# Generative AI and Public Deliberation: A Framework for LLM-augmented Digital Democracy

Nikos Karacapilidis[1*], Evangelos Kalampokis[2], Nikolaos Giarelis[1] and Charalampos Mastrokostas[1]

[1] *Industrial Management and Information Systems Lab, MEAD, University of Patras, 26504 Greece*
[2] *Information Systems Lab, Department of Business Administration, University of Macedonia, 54636 Thessaloniki, Greece*

**Abstract**
Aiming to augment the effectiveness and scalability of existing digital deliberation platforms, while also facilitating evidence-based collective decision making and increasing citizen participation and trust, this article (i) reviews state-of-the-art applications of LLMs in diverse public deliberation issues; (ii) proposes a novel digital deliberation framework that meaningfully incorporates Knowledge Graphs and neuro-symbolic reasoning approaches to improve the factual accuracy and reasoning capabilities of LLMs, and (iii) demonstrates the potential of the proposed solution through two key deliberation tasks, namely fact checking and argument building. The article provides insights about how modern AI technology should be used to address the equity perspective, helping citizens to construct robust and informed arguments, refine their prose, and contribute comprehensible feedback; and aiding policy makers in obtaining a deep understanding of the evolution and outcome of a deliberation.

**Keywords**
Large Language Models, Public Deliberation, Digital Democracy, Neuro-symbolic AI.

## 1. Introduction

Public deliberation is a rational, interactive, and respectful form of communication [2]; it is a complex process that requires thoughtful examination of diverse issues and listening to others' perspectives, aiming to conclude the public judgement on what represents the common good. In turn, public judgement comes from people working together, in a shared search for effective solutions to their community problems, and it requires information about an issue, knowledge of

* Corresponding author.
Proceedings EGOV-CeDEM-ePart conference, September 1-5, 2024, Ghent University and KU Leuven, Ghent/Leuven, Belgium
EMAIL: karacap@upatras.gr (A. 1); ekal@uom.edu.gr (A. 2); giarelis@ceid.upatras.gr (A. 3); cmastrokostas@ac.upatras.gr (A. 4)
ORCID: 0000-0002-6581-6831 (A. 1); 0000-0003-4416-8764 (A. 2); 0000-0003-2611-3129 (A. 3); 0009-0009-6143-4666 (A.4)

the diverse elements and perspectives of a problem, as well as an understanding of the relationships among them and the consequences and trade-offs associated with different policies.

Current digital platforms for public deliberation rely almost completely on the abilities of participants to generate, interpret, and meaningfully process the associated content. This may significantly limit the effectiveness of these platforms, especially in cases characterized by information overload, incomplete knowledge of participants on the subject under consideration, lack of past memory, inability or hesitation of participants to express their opinion, etc. To thoroughly augment the effectiveness and scalability of these platforms, facilitate evidence-based collective decision making, while also increasing citizen participation and trust, we need to thoroughly advance the synergy between human and machine reasoning that is offered by the current public deliberation platforms.

To address the above issues, this article proposes a highly pragmatic framework that builds on and significantly advances state-of-the-art approaches from the areas of Generative AI, Large Language Models (LLMs), Knowledge Graphs (KGs), and Explainable AI (xAI) for the public sector. The proposed solution is based on a neuro-symbolic AI architecture that improves natural language processing tasks, such as question answering, machine translation, and text generation, by combining the strengths of deep learning and symbolic reasoning. We argue that the combination of these cutting-edge AI technologies can play a major role in the establishment of the desired human-machine partnership and may boost the creativity and collaborative strengths of humans in digital democracy [8].

In this line, the proposed framework includes a highly interactive and user-centric toolkit, which will assist stakeholders (both citizens and policy makers) in understanding, arguing and reasoning during a public deliberation, without requiring them to be familiar with argument structures, inference rules and AI/NLP technologies. It can be considered as a digital assistant that meaningfully incorporates a series of novel functionalities and integrates them with traditional deliberation methods. These functionalities include: (i) information retrieval for evidence seeking purposes; (ii) fact checking, to validate one's feedback; (iii) knowledge elicitation, to capture one's expertise and experience; (iv) detection of contradictions; (v) argument building, elaborating pairs of claims and premises; (vi) recommendation of speech acts, prompting a user to refute, corroborate, or clarify his/her feedback; (vii) role playing, assigning roles to LLM-based agents aiming to get insights on the evolution and conclusion of a deliberation; (viii) explanations building, to interpret the algorithm behind and explain the chain of inference; (ix) report generation, to prepare concise summaries of the overall process.

