple languages are spoken, ChatGPT can assist in real-time translation, allowing smoother communication between members who speak different languages.
 4. **Public Engagement:** ChatGPT can be employed in interactive platforms where constituents can ask questions about legislative matters, providing instant answers and helping to increase public understanding of parliamentary proceedings.
 5. **Minute-taking and Transcription:** With the capability to understand context and relevance, the AI could assist in minute-taking, capturing the essence of discussions, and providing detailed transcripts of parliamentary sessions.
 6. **Predictive Analysis:** While it can't predict the future, ChatGPT can provide analysis based on historical data. This might help in forecasting the implications of certain policies or in understanding the past effects of similar legislation in other jurisdictions.
 7. **Training and Onboarding:** New members of parliament or staff can be introduced to parliamentary procedures, ethics, and other relevant topics using ChatGPT as an interactive training tool.
 8. **Speech Preparation:** MPs often give speeches on a variety of topics. ChatGPT can assist in drafting speeches by providing data, historical references, or even suggesting rhetoric techniques to make the speech more compelling.
 9. **Scheduling and Administration:** An AI assistant can handle scheduling of meetings, reminders for deadlines related to legislative processes, and coordination of various administrative tasks within the parliament.
 10. **Accessibility:** For those with visual or auditory impairments, ChatGPT can be adapted to read out documents, explain visual content in accessible terms, or transcribe spoken words into text.
-

Original English Version 08.12.2023

(GPT4 Premium Dec. Version 2023), T: 08.12.2023

Prompt: Propose 10 fields of application for ChatGPT in a parliament

1. **Legislative Research Assistant:** ChatGPT can assist in researching and summarizing relevant information on proposed bills, historical legislation, and comparative legal studies from different jurisdictions.
 2. **Drafting Support for Bills and Amendments:** It can help in drafting initial versions of bills, amendments, and legislative
-

- proposals by providing language templates and ensuring legal and procedural consistency.
3. **Constituent Communications Management:** ChatGPT can assist parliamentarians in managing and responding to constituent communications, ensuring timely and informative responses to public queries and concerns.
 4. **Meeting and Debate Transcription:** It can transcribe parliamentary debates, committee meetings, and public hearings, making the transcripts available for record-keeping, analysis, and public dissemination.
 5. **Language Translation and Interpretation:** ChatGPT can provide real-time translation and interpretation services, aiding communication in multilingual parliamentary settings.
 6. **Policy Analysis and Briefing:** It can analyze policy documents, international treaties, and governmental reports, summarizing key points for briefings and decision-making.
 7. **Training and Education for Parliament Staff:** ChatGPT can be used for training new staff members, providing information on parliamentary procedures, ethics, and legislative processes.
 8. **Public Engagement and Education:** It can aid in creating educational content about the parliamentary process and current legislation for public awareness campaigns and school programs.
 9. **Issue Tracking and Monitoring:** ChatGPT can monitor ongoing issues, media coverage, and public opinion trends, providing parliamentarians with up-to-date information relevant to their legislative and constituency work.
 10. **Speech Writing Assistance:** It can assist in drafting speeches and talking points for parliamentarians, tailored to various occasions and audiences.
-

The divergent responses might be due to the inadequacy of limiting to only 10 suggestions when addressing a complex question, as this constrains the range of possible correct answers and fails to encompass the full breadth of topics. In the case of a very complex range of answers, only those areas that are considered particularly important and relevant by the transformer model are selected. These assessments can change constantly as LLMs may be improved by the providers and grow dynamically while learning from newly fed texts. A direct explanation for the different order and the potentially associated weighting cannot be provided. One possible reason could be that random decisions were also made here, and therefore, the order might be random as well. Furthermore, the emphasis of the underlying training data could also provide a basis for explanation.

4.2 Results for Research Question RQ2

To answer the second research question RQ2, a convincing and simple question is needed to see how the systems deal with negative argumentation that would damage themselves, their reputation and their application possibilities. The prompt: "Why should I not use ChatGPT in parliaments?" seems to fit best regarding other proposals. Care was taken to have a concise prompt and to ensure consistency with the research questions of this article. This selected prompt raises the question whether these LLMs show self-censorship by only presenting negative arguments that are not harmful or use harmless formulations to avoid any self-damage. This could be programmed in the LLM for the final quality check. Are LLMs capable of honestly presenting damaging arguments? If so, to what extent users get the impression that LLMs might be able to reflect, criticize their own data basis, and classify and correct any data leaks (gender bias, ethnic bias, disability bias)? How useful are LLMs and their answers in the daily parliamentary routine, when reality is not just about pleasantries, but about hard-hitting arguments about power and political decisions? If GPT-X systems do not give negative, unfavorable, unpleasant, or self-defeating answers, would this be in the system's self-interest? On the other side, an LLM is not a thinking system, but a system that suggests the next word on the foundation of probabilities. However, in view of the target groups and framework conditions in the parliamentary arena (disputes, delays, facts, truth, manipulations, reports, and poor

data bases), this would only be of limited help. With a view on the weaknesses of text-generating AI (see the SWOT analysis in chapter 5), the results of the second prompt (Table 2) will show which of these concerns are really relevant.

