



HAL
open science

How to Streamline AI Application in Government? A Case Study on Citizen Participation in Germany

Dian Balta, Peter Kuhn, Mahdi Sellami, Daniel Kulus, Claudius Lieven,
Helmut Krcmar

► **To cite this version:**

Dian Balta, Peter Kuhn, Mahdi Sellami, Daniel Kulus, Claudius Lieven, et al.. How to Streamline AI Application in Government? A Case Study on Citizen Participation in Germany. 18th International Conference on Electronic Government (EGOV), Sep 2019, San Benedetto del Tronto, Italy. pp.233-247, 10.1007/978-3-030-27325-5_18 . hal-02445808

HAL Id: hal-02445808

<https://inria.hal.science/hal-02445808>

Submitted on 20 Jan 2020


HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

How to Streamline AI Application in Government? A Case Study on Citizen Participation in Germany

Dian Balta ¹, Peter Kuhn ¹[0000-0001-6774-2904], Mahdi Sellami ¹, Daniel Kulus ²,
Claudius Lieven ², and Helmut Krcmar ³

¹ fortiss GmbH, Guerickestr. 25, 80805 Muenchen, Germany
balta @ fortiss.org

² Behörde für Stadtentwicklung und Wohnen, Neuenfelder Str 19, 21109 Hamburg, Germany

³ Informatics 17 - Chair for Information Systems, Technical University of Munich,
Boltzmannstr. 3, 85748 Garching, Germany

Abstract. Artificial intelligence (AI) technologies are on the rise in almost every aspect of society, business and government. Especially in government, it is of interest how the application of AI can be streamlined: at least, in a controlled environment, in order to be able to evaluate potential (positive and negative) impact. Unfortunately, reuse in development of AI applications and their evaluation results lack interoperability and transferability. One potential remedy to this challenge would be to apply standardized artefacts: not only on a technical level, but also on an organization or semantic level. This paper presents findings from a qualitative explorative case study on online citizen participation in Germany that reveal insights on the current standardization level of AI applications. In order to provide an in-depth analysis, the research involves evaluation of two particular AI approaches to natural language processing. Our findings suggest that standardization artefacts for streamlining AI application exist predominantly on a technical level and are still limited.

Keywords: natural language processing, standardization, government

1 Introduction

AI represents a concept that incorporates various characteristics of intelligent systems that follow particular goals, including a formal representation of (incomplete) knowledge and an automated logical inference based on that knowledge [1]. Application domains have been discussed as far as the 1960s (e.g. [2]), followed by principles for design and application in the 1980s (e.g. [3]). Currently, if designed in an ethical and a trustworthy manner, AI is expected to represent a huge leap from data analysis to high quality and efficiency predictions, and increases the value of informed judgements decisions by humans [4].

Still, being in the trends spotlight for more than 10 years (e.g. [5]), AI applications prove challenging in government practice (cf. e.g. [6]). While application examples

continuously provided indications of AI's potential (cf. e.g. [7, 8]), they have yet to deliver sustainable and reproducible results in the government domain. In particular, online citizen participation has been in the focus of research and practice in the last decades regarding natural language processing (e.g. text mining), that would provide a more efficient and effective service to citizens and government (cf. e.g. [9–11]).

This paper states the argument that standardized artefacts (e.g. business processes, models, shared terminologies, software tools etc.) are required in order to streamline AI application in government. To support this argument, we present findings from a case study in Germany and discuss their implications.

The rest of the paper is structured as follows. First, we introduce the required theoretical background. Next, we present details and background information on the case study and describe our research approach. After this, we present our findings and discuss implication for AI application in government.

2 Theoretical Background

2.1 Standardization

We apply the definition of a standard as “a uniform set of measures, agreements, conditions, or specifications between parties” [27], and standardization represents the process of reaching a standard encompasses stabilizing and solidifying its definition and boundaries [12, 28]. Hence, standardization can be described in detail as “the activity of establishing and recording a limited set of solutions to actual or potential matching problems directed at benefits for the party or parties involved balancing their needs and intending and expecting that these solutions will be repeatedly or continuously used during a certain period by a substantial number of the parties for whom they are meant” [29].