This article also aims to provide valuable knowledge about how modern AI-based public deliberation systems should be developed to address the equity perspective, helping citizens to construct robust and informed arguments, refine their prose, and contribute comprehensible feedback; and aiding policy makers in obtaining a deep understanding of the evolution and outcome of a deliberation. At the same time, the proposed framework aims to significantly increase the social impact of digital public deliberation platforms, enabling the elaboration of complex societal problems calling for collective intelligence. The equal and documented participation of all, assisted through structured information and recommendations generated by LLMs, as well as informative explanations provided by xAI, may have a great potential to restore the faith of both citizens and policy makers in digital democratic deliberations.

The contribution of this article is threefold: (i) it reviews state-of-the-art ap-plications of LLMs in diverse public deliberation related issues, aiming to reveal their strengths, drawbacks, and limitations (Section 3); (ii) it proposes a novel digital deliberation framework that meaningfully incorporates Knowledge Graphs and neuro-symbolic reasoning approaches to improve the factual accuracy and reasoning capabilities of LLMs (Section 4); (iii) it demonstrates the potential of the proposed framework through two key deliberation tasks, namely fact checking and argument building (Section 5).

## 2. Research Methodology

For the development of the proposed digital deliberation framework, we adopted the design science paradigm. This research paradigm aims to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts. When applied to the ICT domain, it results in purposeful technological artifacts created to address important organizational problems [6]. In our case, the problem we aim to address is the low level of quality and trustworthiness of existing platforms supporting digital democratic deliberations, and their inadequacy to address the associated scaling issues, by paying much attention to the value of effective deliberation in producing rich and legitimate outcomes. The overall objective of our research is to develop a new digital toolkit that thoroughly augments the functionality offered to citizens and policy makers by the existing deliberation platforms. For this purpose, we have used the design science research methodology proposed by Peffers et al. [12] for research in the domain of information systems. In our research, we combined the above paradigm with the action research methodology [13]. Action research has been proven to enable the design, implementation and evaluation of ICT-based actions and changes in organizations, which address specific problems and needs (of high interest for practitioners), and at the same time create scientific knowledge (of high interest for the researchers). The complementarity between these two research paradigms, i.e. the design science and the action research, and the potential of integrating them, has been comprehensively discussed in the literature [6].

We have also enhanced our research methodology with a literature review concerning LLM applications in diverse functionalities associated with public deliberation, which is presented in Section 3. A comprehensive literature review has been proven to be critical in the action research process, enabling the extraction of valuable insights, putting the overall study into context, and providing grounds to ensure its relevance and efficacy. Although action research is practice-oriented, such a review serves a series of purposes, including the analysis and synthesis of existing knowledge on the topic, the evidence-based identification of research gaps and areas where further investigation and intervention are required, and the assurance that the study contributes to advancing knowledge in the field.

## 3. Related LLM Applications

Large Language Models (LLMs) are state-of-the-art text generation ML models that possess various Natural Language Understanding (NLU) capabilities. Recent research has demonstrated successful applications of LLMs, ranging from simple tasks such as the summarization of text data for argumentation purposes to more complex ones where a multitude of LLMs act as dialogue agents for large-scale public deliberation. Popular ML frameworks (e.g., PyTorch, HuggingFace's Transformers) enable the development, finetuning and deployment of various LLM-based approaches. LLMs can be valuable in digital deliberation since they may provide meaningful recommendations to augment the co-creation of ideas and assist policy makers in their decision-making processes. LLMs can generalize and produce new information that is not part of their training knowledge. However, this knowledge is stored in a non-interpretable manner, due to their black-box architecture; moreover, their generalization capabilities can often lead to hallucinations, in cases where there is no proper context in their prompt.

Aiming to reveal strengths, drawbacks, and limitations of LLMs when addressing various tasks in a deliberation setting, and accordingly provide insights for future research, the rest of this section presents and comments on such applications. As expected, most of the selected works were (pre-)published in 2023, a year that is broadly considered as a landmark for deep learning research due to the introduction and practical application of this technology. To start with, Chen et al. [5] distinguish various argumentation related functionalities (claim, evidence and stance detection, evidence classification, counter-argument generation, argument summarization) and evaluate the corresponding capabilities of various LLMs architectures in zero-shot and few-shot settings. Their experimental analysis shows the strong potential of LLMs in computational argumentation, while also highlighting existing limitations such as the reduced LLMs performance in a zero-shot (vs. a few-shot setting), and performance variations of different sized LLMs when deployed in a few-shot setting.

de Wynter and Yuan [20] evaluate GPT-3 and GPT-4 in terms of their argumentative reasoning capabilities by changing the input and output data representations for the LLMs and then measuring their performance on two argumentation tasks, namely argument mining and argument pair extraction. The authors dis-cover that the input and output representations had significant impact on the downstream performance of these LLMs. This sensitivity suggested that the application of the model to critical cases needs much attention. However, when they applied Chain-of-thought (CoT) prompting techniques, their results were more consistent despite the changes in the input and output data representations.