Table 3: Results of the Second Prompt

Original English Version 23.06.2023 (ChatGPT May 24 Version 2023), T: 23.06.2023	
Prompt: Why should I not use ChatGPT in parliaments?	
1. Lack of human judgment: ChatGPT operates based on patterns and data it has been trained on, without the ability to exercise critical thinking or make subjective judgments. This can lead to limitations in handling complex or nuanced situations that require human intuition and contextual understanding.	
2. Bias and misinformation: As an AI language model, ChatGPT learns from the data it is trained on, which can inadvertently include biases present in the training data. If not carefully monitored and controlled, ChatGPT may unintentionally propagate or amplify biased information, reinforcing existing prejudices or misinformation.	
3. Legal and ethical implications: Using AI technology like ChatGPT in parliamentary settings raises legal and ethical concerns. Privacy of personal data, accountability for decisions made with AI assistance, and potential misuse of AI capabilities are among the issues that need to be addressed through robust regulations and guidelines.	
4. Lack of empathy and emotional intelligence: ChatGPT lacks the ability to understand emotions, empathy, and social dynamics in the same way humans do. This can limit its effectiveness in situations that require emotional support, sensitive communication, or conflict resolution.	
5. Dependence on technology: Overreliance on AI systems like ChatGPT may lead to a reduction in human skills, critical thinking, and engagement in parliamentary processes. It is essential to strike a balance between leveraging AI for efficiency while ensuring the active participation and decision-making abilities of elected representatives.	
6. Transparency and explainability: AI systems like ChatGPT often operate as black boxes, making it challenging to understand the reasoning behind their responses or decisions. This lack of transparency can undermine trust and accountability in parliamentary proceedings.	
7. Technical limitations: ChatGPT, like any AI system, has its limitations. It may struggle with understanding ambiguous queries, providing accurate information on rapidly evolving events, or handling highly specialized or domain-specific topics.	
Original English Version 18.09.2023 (GPT4 Premium Aug 3 Version 2023), T: 18.09.2023	
Prompt: Why should I not use ChatGPT in parliaments?	
1. Over-reliance on AI: Parliamentarians could become overly reliant on the AI for information and guidance. This could lead to a decline in the emphasis on human expertise and experience, which are essential for nuanced understanding and decision-making.	
2. Bias and Accuracy Concerns: AI models can inadvertently exhibit biases present in their training data. Misinformation or skewed perspectives could influence legislative processes and discussions if unchecked.	
3. Data Security and Privacy: Given that ChatGPT can log queries for operational purposes, there might be concerns about sensitive parliamentary information being exposed or misused.	
4. Loss of Human Touch: The human element in public service, including empathy, subjective judgment, and understanding of local contexts, is irreplaceable. There's a risk that an over-dependence on AI could diminish these valuable qualities in legislative proceedings.	
5. Authenticity and Accountability: There might be concerns about whether an AI's response genuinely reflects the views or intentions of a parliamentarian or body. This can lead to accountability issues if the AI provides a contentious or incorrect stance.	
6. Erosion of Skills: If parliamentarians or parliamentary staff become overly reliant on AI for tasks like research or speech writing, their own skills in these areas might atrophy over time.	
7. Transparency Concerns: AI's decision-making process is often seen as a "black box." The lack of transparency in how AI reaches its conclusions could be a concern in settings where decision-making processes should be open to scrutiny.	
8. Manipulation and Misuse: There's potential for misuse if individuals learn how to manipulate the AI's responses to suit	

Original English Version 18.09.2023

(GPT4 Premium Aug 3 Version 2023), T: 18.09.2023

particular agendas, leading to misinformation.

9. **Economic and Employment Impact:** Relying on AI for tasks traditionally done by humans could lead to job displacement in administrative roles.

10. **Technological Limitations:** ChatGPT, while advanced, does not understand context in the same way humans do. It might provide answers that are technically correct but not appropriate or nuanced enough for complex legislative issues.

Original English Version 08.12.2023

(GPT4 Premium Dec. Version 2023), T: 08.12.2023

Prompt: Why should I not use ChatGPT in parliaments?

1. **Accuracy and Reliability:** AI models like ChatGPT may not always provide completely accurate or up-to-date information. Relying on them for critical decisions, especially without human verification, could lead to misinformation or errors.
 2. **Bias and Ethical Concerns:** AI systems can inadvertently perpetuate biases present in their training data. This could lead to biased policy suggestions or communications, which is particularly concerning in a parliamentary context where fairness and impartiality are paramount.
 3. **Confidentiality and Security Risks:** Using AI in handling sensitive or confidential information poses security risks. ChatGPT and similar tools might be vulnerable to data breaches or misuse, potentially compromising confidential parliamentary communications.
 4. **Dependence on Technology:** Over-reliance on AI for tasks like research, drafting, and communication could reduce the development of these skills among parliamentary staff and members, making them overly dependent on technology.
 5. **Lack of Contextual Understanding:** AI models may not fully grasp the cultural, social, and political nuances important in parliamentary work. Misunderstanding context can lead to inappropriate or irrelevant suggestions.
 6. **Accountability and Transparency:** Decisions influenced by AI might lack transparency, making it difficult to understand how certain conclusions were reached. This could raise accountability issues, especially in a political context where decision-making processes should be clear and justifiable.
 7. **Reduced Human Interaction:** Relying on AI for tasks like constituent communication could reduce personal interaction, which is vital for understanding public opinion and maintaining public trust.
 8. **Legal and Ethical Regulations:** The use of AI in government processes may be subject to strict legal and ethical regulations. Ensuring compliance and adapting these frameworks to keep pace with technological advancements can be challenging.
 9. **Potential for Misuse:** There is a risk that AI tools could be misused for political manipulation, such as spreading propaganda or manipulating public opinion, which could undermine democratic processes.
 10. **Impact on Employment:** Implementing AI solutions in parliamentary settings might lead to concerns about job displacement or the devaluing of human expertise.
-

Only seven answers were given in the first query. This suggests, to some extent, that the LLM may not have data of all concerns at that time and did not need to make a selection as it did in the responses from the first prompt. In the subsequent queries, the LLM provided 10 responses. This could be a result of another LLM, more training (processing more texts), of publicly debated risks associated with generative AI models, and of suggestions that improvements may have been made in terms of ethical considerations. The order of the responses changes when comparing the first query to the two following queries. However, the order of the responses in the second and third queries remains largely consistent until the fifth response. From the fifth response onwards, different orderings can be observed even when comparing query 2 and 3. Nevertheless, the responses are overall very similar in content, with only varying weightings and phrasings in the headings. The answers fit to the prompt and can be assessed as fair. There is no obvious evidence of self-censorship or hallucinations. How any noticeable anomalies would look like in a parliamentary context, for example triggered by a bias in the basic data set or in the algorithms, ironically intended statements or deliberate manipulation, cannot be answered in general terms. The suppression of certain answers or the

selection of harmless suggestions would be approaches. Regarding the self-interests of the originators and the effects to be achieved in parliament, manipulation should ideally not be noticed, but should contribute to forming an opinion. In a parliamentary context and for identifying these incidents, this suggests that every generated answer should be reviewed and reflected with the necessary expertise in general.