In order to analyze IT standardization artefacts in government, the following framework that consist of two dimension can be applied [12, 13]. The first dimension includes three levels of interoperability and the second dimension includes five functional views (cf. table 1, with exemplary artefacts). The interoperability dimension is structured along three layers. First, the interoperability of business processes applied in delivering public service is found on the organizational layer. Second, interoperability regarding exchange of information and data as well as to their meaning between parties involved is found on the semantic layer. Third, interoperability regarding data structure and format, sending and receiving data based on communication protocols, electronic mechanisms to store data as well as software and hardware is situated on the technical/ syntactic layer.

The second dimension includes five functional views [12]. The administration view includes predominantly non-technical standards. They affect personnel and process aspects as well as communication within or between public administrations. For instance, a standardized business process definition or standardized shared terminology to describe public services or a standardized business reporting standard represent particular artefacts in this view. Second, the modeling view includes reference models and architectures, as well as modelling languages for each corresponding interoperability level.

For instance, an ontology can be applied to model a semantic standard towards the creation of a shared terminology and its sustainable use. Third, standards that focus the computation of data are included in the processing view. Exemplary artefacts in this view include a specific software application such as an information search service or a tax accounting software application. Fourth, corresponding standards for data and information exchange between different public administrations is handled in the communication and interaction view. For instance, a common metadata definition (i.e. data describing other data) is applied in a shared methodology in order to allow for an effective information search service. Fifth, the security and privacy view contains standards that aim at addressing issues such as definition of access management policies, application of cryptography methods or requesting a minimum of personal data and respecting privacy.

Table 1. Analysis framework and exemplary standardization artefacts in e-government

| | Administration | Modeling | Processing | Communication & Interaction | Security & Privacy |
|-----------------------|--------------------|--------------------------|-------------------------------------|-------------------------------|---------------------------|
| Organizational | Business Process | Reference Process | Business Process Modeling Tool | Process Model Exchange Format | Information Access Policy |
| Semantic | Shared Terminology | Ontology | Information Search Service | Common Metadata Definition | |
| Technical / Syntactic | Business Reporting | Application Architecture | Tax Accounting Software Application | Data Format | Encryption Algorithm |

An analysis of standards for e-government based on the described framework would include assigning them to one or several cells along the two dimensions. An assignment to multiple cells is possible, since a standardized solution might address different interoperability layers and functional views at the same time. Consequently, we apply the framework to analyze challenges (e.g. organizational and managerial, data as well as technological challenges [6]) of AI application in government and, in particular, in online citizen participation.

2.2 Online Citizen Participation

Online citizen participation can be described as a form of participation that is based on the usage of information and communication technology in societal democratic and consultative processes focused on citizens [14, 15]. Given the fact that different levels of online citizen participation and models to describe them exist [16], the participation

referred to in our research can be described as collaboration between citizens and government. In particular, government employees and/ or politician are still the ultimate decision makers, but a two-way communication between government and citizens takes place and the latter play an active role in proposing and shaping policy and decisions [17].

Implementation of online citizen participation is a challenging task [18] that includes the application of various techniques and technologies [19, 20]. With regard to existing models and frameworks of online citizen participation [21–23], the online citizen participation that lays the ground for our analysis can be described as follows. First, the participation process is steered top-down and is government led [23], where a public administration invites users to provide feedback by providing a set of topics and applies a set of participation techniques over a particular platform [19].