Wilson et al. [19] investigate the performance of LLMs for argument structure generalization tasks. Their experiments reveal that LLMs were in most cases capable of predicting novel arguments to known verbs in correct positions and different structures than those in their fine-tuning data. The authors argue that despite their great performance on various NLP tasks, LLMs have limited generalization capabilities for human-like argument structuring. For this reason, they propose their training in various differing contexts, which - in terms of the amount of training data involved - is only possible for a few high resource languages.

Tuvey and Sen [15] explore the ability of open-sourced LLMs to generate arguments based on the factual content found in legal cases. Their approach considers an analysis of legal cases that assigns a rhetorical role to each sentence of the document. This breakdown facilitates the extraction of "fact-argument" pairs from legal documents, enabling the training of generative models like Flan-T5 and GPT-2 for the task of argument generation. The authors assess the performance of these models using a dataset that includes 100 annotated documents. The experimental results reveal that the models, which are trained on longer sentence summaries, generate high quality arguments. With respect to limitations, the authors highlight the importance of using datasets of high quality.

Castagna et al. [4] present a comprehensive survey about argumentation-based chatbots and their abilities. Although their study focuses on earlier chatbot architectures, they also examine the benefits of using LLMs for computational argumentation. This work stresses that despite their NLU capabilities, LLMs exhibit a set of limitations, including that they: (i) struggle to explain their outputs even in the case of similar inputs; (ii) present factually incorrect information (hallucinations) based on false training data or mistakes in their reasoning process; (iii) have weak reasoning skills, being unable to handle complex tasks; (iv) may generate toxic and offensive language in their outputs, upon data used in their training. According to the authors, the techniques that have been proposed so far to mitigate these limitations are not fully successful.

Other works focus on deliberation assistance issues. For instance, Argyle et al. [1] propose a digital assistant that was designed to guide users in real-time discussions on divisive political topics. The proposed solution provides refined suggestions without changing the fundamental content or the stance of the messages. The assistant's suggestions are based on three rephrasing techniques concerning the restatement, politeness, and validation of a user's message. The authors conducted an experiment with more than 1500 participants and a total of 2742 rephrased message suggestions, of which two thirds were accepted by users, in that they succeeded to change the tone of a message without significantly altering its topic. The authors also examined the impact of message rephrasings on the conversation quality and reported improvements in quality without influencing respondents to adopt any specific perspective.

In a similar direction, Li et al. [10] explore the potential of autonomous co-operation among communicative agents and propose a novel framework, named role-playing, that enables them to collaborate toward completing tasks while re-quiring minimal human intervention. The proposed approach can guide chat agents toward task completion while maintaining consistency with human intentions. It can generate conversational data to assist the study of the behaviors and capabilities of multi-agent systems, providing a valuable resource for investigating conversational language models.

Wang et al. [16] investigate the reasoning capabilities of LLM models by experimenting with debate-like conversations between ChatGPT and users. The goal is to determine whether the LLM can consistently maintain and defend its belief throughout a debate, without being misled by the user. The authors propose an evaluation framework that utilizes various benchmarks to evaluate the failure rate of ChatGPT across different types of reasoning tasks, including mathematics, logic and commonsense. Their results indicate that ChatGPT is susceptible to being misled into accepting falsehoods, revealing vulnerabilities not captured by traditional benchmarks.

Zhu and Wang [21] investigate the process of human problem-solving with the support of LLMs. They use a simple software installation task as their case study. The task is performed by users of various expertise levels who are provided with the support of ChatGPT. The authors evaluate the influence of LLMs based on the results of the problem-solving task and the execution time needed to complete the task. They also observe the sequence of log prompts to understand how users utilize the LLMs. As far as limitations are concerned, they report that some users highlighted LLMs inaccuracies, such as repeated outputs, and lack of complex understanding.