4.3 Results for Research Question RQ3

The business model canvas is an established business management tool, an analysis method and a strategic management template that Alexander Osterwalder designed and introduced in 2004 and 2005 for developing and improving business models [50][51]. It could be adapted to specific business scenarios and applications.

Good LLMs are convincing in practice because they master proven management models and business templates, but also apply them well in different contexts, so that this leads to appropriate results. The convincing advantages of a well-trained LLM are the high quality of the proposals and the speed with which results can be generated for a business canvas, in minutes rather than weeks, combined with the rich linguistic potential of an LLM. The quality can be demonstrated with a challenging task. For this reason, it should be tested how well the LLMs of GPT-3.5 and GPT-4 react in the application context for a new to be established AI department in a parliament and which constructive suggestions it generates. This is a non-trivial challenge for which only a few models or templates exist.

Three business model canvas for an AI department of a parliament (Table 4) were generated in June 2023, September 2023 and December 2023. The speed at which LLMs were able to generate comprehensive business canvases was remarkable. ChatGPT and GPT-4, both can handle the concept of a business canvas and apply it to various fields of application, also in a parliamentary context for the establishment of a new AI department in a parliament. Osterwalder's classic business canvas has nine sections that need to be filled in [50][51]. The two LLMs choose ten categories for their answers. In the first attempt (GPT-3.5 in June 2023), there was an additional section "Key Metrics", in the later attempts with GPT-4, the section "Potential Challenges" or "Potential Challenges and Risks" were added.

Table 4: Results of the Third Prompt

Original English Version 23.06.2023 (ChatGPT May 24 Version 2023), T: 23.06.2023
Prompt: Please create a business canvas for an AI department in a parliament
1. Key Partners: <ul style="list-style-type: none"> • Parliament members and committees • IT department of the parliament • External AI experts and consultants • AI solution providers and vendors • Research Institutions and universities
2. Key Activities: <ul style="list-style-type: none"> • Developing and maintaining AI applications for parliamentary use • Collaborating with parliament members to understand their needs and requirements • Conducting research on AI technologies and their potential applications in parliamentary processes • Providing training and support for parliament members and staff on AI tools and systems • Ensuring data privacy and security in AI applications

- Monitoring and evaluating the performance and impact of AI initiatives

3. Key Resources:

- Skilled AI professionals (data scientists, machine learning engineers, natural language processing experts, etc.)
- High-performance computing infrastructure
- Diverse and reliable data sources related to parliamentary activities
- Partnerships with external AI experts and solution providers
- Budgetary allocation for AI department operations

4. Value Proposition:

- Enhanced efficiency and productivity in parliamentary processes through AI automation
- Improved decision-making through data-driven insights and analysis
- Real-time access to legislative information and research
- Enhanced constituent engagement and communication through AI-powered tools
- Increased transparency and accountability through AI-based monitoring and analysis

5. Customer Segment:

- Parliament members and committees
- Parliamentary staff and administrators
- Constituents and the public

6. Channels:

- Internal communication channels within the parliament (intranet, emails, newsletters, etc.)
- Training sessions and workshops for parliament members and staff
- Online platforms and portals for access to AI tools and information
- Collaborative platforms for engagement with external partners and stakeholders

7. Customer Relationships:

- Providing ongoing support and assistance to parliament members and staff
- Collaborative approach to understanding and addressing their needs
- Regular feedback collection and evaluation of AI initiatives
- Ensuring data privacy and security in AI applications

8. Revenue Streams:

- Government funding and budget allocation for the AI department
- Partnerships and collaborations with external organizations for research and development projects
- Potential revenue generation through commercialization of AI solutions developed for parliamentary use

9. Cost Structure:

- Salaries and benefits for AI department personnel
- Infrastructure and technology costs (computing resources, software licenses, etc.)
- Training and development programs
- Research and development expenses
- Operational and administrative costs

10. Key Metrics:

- Adoption rate of AI tools and systems by parliament members and committees
 - Efficiency gains in parliamentary processes
 - Constituent satisfaction and engagement levels
 - Cost savings or cost-effectiveness achieved through AI implementation
 - Impact assessment of AI initiatives on decision-making and policy outcomes
-

Prompt: Please create a business canvas for an AI department in a parliament

1. Customer Segment:

- MPs (Members of Parliament) and their staff.
- Parliamentary committees.
- Administrative staff of the parliament.
- The general public and constituents.

2. Value Propositions:

- Faster and more efficient analysis of legislative data.
- Predictive modeling for policy outcomes.
- Automated summaries of long documents and legal texts.
- Real-time translation services for multilingual parliaments.
- Improved public engagement through AI-driven tools.

3. Channels:

- Internal parliament intranet and software systems.
- Public-facing parliament website.
- AI-driven communication tools for MPs.
- Training sessions and workshops.

4. Customer Relationship:

- Dedicated support for MPs and committees.
- Regular training sessions on AI tools.
- Feedback loops for continuous improvement.
- Transparent data handling and privacy assurance.

5. Revenue Stream:

- (This might be different since most parliamentary departments are publicly funded) Allocation from government budgets.
- Potential external partnerships or grants for specific AI research projects.

6. Key Resources:

- AI experts and data scientists.
- Large datasets from parliamentary records.
- Advanced computational infrastructure.
- Proprietary and open-source AI software.

7. Key Activities:

- Developing and deploying AI models.
- Data cleaning and preprocessing.
- Continuous research on AI's role in policy-making.
- Training and capacity building for parliamentary staff.

8. Key Partnerships:

- Technology companies and AI research institutions.
- Universities and academic researchers in governance and AI.
- International parliaments for best practice sharing.
- NGOs focused on transparency, governance, and civic tech.

9. Cost Structure:

- Salaries for AI experts and support staff.
- Infrastructure maintenance and cloud computing costs.
- Research and development.
- Training and workshops.
- Licensing for certain software tools.

10. Potential Challenges:

- Data privacy and security concerns.
-

Original English Version 18.09.2023

(ChatGPT4 Premium Aug. 3 Version 2023), T: 18.09.2023

- Resistance to technology adoption among some MPs or staff.
 - Ethical considerations in automating decision-making.
 - Keeping up with the rapidly evolving field of AI.
-

Original English Version 08.12.2023

(ChatGPT4 Premium Dec. Version 2023), T: 08.12.2023

Prompt: Please create a business canvas for an AI department in a parliament

1. Value Propositions:

- Streamline legislative processes.
- Enhance decision-making with data-driven insights.
- Improve communication and engagement with constituents.
- Increase efficiency in policy research and drafting.