The study of the above works revealed that despite the well-known strengths of LLMs in data generation, their application to digital deliberation comes with the following issues and limitations:

- The development of LLMs follows a rapidly evolving pace, resulting to that some models perform better than others depending on the argumentation context;
- The majority of the works considered lack any form of fact checking functionalities to mitigate the problem of hallucinations that occurs in every LLM architecture;

- With respect to prompt engineering techniques, there is no one-size-fits-all solution, thus certain user questions may not be well understood by LLMs, which in turn leads to poor outputs;
- •Most of the applications considered do not provide meaningful interpretations and explanations of the produced LLMs results; this is partially because they build on commercial LLMs (e.g., OpenAI's GPT models), which are closed-source in terms of code and weights; however, even in the case of open-source LLMs, the integration of functionalities elaborated in the field of xAI has not been thoroughly explored;
- Many applications rely on LLMs that were trained on toxic and offensive data; while several mechanisms have been already proposed to address this issue, none of the applications considered in this study has incorporated them.

## 4. A Unified LLM-KG Framework for Digital Deliberation

While LLMs are actually transforming the way that research is carried out, as well as the way that citizens and policy makers use them in diverse public sector applications, it is also becoming clear that there are many challenges to be ad-dressed when dealing with natural argumentation and deliberation settings, which are characterized by a complex structure, nuanced presentation and (re)framing of ideas, context-based sensitivity, need for evidence-based resolution of differences and conflicts of opinion, and need for transparency while fostering and promoting public judgement.

Taking the above into account, we propose a framework for LLM-augmented public deliberation that builds on the synergy of human and machine reasoning to give citizens a significant role through direct and impactful deliberation, improve deliberation quality and promote democratic reciprocity, i.e. the willingness to grant everyone the same right to express and advocate their views in the public sphere that we hope they will grant us. At the same time, the proposed framework provides policy makers with greater legitimacy to better understand public priorities and the reasons behind them, identify conflicts and areas where consensus is feasible or not, and accordingly conclude a deliberation. The proposed solution meaningfully incorporates Knowledge Graphs and neuro-symbolic reasoning approaches to improve the factual accuracy and reasoning capabilities of LLMs and, consequently, their trustworthiness.

Knowledge Graphs (KGs) enable the structuring and linking of knowledge into semantic representations in a transparent and scalable way. Recent KG-based approaches can represent real-world knowledge extracted from heterogeneous sources with varying form and structure. Overall, such approaches outperform classical rule-based ones in data-driven applications, since they allow for easy data integration, while accounting for future data transformation and expandability. At the same time, many Graph-ML algorithms facilitate the discovery of hidden insights from KG data. KGs have structural knowledge that is stored in the form of accurate and interpretable domain-specific facts; however, they are unable to handle cases of missing or incomplete facts, and they do not possess any NLU capabilities.

By integrating structured knowledge from KGs and neuro-symbolic reasoning, LLM systems will be more transparent and explainable, suitable for sensitive applications such as that of digital deliberation. Specifically, the proposed framework will facilitate the synergy between KGs and LLMs in the following ways [11]: (i) the structural domain-specific knowledge stored in KGs will be utilized as contextual facts for LLMs, mitigating possible hallucinations; (ii) to solve the indecisiveness issue, where LLMs generate different plausible answers for the same input, we will take into account domain-specific knowledge from KGs, which will guide LLMs' output towards the correct answer (out of many plausible ones); (iii) facts stored in KG can become incomplete due to the constant evolution of knowledge; NLU capabilities provided by LLMs will enable the inference of new and unseen facts and data from public deliberations, which in turn can be used to dynamically update the entities and relationships of the KG.

On top of a neuro-symbolic AI architecture, the proposed unified LLM-KG framework is able to capture factual knowledge, created during digital deliberation, and stored in a structured format. The foreseen solution facilitates the automatic creation of dynamic KGs, together with novel LLM-enhanced mechanisms to encapsulate various aspects of public deliberation, thus addressing the limitation of non-evolving and stale knowledge of traditional approaches. Moreover, the proposed framework addresses limitations that affect current approaches including: (i) scalability of KGs; (ii) fact-checking extracted information of LLMs using multiple sources and models to eliminate hallucinations; (iii) the combination of the strengths of KGs and LLMs to produce beyond state-of-the-art semantic models and KG-based approaches. The ubiquitous trade-off between computational efficiency and model expressiveness is addressed through the abovementioned unified LLM-KG framework that semantically structures factual knowledge. This structuring process enables the creation of a robust KG, while offering data interpretability. This interpretability aspect assists stakeholders to acquire a complete and informed understanding of the underlying decision-making algorithms, thus promoting transparency and trust of their produced results and recommendations. The conceptual architecture and components of the proposed solution, namely the Unified LLM-KG Framework, the Digital Deliberation Assistant Services, and the Data Management Services (Figure 1), are presented below.

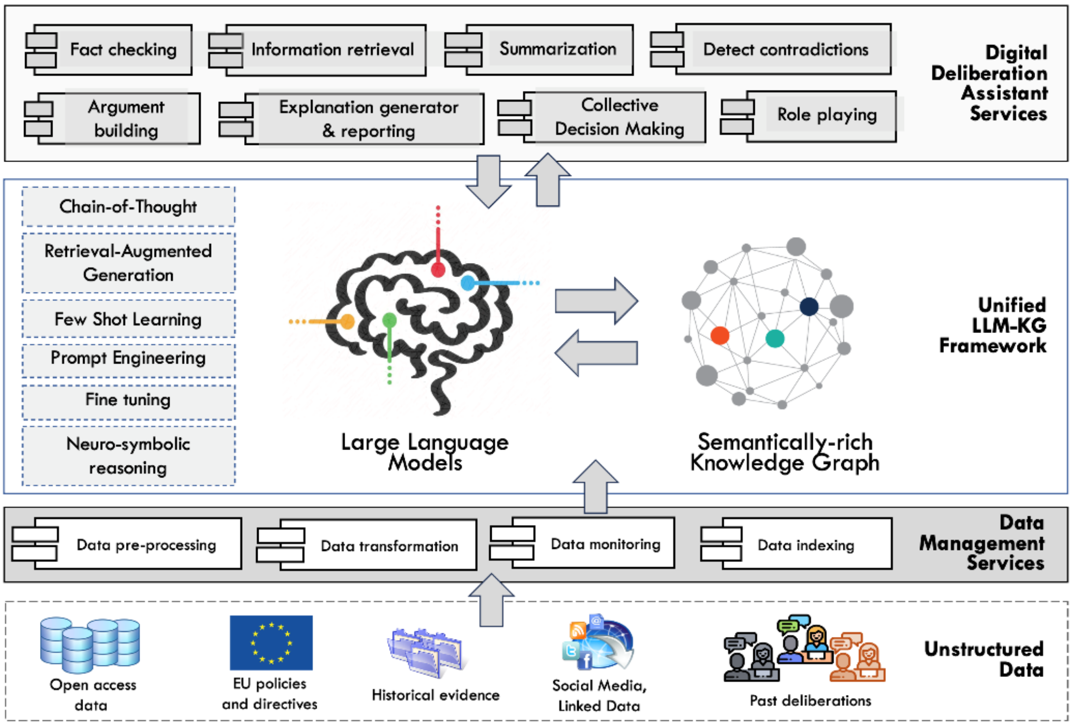


**Figure 1:** The proposed solution.

### 4.1. The Unified LLM-KG framework

The proposed solution explores novel techniques aiming to integrate LLMs with Knowledge Graphs and neuro-symbolic reasoning architectures. More specifically, it concerns: (i) a unified LLM-KG framework that combines the generative abilities of LLMs with the logical and factual coherence of KGs; (ii) techniques for compression and vectorization of KGs for neural networks, pattern extraction to link neural patterns with symbolic knowledge, and neuro-symbolic mapping in the foreseen LLM-KG architecture. These techniques will address the opaque-ness of LLMs and ensure that the decision-making process in the unified framework is transparent, explainable, and trustworthy. Techniques to be elaborated in this framework include:

- ***Few-shot Learning:*** A machine learning paradigm that focuses on training models to perform a task with only a small amount of labelled training data. It enables models to generalize and make accurate predictions even when provided with only a few examples, contrary to conventional methods that might face challenges in such scenarios due to data limitations [14].
- ***Chain-of-Thought (CoT):*** A type of few-shot learning, in which examples of chain-of-thought reasoning are provided in the model. It provides the model with a series of input statements or questions that lead it through a logical progression of information, enabling it to generate coherent and contextually relevant responses. CoT prompting allows LLMs to address intricate tasks involving arithmetic, commonsense, and symbolic reasoning [17].
- ***Retrieval-Augmented Generation (RAG):*** A technique that utilizes external data resources to augment user prompts with domain specific information. The user's original input is used to retrieve related text documents, which are processed and then used as supplementary context for the LLM to generate the desired output. This knowledge retrieval technique diminishes the likelihood of hallucinations, as the data is included in the prompt itself instead of relying solely on the internal knowledge of the LLM [9].
- ***Neuro-Symbolic Reasoning:*** A hybrid approach that allows machines to reason symbolically while leveraging the powerful pattern recognition capabilities of neural networks [7].
- ***Prompt Engineering:*** It refers to the practice of designing and refining the input prompts used in LLMs. Effective prompt engineering involves crafting input prompts in a way that evokes the desired information or behavior from the model, since general purpose LLMs may not always produce the expected output with a generic prompt [18].
- ***LLM Fine-tuning:*** It is a technique that refers to the process of taking a pre-trained model and further training it on a specific task or domain to improve its performance. Despite its benefits, fine-tuning has increased data and computational resource requirements compared to the aforementioned techniques [3].