2. Customer Segment:

- Parliamentarians and their staff.
- Parliamentary committees.
- Constituents seeking engagement with their representatives.
- Parliamentary administrative and support staff.

3. Channels:

- Internal communication networks within the parliament.
- Public-facing websites and social media platforms for constituent engagement.
- Training workshops and seminars for staff and parliamentarians.
- Integration with existing parliamentary IT systems.

4. Customer Relationship:

- Continuous support and training for parliament members and staff.
- Feedback loops to understand and address the needs of users.
- Transparency in AI operations and decision-making processes.

5. Revenue Stream:

- Government funding allocated for parliamentary operations.
- Possible grants or funding for innovative technology use in government.

6. Key Resources:

- AI technology and software.
- Data storage and processing infrastructure.
- Skilled personnel in AI, data analysis, and legislative processes.
- Access to legislative and public policy databases.

7. Key Activities:

- Developing and maintaining AI tools.
- Training staff and parliament members to use AI systems.
- Conducting research and analysis on legislative matters.
- Ensuring data security and compliance with legal standards.

8. Key Partnerships:

- Technology providers and AI development firms.
- Academic institutions for research collaboration.
- Data security and privacy experts.
- Government and non-governmental agencies for information sharing.

9. Cost Structure:

- Development and maintenance of AI systems.
 - Personnel salaries.
 - Training and development costs.
-

- Data security and infrastructure expenses.

10. Potential Challenges:

- Ensuring data privacy and security.
 - Addressing ethical concerns and bias in AI.
 - Keeping up with rapid technological advancements.
 - Balancing transparency and confidentiality needs.
-

4.4 Results for Research Question RQ4

The results were positively surprising, especially with regard to a general LLM. At a first glance, the business canvas covered breadth and depth. The suggestions could be assessed as practical and as accurate. In addition to many overlaps, there are also some differences, such as arguments of another nature with an emphasis on other advantages. The comparison of the content of the 10 response fields revealed minor differences, but no completely diverse arguments nor fundamental distinct approaches.

Each prompt generated ten well-filled business canvas boxes. None of the fields differed fundamentally in wording from the other two fields, which were created at a different time. This may perhaps be due to the limited options for organizational design in the parliamentary environment. Errors, hallucinations, or purposeful deceptions could not be detected in the variety of results. Also, no ethical concerns and issues were addressed by the LLMs. The proposal for "potential revenue generation through commercialization of AI solutions developed for parliamentary use" may cause irritation among insiders. To be honest, such a proposal is not entirely far-fetched and can certainly be considered realistic in times of limited budgets. Proposals that deliberately mislead stakeholders were not found in the business canvas results. The LLMs suggest external AI experts and consultants, international parliaments for best practice sharing, NGOs focused on transparency, governance, and civic tech as well as data security and privacy experts. The proposals to involve external partners, which would open the door to trojan horses, i.e. manipulators and other enemies of the state, parliament, and democracy, need to be reviewed. None of the proposals are irrational at this point. In addition, there are established security checks to protect state institutions, which should always be applied in a partnering context. Overall, the results can be justified with regard to the task and its context (public management and digital governance). It is merely the application of a business canvas with reference to a parliament and its organization, but not the creation of a law or regulation based on an LLM. Opportunities and risks must nevertheless be adequately reflected. In this respect, it is advisable to review the generated proposals for the concretization of a business model by real people in charge.

This exemplary confrontation of an LLM with a thoroughly unfamiliar environment and the resulting answers, which can be considered as relevant, leads to the question of what else an LLM can be enabled to do in a parliamentary context. Without wanting to give a definitive answer at this point, it depends on which prompt is entered and secondly how the overall system has been fine-tuned so that critical or dangerous questions are not answered or only answered inaccurately. In fact, an LLM is just an AI-based IT system that has a large, trained language model and uses probabilities to determine which could be the best output for an input. The test case was developed on the basis of general language models, that can probably only be marginally trained for the specifics of a parliament. However, without access to the

detailed training material, it is not possible to make more substantiated statements. Parliamentary data and texts are luckily not particularly protected under data protection and copyright laws. As common property, they are therefore perfectly suitable as training material for LLMs. If they are specifically trained LLMs using parliamentary texts and laws, they will be able to deliver much better results. Further points and details, which are not apparent at first glance, could be generated by the user at any time later through a dialog-like deepening and deliberate provocations in the sense of a devil's advocate.

5 DISCUSSION

As part of an academic workshop on the use of AI in parliaments[24], colleagues from academia and parliamentary practice met in July 2023 to discuss the use of LLMs in a parliamentary context. Interesting insights were gained, but several concerns were also expressed. There was a debate about whether these systems should be used, in which cases, how and by whom these systems are used, and how this can be detected. There was concern about the emergence of new dependencies. The lack of transparency of LLMs, the lack of knowledge about how to feed and improve them and the overwhelming challenge for users to correctly assess the answers and their quality were criticized. Uncertainty and serious interest also characterized the discussion about how the skill sets are changing and which training measures are needed to teach the necessary new skills. The consequences for the understanding of the state and democracy were also debated.

Based on the results of the prompts and the subsequent considerations, an overall picture emerges that could be transferred into a SWOT analysis. Despite all the strengths and opportunities that these technologies offer, the weaknesses and threats must be considered. LLMs can react quickly to requests submitted via prompts and generate responses in seconds. They generally convey an understandable, linguistically skilled impression. However, there is no intelligence behind this, but rather a language model with a lot of knowledge that provides characters, words, data, answers, and information in short time. With the size of the LLM, the adaptability of the systems to topics and requirements increases. Due to the learning ability, they can also absorb new content in the future. Regarding weaknesses, classical LLM transformer systems lack the necessary sensitivity to false information. They take learned information as given, regardless of whether these are accurate or incorrect. This naturally makes them vulnerable to disinformation, irony, and sarcasm. Sometimes they even produce hallucinations, i.e. linguistically appropriate answers to a posed question, which, however, do not correspond or only partially correspond to the truth. Neutral observers would call this lying. LLM transformer systems unfold many new opportunities. Language translations are possible nearly in real time. Assistance services facilitate the conduct of meetings by providing valuable support in agenda setting and minute taking. LLMs can also generate pre-wordings for appropriate analyses. They also provide valuable assistance in the creation of creative content and brainstorming. However, users must not ignore the risks of LLM transformer systems. Such systems could have been manipulated by malicious people without this being directly noticeable in the answers. For example, clues could be deliberately built into the information learned that would enable (political) influence on the answers. This can also lead to discrimination against certain groups. Security risks and vulnerabilities must be considered, as well as a lack of data protection and privacy if these systems are learning based on input with real personal data.