### 4.2. Digital Deliberation Assistant Services

Building on the functionalities of the proposed unified LLM-KG framework, the proposed solution offers a set of services aiming to assist citizens and policy makers in various aspects of a digital deliberation process. These include:

- ***Fact checking:*** This service will enable stakeholders to verify and assess the accuracy and validity of information, claims, or statements presented during a deliberation (including those asserted without references and those referring to other outside sources). Moreover, it will facilitate citizens in creating accurate and persuasive arguments.
- ***Information retrieval:*** This service will act as an intermediary between digital deliberation stakeholders and large repositories of (unstructured) data. It will aid them to find relevant information by processing their queries, retrieving the appropriate documents, and presenting results in a way that meets their specific needs.
- ***Argument building:*** This service will assist citizens in creating robust and convincing arguments through novel mechanisms of argument generation and style transfer. It will build on human-machine learning techniques, using active learning for both the system and users, to improve and personalize argument classification and generalization. Much attention will be paid to the preservation of critical deliberation values such as openness, respect, reasoned discourse, and reliability.
- ***Summarization:*** This service will provide summaries of a deliberation (following a hybrid extractive and abstractive approach) that capture the different citizens' perspectives in an accurate and equitable way, and accordingly return a finite set of alternative options. The service will be based on criteria derived from deliberation theory; it involves assessment of clustering and classification algorithms to cross-validate its outcomes for potential biases.
- ***Detection of contradictions:*** This service will point out inconsistencies and contradictions within a digital deliberation process. It builds on the project's unified LLM-KG framework to extract and analyze statements, graph algorithms to check relationships within the KG, as well as formal logic and rule-based reasoning.
- ***Role playing:*** This service will enable participants to simulate different deliberation personas, representing various points of view, roles or perspectives on a certain topic. In this way, they can mimic the behavior, language, and responses of a particular persona, and get valuable insights about the evolution and outcome of a deliberation. The service can be also used to model facilitators that guide a deliberation, by posing probing questions, and stimulating a deeper exploration of certain arguments and statements.
- ***Explanation Generator and Reporting:*** This service aids citizens and other stakeholders to get a complete and informed understanding of the inferential process of the underlying machine learning algorithms and decision making mechanisms and promotes trust for the deliberation outputs. The service utilizes the KG structure, contents and semantics to offer diverse explainability and reporting functionalities.
- ***Collective Decision Making:*** This service adopts a knowledge-based decision-making view, enabled by the LLM-KG framework. It will deploy input aggregation and voting rules that work effectively with AI-assisted deliberation; it will also augment collective decision making with effective interplays between AI-based deliberation and decision support. The overall approach efficiently addresses the associated scaling issues.

## 5. Examples of Use

A series of application scenarios of the proposed solution have been developed in collaboration with representatives from citizen assemblies, governmental agencies at the regional and national level, and two think tanks working on democracy and public policy issues. It is noted here that we conducted three hourly workshops with these representatives, adopting a qualitative approach with in-depth discussions to collect relevant information. Specifically, we first presented them the basic idea, and asked them to elaborate on its feasibility and usefulness. We asked them to envision the proposed framework in their specific context, asking them to identify and prioritize the associated functionalities required. Then, we collaboratively developed with them specific application scenarios for a subset of these functionalities, also defining the main types of questions to be asked by the users of the proposed framework (citizens and policy makers). Finally, we asked them to think about data sources (structured and unstructured) that are particularly useful for the framework's data management services.

To describe the use of the proposed digital deliberation framework, this section illustrates two application scenarios concerning fact checking (performed by a policy maker) and argument building (performed by a citizen). The mockup shown in Figure 2 illustrates an instance of the foreseen functionality concerning a request from a policy maker for fact-checking. The digital assistant utilises its LLM part to extract the possible claims from the user's input, and formulates them into triplets (e.g., `[pesticides, banned, Europe]`, `[pesticides, exported, developing countries]`). These triplets are used to query the proposed framework's Knowledge Graph, which returns a series of triplets containing related facts (e.g., `[pesticides, associated, cancer]`, `[pesticides, associated, neurodegenerative diseases]`). The digital assistant then uses these factual triplets as input to the LLM to appropriately form an answer. In the instance shown in Figure 2, two claims have been identified and for each of them the LLM provides a list with the supporting facts.

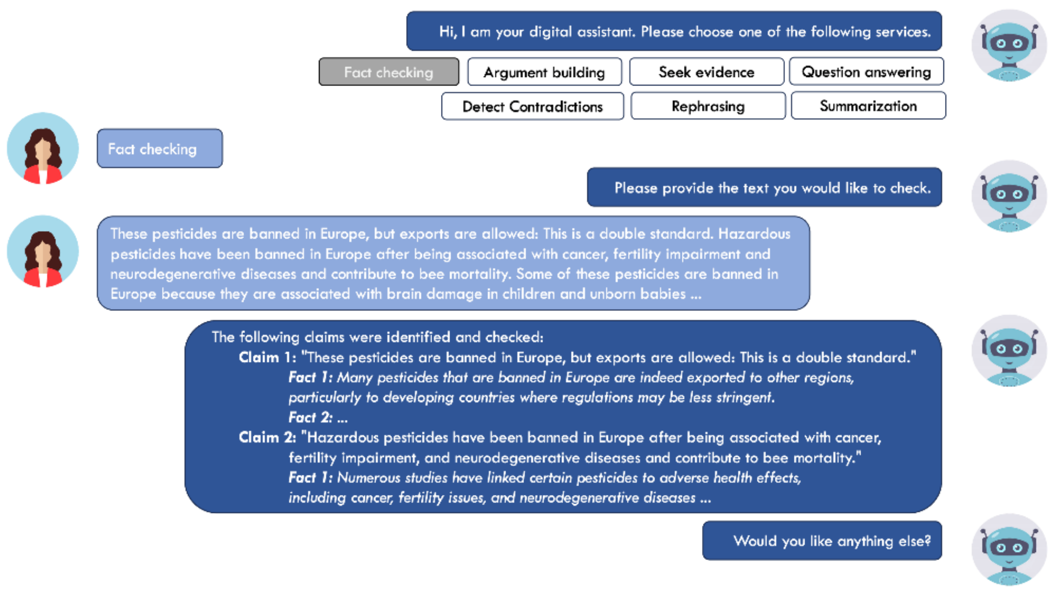


**Figure 2:** An instance of the "fact checking" functionality.

The mockup shown in Figure 3 illustrates a second instance of the foreseen functionality, this time concerning a request from a citizen for argument building. The digital assistant utilizes its LLM-part to extract the possible premises from the input using NLU, and formulate them into triplets (e.g., `[pesticides, contribute, environmental pollution]`, `[pesticides, pose,` `health risks]`). These triplets are used to query the Knowledge graph, which returns a series of triplets containing related evidence to the digital assistant (e.g., `[pesticides, cause, contamination]`, `[pesticides, associated, health issues]`). As in the previous example, the digital assistant uses these triplets as input to the LLM to form an appropriate answer. In the instance shown in Figure 3, the LLM suggests the building of an argument "supporting the ban on the export of toxic pesticides and advocating for innovation towards more sustainable plant protection methods" and returns evidence about the premises identified.