Based on this foundation, a brief impact assessment is recommended with a view to the long-term use of LLMs in parliament, an increase in critical stimuli as well as acceleration and diversity of perspectives. Regarding the field of application of parliament and legislation, this is a real high-risk area for the use of AI. Each LLM must be assessed separately regarding its specific characteristics, scope of services, language scope and learning capabilities. If there are no regulations or bans, employees and MPs will experiment and work with the tools and LLMs available. The results of prompts will be accurate, partly accurate or incorrect. The better are the results, the more the services are used. The users must be able to handle the results, evaluate them for their suitability and then use, modify, or reject them. This requires additional skills in dealing with generative AI. If users in parliament, where the correctness and accuracy of results is an essential requirement, do not have these skills, the use of LLMs must be strongly discouraged. Training programs help to close this gap and to deal appropriately with the opportunities and risks. Using a search window or verbal query, known and unknown data treasures can be uncovered by a trained LLM and utilized with templates and prompt specifications. Not only positive statements, but also critical, unpopular positions can be collected and introduced in a targeted manner. This exposes areas for attack. In any parliament, an LLM could be a promising way to manipulate. The interest of vendors, lobbyists, and other states, but especially hostile powers and their intelligence services and military, in influencing a parliament will remain strong. It cannot be ruled out that a manipulator with the necessary resources (military, intelligence service, organized crime) will succeed, either obviously or unnoticed. In this respect, every prompt result must be checked by the users for its suitability, relevance, and consequences before it is used. The advantages of an LLM in terms of speed, usefulness, and diversity of perspectives encourage increased use. Without great effort and with an excellent prompt, different perspectives can be seen, other points of view can be reflected, and new insights can be gained. Professional users work specifically towards the desired results by anchoring the necessary commands in the prompt, generating a wide variety of different results, making a selection, and then drilling down at the appropriate points. This will improve results and speed up processes in the organizations. In this way, LLMs open approaches for providing new, even surprising or disruptive stimuli and filling gaps that have not yet been considered.

6 CONCLUSION AND OUTLOOK

The test settings have shown that an LLM can already produce high-quality responses. The comparative analysis demonstrates that correct results can be generated, but that these may differ from each other. There is no 100% certainty of accurate answers. LLMs are only as good as their database and programming. They are not knowledge systems, but merely language models that often lead to plausible results based on probabilities. In view of the strengths and weaknesses, opportunities and risks, the question of whether LLMs can be used in the parliamentary arena and thus in politics and legislation can be answered with: "Yes, but only with restrictions".

Based on the investigation, the team was able to determine that ChatGPT in combination with GPT-3.5 and GPT-4 is already capable of providing qualitative and partially creative responses, which are suitable in the parliamentary routine for preparing decisions, texts, and meetings. However, due to the variations and the fact that there is no certainty about the truthfulness of the answers, which is of utmost priority in the parliamentary context, the generated content should now only be used as a basis for reflection and further processing by humans. Therefore, the current capabilities of ChatGPT should be viewed more as

preparatory or stimulating inputs that serve as a basis for more in-depth elaborations and considerations. Translations and summaries could be subject to a further review.

Through the compact SWOT analysis, the team was able to address the internal and external risks of using LLM's in the parliamentary context. These insights about LLM's in general help to understand and classify the insights about ChatGPT in particular. In the investigation, the team found that ChatGPT itself is familiar with some of the logics and procedures of the parliamentary environment. However, due to its technical design and the probabilistic function based on a broad LLM, it cannot take on the role of an in-depth process and structural analysis, and its responses tend to be more general. Therefore, the answers can provide a first impression of potential improvements, but they cannot deliver results that analyze and consider the specific framework conditions of individual parliaments. However, if the specific application challenges of a parliament are known, suggestions can be made based on the naming of these problems on how these challenges could be solved. From this perspective, a review and reflection by human actors remain urgently advised.

It is important to note that the underlying language model can improve, and therefore, the understanding of risks and concerns can become more comprehensive. During the investigation, the team did not get the impression that ChatGPT responds manipulatively or negligently in the context of risks and challenges. Even the indirect investigation through the inquiry of a business canvas led to ChatGPT providing some risks and concerns at the end of the answer without being asked to do so. Here too, ChatGPT provides some important points and suggestions for an assessment by human actors. However, the questions of risks and concerns remain highly political and should continue to be critically questioned and democratically negotiated. Especially in this context, the limitations of ChatGPT should not be blindly adopted, but further dangers should be thoroughly considered in each use.

The responses generated by GPT-X with the same prompt at different times are sometimes different, but they can also be very similar. Only the headlines and focal points varied, with relatively consistent content. Also, the order of the delivered responses was sometimes the same and sometimes different. Certain arguments, such as the erosion of skills, were exclusively identified in responses to specific inquiries during our investigation. Despite the probabilistic functioning of the system, there seems to be a certain continuity in the answering of questions, not arbitrariness, but a low variance in the alternation and weighting of arguments.

LLMs can generate valuable responses in and around parliaments if they have been well trained and kept up to date. With well-placed prompts, they can provide direct, valuable suggestions that can be used to improve and optimize parliamentary proceedings. Nevertheless, some weaknesses and risks are glaring, especially hallucinations and the possibility of (unnoticed) manipulation by third parties. It makes sense to use LLMs as part of decision-support systems, as results can be produced quickly and in high quality. But only competent and trained people can be trusted to deal with the results, to evaluate and use them, to correct or reject them. The use of LLMs as decisive systems in the parliamentary arena without further human intervention, on the other hand, is not advisable. The risks for the democratic system would simply be too high in such a case, especially as there are visible or invisible opportunities for manipulation.