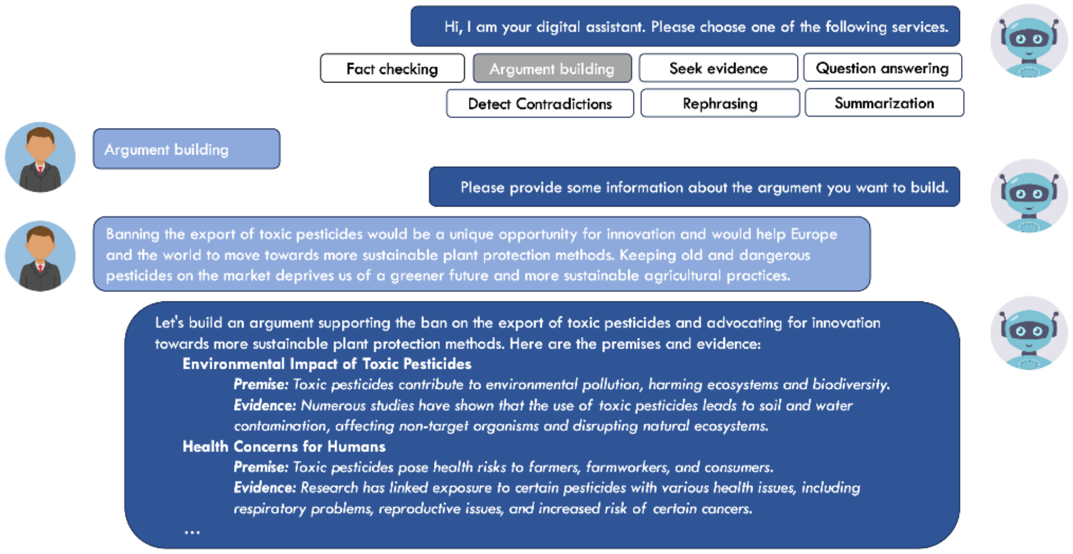


**Figure 3:** An instance of the "argument building" functionality.

## 6. Conclusions

Arguing that public deliberation processes should build on the synergy of human and machine reasoning, this paper has presented a digital deliberation framework that adopts a neuro-symbolic AI approach to combine the strengths of LLM-based processing of diverse deliberation items with symbolic representations facilitated by Knowledge Graphs. Our work contributes and provides valuable knowledge about how LLM-based public de-liberation systems should be developed to encourage citizens to contribute more argumentative and comprehensible contributions.

Preliminary assessment results demonstrate that the proposed approach may significantly increase the social impact of digital public deliberation platforms, enabling the elaboration of complex societal problems calling for collective intelligence. The equal and documented participation of all, assisted through consultative and structured information and recommendations generated by LLMs, has a great potential to restore the faith of citizens in democratic deliberations. The main limitation of our study is that though the proposed framework has gone through a first level assessment and validation by experienced practitioners, which has been positive, its application has to be carefully planned based on the available resources of different public sector organizations; this constitutes the basic direction for future work. In addition, its application has to be thoroughly assessed through a set of carefully defined Key Performance Indicators, which fit well to the field of digital democracy.

## References


[1] Argyle, L. P., Bail, C. A., Busby, E. C., Gubler, J. R., Howe, T., Rytting, C., ... & Wingate, D. (2023). Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale. *Proceedings of the National Academy of Sciences*, *120*(41), e2311627120.

[2] Bächtiger, A., & Pedrini, S. (2010). Dissecting deliberative democracy: A review of theoretical concepts and empirical findings. *Political discussion in modern democracies: A comparative perspective*, 9-25.

[3] Bakker, M., Chadwick, M., Sheahan, H., Tessler, M., Campbell-Gillingham, L., Balaguer, J., ... & Summerfield, C. (2022). Fine-tuning language models to find agreement among humans with diverse preferences. *Advances in Neural Information Processing Systems*, *35*, 38176-38189.

[4] Castagna, F., Kokciyan, N., Sassoon, I., Parsons, S., & Sklar, E. (2024). Computational Argumentation-based Chatbots: a Survey. *arXiv preprint arXiv:2401.03454*.

[5] Chen, G., Cheng, L., Tuan, L. A., & Bing, L. (2023). Exploring the potential of large language models in computational argumentation. *arXiv preprint arXiv:2311.09022*.

[6] Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22). Springer Science & Business Media.

[7] Hitzler, P., Eberhart, A., Ebrahimi, M., Sarker, M. K., & Zhou, L. (2022). Neuro-symbolic approaches in artificial intelligence. *National Science Review*, *9*(6), nwac035.

[8] Kalampokis, E., Karacapilidis, N., Tsakalidis, D., & Tarabanis, K. (2023). Understanding the use of Emerging technologies in the public sector: A review of Horizon 2020 projects. *Digital Government: Research and Practice*, *4*(1), 1-28.

[9] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

[10] Li, G., Hammoud, H., Itani, H., Khizbullin, D., & Ghanem, B. (2023). Camel: Communicative agents for" mind" exploration of large language model society. *Advances in Neural Information Processing Systems*, *36*, 51991-52008.

[11] Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*.

[12] Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, *24*(3), 45-77.

[13] Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 37-56.

[14] Song, Y., Wang, T., Cai, P., Mondal, S. K., & Sahoo, J. P. (2023). A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *ACM Computing Surveys*, *55*(13s), 1-40.

[15] Tuvey, O., & Sen, P. (2023). Automated Argument Generation from Legal Facts. *arXiv preprint arXiv:2310.05680*.

[16] Wang, B., Yue, X., & Sun, H. (2023, December). Can ChatGPT defend its belief in truth? evaluating LLM reasoning via debate. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 11865-11881).

[17] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, *35*, 24824-24837.

[18] White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., ... & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382*.

[19] Wilson, M., Petty, J., & Frank, R. (2023). How abstract is linguistic generalization in large language models? Experiments with argument structure. *Transactions of the Association for Computational Linguistics*, *11*, 1377-1395.

[20] de Wynter, A., & Yuan, T. (2023). I wish to have an argument: Argumentative reasoning in large language models. *arXiv preprint arXiv:2309.16938*.

[21] Zhu, Q., & Wang, H. C. (2023, October). Leveraging Large Language Model as Support for Human Problem Solving: An Exploration of Its Appropriation and Impact. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 333-337).