This study has some limitations. For example, it must be fundamentally questioned whether qualitative questions can be answered at all with the help of a system that works predominantly according to probabilities and is not actually capable of answering qualitative questions per se because it is just a large language model. Nevertheless, the team embarked on this experiment to see whether this tool, due to its size, its vocabulary, and its programming, is already able to leap over its shadow and its expectations. Regarding the questions posed, the results were really good, even if this was not necessarily to be expected. However, these qualities depend not only on the system and its language model, but also on the prompts posed and the background knowledge of the creator of the prompts. Another justified point of criticism is that only three queries were created in June, September, and December 2023 and these are used as source documents. Actually, the same query should have been asked 100 or 1000 times in order to seriously evaluate the lasting response quality of the LLM and its response breadth (e.g. by the number of repetitions). Such an approach with a long-term measurement could be the subject of further studies. However, this study was not concerned with the overall quality of an LLM such as GPT-3.5 or GPT-4, measurable via quality tests or an established benchmark, but with the existing quality for parliamentary use, which in practice is limited to only one or two queries.

The insights gained do not cover the entire range of possible issues that arise with the use of LLMs in a parliamentary context. From a linguistic perspective, for example, the qualitative differences in the generation of texts in different languages need to be deeper analyzed in further studies. From an information management perspective, the question must be answered which LLM systems from which provider may be used in a parliamentary context and which should be avoided. Additional options include operating an LLM locally, in the data center, at a hosting partner or using a service from the cloud. Both open source LLMs and close source LLMs are relevant here, whereby open source LLMs can be further trained and fine-tuned by the parliament administration. Parliaments have appropriate training material in the official languages. With the laws, ordinances, questions, reports, and minutes that have been ratified, they have a pool that can be used for further training without data protection and copyright barriers. Proven and recognized benchmarks for LLMs that can make reliable statements about response quality would certainly be desirable. The development and establishment of these benchmarks is another area for future research. This requires a set of standardized methods that result in reliable, evaluative statements about response behavior. Future LLM benchmarks for parliaments will make it possible to check the quality of LLMs freely available on the market against the requirements of a parliament. Benchmark services could also evaluate the quality of internal parliamentary LLMs that have been trained and fine-tuned with a package of parliamentary documents. A comparison would reveal differences, for example regarding cybersecurity level, national security requirements, official languages, linguistic diversity, or explainable AI. Recommendations can also be made. Comparison rings would help to set minimum standards and identify frontrunners. Not only national parliaments could benefit from this, but also sub-national houses of representatives and local councils, which might be overwhelmed by the diversity of LLM offers. As another requirement, it should be ensured that it will be possible to change the LLM provider at any time. The implementation of LLMs in parliaments requires the use of guidelines for the introduction of AI in parliaments [25]. Considering the weaknesses and risks on the one hand and the special requirements of

LLMs in parliaments on the other side, the development of guidelines for the use of LLMs in parliaments [58] is another important field for future research.

The long-term implications of LLMs in parliaments also need to be explored further: Do LLM-generated agendas already improve the quality of meetings and time management in MPs' offices and parliamentary groups? Are MPs' speeches becoming more comprehensible, convincing, and relevant thanks to LLMs? Are the outcome and impact of debates improving thanks to AI-based technology? Can better and more comprehensible laws be drafted with the help of LLMs? Do they allow a better assessment of the impact of legislation, for example regarding bureaucracy, contradictions, and future lawsuits? How does the rapid, LLM-based generation of impact assessments change the way of dealing with opportunities and risks? Do LLMs support the negotiations in the mediation process between the chambers with convincing compromise proposals? Or will politicians reject this because they can no longer become fully involved?

In summary, LLMs can already play a role in various aspects of parliaments today. However, those responsible should start using them carefully, gradually gain experience, and always use the results in a reflective manner.

HISTORY DATES

Received January 2024; revised March 2024; accepted May 2024.

This open access publication was funded by the Zeppelin University Open Access Publication Fund.

REFERENCES

- [1] Adams, R. (2021). Can Artificial Intelligence Be Decolonized? In: *Interdisciplinary Science Reviews*, 2021, 46(1-2), pp. 176-197.
- [2] Albrecht, S. (2023). ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen. TAB-Hintergrundpapier Nr. 26. Berlin.
- [3] Al-Mushayt, O. (2019). Automating E-Government Services With Artificial Intelligence. *IEEE Access*, 7, 146821-146829. <https://doi.org/10.1109/ACCESS.2019.2946204>.
- [4] Bartl, G. (2023). Krise und technologischer Solutionismus: Die politische Dimension des digitalisierten Umgangs mit Unsicherheit. In: Wagener, A., Stark, C. (Eds.) *Die Digitalisierung des Politischen. Sozialwissenschaften und Berufspraxis*. Springer VS, Wiesbaden, pp. 45-62.
- [5] Bubeck, S. et al. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. *ArXiv*, <https://arxiv.org/abs/2303.12712>.
- [6] Budhwar, P. et al. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. In: *Human Resource Management Journal*, 33(03), pp. 606-659.
- [7] Bundesregierung der Bundesrepublik Deutschland (2018). *Strategie Künstliche Intelligenz der Bundesregierung*. Berlin.
- [8] Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Science*, 376:20180080, pp. 1-8. <http://dx.doi.org/10.1098/rsta.2018.0080>.
- [9] Cirillo, D., Catuara-Solarz, S., Guney, E. (2022). *Sex and Gender Bias in Technology and Artificial Intelligence*. Academic Press.
- [10] Committee on Artificial Intelligence (2023). Consolidated working draft of the framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law. Strasbourg.
- [11] Commonwealth Parliamentary Association (2023). *Parliamentary Handbook on Disinformation, AI and synthetic Media*. London. <https://www.cpahq.org/media/ivih25ue/handbook-on-disinformation-ai-and-synthetic-media.pdf>.
- [12] Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. In: *Journal of Science Education and Technology*, vol. 32, pp. 444-452.
- [13] Council of Europe (2021). *Artificial Intelligence, Human Rights, Democracy, And The Rule Of Law – A Primer*, Council of Europe and The Alan Turing Institute, Strasbourg.
- [14] c't (2023): *KI-Praxis – Mit künstlicher Intelligenz produktiver arbeiten*, Heise Verlag.
- [15] de Swarte, T., Boufous, O. & Escalle, P. (2019). Artificial intelligence, ethics and human values: the cases of military drones and companion robots. In: *Artificial Life and Robotics*, volume 24, pp. 291-296.
- [16] Di Fede G., Rocchesso D., Dow S.P., Andolina S. (2022). *The Idea Machine: LLM-based Expansion, Rewriting, Combination, and Suggestion of Ideas*.

In ACM International Conference Proceedings Series. Association for Computing Machinery, New York, pp. 623-627.

- [17] Dogru, T. et al. (2023). Generative Artificial Intelligence in the Hospitality and Tourism Industry: Developing a Framework for Future Research. In: *Journal of Hospitality & Tourism Research*, 0(0). <https://doi.org/10.1177/109634802311886>.
- [18] European Commission (2020). White Paper on Artificial Intelligence – A European Approach to excellence and trust. Brussels.
- [19] European Commission (2021). Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Brussels.
- [20] European Parliament. (2023). Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI. Retrieved from <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>.
- [21] Executive Office of the President (2019). Maintaining American Leadership in Artificial Intelligence. Executive Order 13859 of February 11, 2019. Federal Register Vol. 84, No. 31. Washington D.C. pp. 3967
- [22] Feuerriegel, S., Hartmann, J., Janiesch C., Zschech P. (2023). Generative AI. *Business & Information Systems Engineering*, 66(01), pp. 111-126.
- [23] Fitsilis, F., von Lucke, J. (2023).: Beyond Contemporary Parliamentary Practice - Unfolding the institutional potential of Artificial Intelligence. In: *The Parliamentarian - Journal of the Parliaments of the Commonwealth*, 104(01): pp. 58-59.
- [24] Fitsilis, F., von Lucke, J., & Frank, S. (2023a). A Comprehensive Research Workshop on Artificial Intelligence in Parliaments. *International Journal of Parliamentary Studies*, 3(2), pp. 316-324.
- [25] Fitsilis, F., von Lucke, J., Mikros, G., Ruckert, J., Alberto de Oliveira Lima, J., Hershowitz, A., Philip Todd, B., Leventis, S. (2023b). Guidelines on the Introduction and Use of Artificial Intelligence in the Parliamentary Workspace. Version 1.
- [26] Floridi, L., Chiriat, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds & Machines* 30, pp. 681–694
- [27] Fountain, J., E. (2022). The moon, the ghetto and artificial intelligence: Reducing systemic racism in computational algorithms. In: *Government Information Quarterly*. 39(02).
- [28] Gartner (2023). Hype Cycle for Artificial Intelligence, 2023. <https://www.gartner.com/en/documents/4543699>.
- [29] Geertsema, P. G.; Bifet, A., & Green, R. (2023). ChatGPT and Large Language Models: What are the Implications for Policy Makers? <http://dx.doi.org/10.2139/ssrn.4424048>.
- [30] Heinrichs, B. (2022). Discrimination in the age of artificial intelligence. In: *AI & Society*, vol. 37, pp. 143–154.
- [31] High-Level Expert Group on Artificial Intelligence (2019). Ethic Guidelines for trustworthy AI. Brussels.
- [32] His Majesty's Government (2018). Industrial Strategy Artificial Intelligence Sector Deal. Crown Copyright. London.
- [33] Jobin, A., Ienca, M., Vayena, E. (2019). The global landscape of AI ethics guidelines. In: *Nature Machine Intelligence*. 1(09), pp. 389-399.
- [34] Jungherr, A. (2023). Artificial Intelligence and Democracy: A Conceptual Framework. In: *Social Media + Society*. 9 (3), pp. 1-14.
- [35] Keber, T., Wacke, J., Herzog, Y., Meding, K., Maslewski, D., Rembold, J., Steinbrück, A. and Skobel, E. (2023). Legal bases in data protection for the use of artificial intelligence. Discussion paper. Stuttgart. <https://www.baden-wuerttemberg.datenschutz.de/legal-bases-in-data-protection-for-ai/>.
- [36] Koryzis, D., Dalas, A., Spiliotopoulos, D., & Fitsilis, F. (2021). Parltch: Transformation framework for the digital parliament. *Big Data and Cognitive Computing*, 5(1), 15, 1-16. <https://doi.org/10.3390/bdcc5010015>.
- [37] Koryzis, D., Margaris, D., Vassilakis, C., Kotis, K. and Spiliotopoulos, D. (2023). Disruptive Technologies for Parliaments: A Literature Review. In: *Future Internet*, 2023, 15(66), 1-21. <https://doi.org/10.3390/fi15020066>.
- [38] Korzynski, P., Mazurek, G., Altmann, A., Ejdy, J., Kazlauskaitė, R., Paliszkiwicz, J., Wach, K. and Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. In: *Central European Management Journal*, 31(1), pp. 3-13.
- [39] Kumar, S., Choudhury, S. (2022). Gender and feminist considerations in artificial intelligence from a developing-world perspective, with India as a case study. In: *Humanities and Social Sciences Communications*, 9(31), pp. 1-9.
- [40] Langer, P.F., Weyerer, J.C. (2020). Diskriminierungen und Verzerrungen durch Künstliche Intelligenz. Entstehung und Wirkung im gesellschaftlichen Kontext. In: Oswald, M., Borucki, I. (eds). *Demokratietheorie im Zeitalter der Frühdigitalisierung*. Springer VS, Wiesbaden, pp. 219-242.
- [41] Lai, V.D., Ngo, N.T., Veyseh, A.P., Man, H., Dernoncourt, F., Bui, T., & Nguyen, T.H. (2023). ChatGPT Beyond English: Towards a Comprehensive Evaluation of Large Language Models in Multilingual Learning. *ArXiv*, <https://arxiv.org/abs/2304.05613>.
- [42] Mannuru, N. R. et al. (2023). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. In: *Information Development*, 0(0).
- [43] Maslej, N., et al. (2023). The AI Index 2023 Annual Report, Stanford.
- [44] Mazzone, M., Elgammal, A. (2019). Art, Creativity, and the Potential of Artificial Intelligence. *Arts* 2019, 8(1): 26, pp. 1-9.
- [45] McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E. (1956). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. In: *AI Magazine* (2006), 27(4), pp. 12-14.
- [46] Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1/2019, pp. 501–507.
- [47] National Cyber Security Centre et al. (2023). Guidelines for secure AI system development, https://www.ncsc.gov.uk/files/Guidelines-for-secure-AI-system-development.pdf?sc_src=email_3765039&sc_lid=361294739&sc_uid=ynhmtE2zh2&sc_lid=116.
- [48] Navigli, R.; Conia, S.; Ross B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. In: *Journal of Data and Information Quality*, 15 (2), pp. 1–21.

- [49] Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., Nguyen, B-P. T. (2023). Ethical principles for artificial intelligence in education. In: Education and Information Technologies, vol. 28, pp. 4221-4241.
- [50] Osterwalder, A. 2004: The Business Model Ontology. A Proposition in a Design Science Approach, Dissertation. University of Lausanne. Lausanne.
- [51] Osterwalder, A. 2005: What is a business model? Business Model Design and Innovation Blog. <https://www.alexosterwalder.com> and <https://web.archive.org/web/20061213141941/http://business-model-design.blogspot.com/2005/11/what-is-business-model.html>.
- [52] Otte, R. (2021). Maschinenbewusstsein: Die neue Stufe der KI – wie weit wollen wir gehen? Campus Verlag. Frankfurt am Main.
- [53] Pavlik, J. V. (2023). Collaborating With ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education. In: Journalism & Mass Communication Educator, 78(01), pp. 84-93.
- [54] Palmirani, M., Vitali, F., Van Puymbroeck, W., & Nubla Durango, F. (2022). Legal Drafting in the Era of Artificial Intelligence and Digitisation. Brussels: European Commission.
- [55] Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., & Turini, F. (2019). Meaningful Explanations of Black Box AI Decision Systems. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01): pp. 9780-9784.
- [56] Peres, R., Schreier M., Schweidel, D., Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. In: International Journal of Research in Marketing, 40(02), pp. 269-275.
- [57] Perlman, A. (2024). The Legal Ethics of Generative AI. Suffolk University Law Review, 1-18, Forthcoming, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4735389.
- [58] Popvox Foundation. (2024). The Modern Intern - 3 Steps For Working In Congress With GenAI. https://static1.squarespace.com/static/60450e1de0fb2a6f5771b1be/t/65ef47b9d15f07697a42eaab/1710180282939/The_Modern_Intern.pdf.
- [59] Prives, L. (2018). AI for All: Drawing Women into the Artificial Intelligence Field. In: IEEE Women in Engineering Magazine, 12(02), pp. 30-32.
- [60] Ruschemeier, H. (2023). Regulierung von KI in: Aus Politik und Zeitgeschehen, 73(42), Bundeszentrale für politische Bildung, Bonn, pp. 20-34.
- [61] Santow, E. (2020). Can artificial intelligence be trusted with our human rights? In: Australian Quarterly, 91(04), pp. 10-17.
- [62] Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., Anderson, R. (2023). The Curse of Recursion: Training on Generated Data Makes Models Forget. ArXiv, <https://arxiv.org/abs/2305.17493>.
- [63] Simanowski, R. (2023). Narrative der Weltbeglückung. In: Blätter, 68(06), pp. 63-73.
- [64] Stahl, B.C. (2021). Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies. Springer Cham, Cham.
- [65] Stahl, B.C., Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology, International Journal of Information Management, Volume 74, 2024, 102700.
- [66] Stanford University (2021). Artificial Intelligence Index Report 2021, Stanford.
- [67] Valle-Cruz, D., García-Contreras, R., & Gil-García, J. R. (2023). Exploring the negative impacts of artificial intelligence in government: the dark side of intelligent algorithms and cognitive machines. International Review of Administrative Sciences, 0(0).
- [68] Vaswani, A., Shazeer, A., Parmar, N., Uszkoreit, J., Jones, L., Gomez, AN., Kaiser, L., Polosukhin, I. (2017). Attention Is All You Need. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
- [69] Vincent-Lancrin, S. and R. van der Vlies. (2020). Trustworthy artificial intelligence (AI) in education: Promises and challenges", OECD Education Working Papers, No. 218, OECD Publishing, Paris.
- [70] von Lucke, J., Fitsilis, F. (2023). Using Artificial Intelligence in Parliament - The Hellenic Case. In: Lindgren, I., et al. Electronic Government. EGOV 2023. Lecture Notes in Computer Science, vol 14130. Springer, Cham, pp. 174-191.
- [71] von Lucke, J., Etscheid, J. (2020a). How artificial intelligence approaches could change public administration and justice. In: Jusletter IT, 21 December 2020.
- [72] von Lucke, J., Etscheid, J. (2020b). Künstliche Intelligenz im öffentlichen Sektor. In: HMD Praxis der Wirtschaftsinformatik 2020, (57): pp. 60-76.
- [73] von Lucke, J., Fitsilis, F.: Einschätzungen aus dem griechischen Parlament zum Einsatz von künstlicher Intelligenz in Parlamenten, in: Gunnar Auth und Tim Pidun (Eds.): 6. Fachtagung Rechts- und Verwaltungsinformatik (RVI 2023) - Nachhaltiges und digitales Regierungs- und Verwaltungshandeln in der vernetzten Gesellschaft, Vol. P-341, Gesellschaft für Informatik, Köllen Druck + Verlag GmbH, Bonn, pp. 122-143.
- [74] von Lucke, J., Fitsilis, F., Etscheid, J. (2023). Research and Development Agenda for the Use of AI in Parliaments, in: David Duenas Cid, Nadzeya Sabatini, Loni Hagen und Hsin-Chung Liao (Hrsg.): DGO '23: Proceedings of the 24th Annual International Conference on Digital Government Research, Association for Computing Machinery (ACM), New York, pp. 423-433.
- [75] White House (2022). Blueprint for an AI Bill of Rights. Washington D.C.
- [76] Zhang, S. et al. (2022). OPT: Open Pre-trained Transformer Language Models. ArXiv, <https://arxiv.org/abs/2205.01068>.
- [77] Zuiderveen Borgesius, F. J. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. In: The International Journal of Human Rights. 24(10).