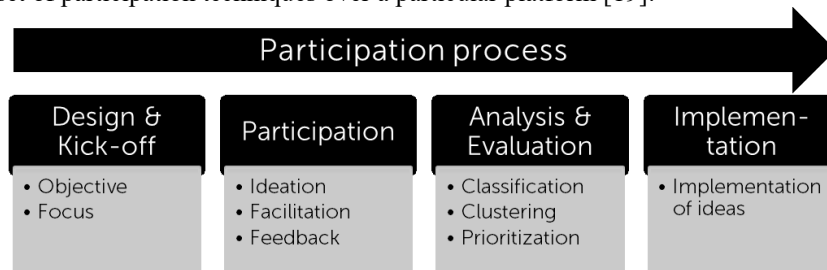


Fig. 1. Four generic steps of a citizen participation process

Second, the participation process can be described based on four generic steps from a government’s perspective (cf. figure 1). In the first step, public administration representatives design and kick-off citizen participation by setting objectives and by adjusting the focus to the participation goals. In a second step, citizens are invited to participate and provide feedback. In this step, ideation techniques are applied, and public administration employees aim at facilitating the process. In the third step, public administration employees or instructed service providers analyse and evaluate the generated citizens’ feedback. This includes classifying ideas based on a predefined set of topics and objectives as well as clustering all ideas into new feasible subsets in order to develop a summary report that includes a prioritization of citizens’ ideas. This step can require substantial effort from the involved employees, given the potentially large amount of citizen input to be analysed and evaluated. Finally, based on predefined policy processes, the implementation of the ideas is triggered (e.g. by providing the summarized report to action-taking parties such as architects’ offices).

2.3 Applied AI

According to the Association for Computing Machinery (ACM) Computing Classification System (CCS), AI is a broad research field with various application areas. In particular, AI employs machine learning methods that are based on algorithms that can learn from data without relying on rules-based programming. These methods emerged in the 1990s, making use of steady advances in digitization and cheap computing

power, and enabling to efficiently train computers to develop models for analysis and prediction. Recent developments in machine learning include novel models for knowledge representation based on neural networks, and logical inference is based on deep learning [24]. Neural networks represent a biologically-inspired programming paradigm which enables a computer to learn from observational data, while deep learning represents a powerful set of techniques for learning in neural networks [24].

An application of AI with a particular relevance for online citizen participation is a methodology named natural language processing (NLP). With NLP, tasks such as information extraction and summarization or discourse and dialogue or even machine translation can be automated to a certain degree. Consequently, goals of applying NLP in online citizen participation include designing a more efficient participation process through supporting the ideation (e.g. suggesting keywords or related contributions during ideation) as well as the analysis and evaluation (e.g. clustering and classifying user contributions). NLP has been already applied in government practice (e.g. [7, 8]) as well as in online citizen participation (e.g. [9–11]). While various tools and automated programmable interfaces (APIs) exist (cf. e.g. [25] for an overview), recent analysis shows that open source tools, that allow for a better control of data privacy and on premise operation of NLP, perform well in comparison with established API providers over closed source logical inference software and knowledge models [26].

Given the great number of applications, methodologies and tools, there are also numerous challenges to be addressed [6]. Hence, it would be of interest to analyse standardized artefacts potentially suitable for application in the government and, in particular, in the online citizen participation domain in order to effectively manage and streamline AI.

3 Online Citizen Participation in the Free and Hanseatic City of Hamburg and the Project Civitas Digitalis

Hamburg is a vivid city in northern Germany, with over 1.8 million citizens that has a strong economic development and a steady growth in terms of urban development and infrastructure projects. Since this growth affects a large number of citizens, citizen participation obtains an important role. In order to intensify information and participation in urban development projects and environmental protection issues and to develop a new planning culture in Hamburg, the ‘Stadtwerkstatt’ (city workshop) was set up as an organizational entity in the state-ministry of urban development and housing in 2012. Consequently, the Hamburg follows the concepts of top-down participation [23] by providing an own platform for participation and thus steering the democratic participation process (cf. chapter 2.2).

Since 2016, the city workshop unit offers a tool for online participation as part of its participation platform. The open source tool was developed in cooperation with the city’s Agency for Geoinformation and Surveying (LGV) and since then has been used in more than 30 participation processes, with a total of over 10,000 contributions created by users. This geodata-based web application allows citizens to gather information

about urban development projects and to submit contributions, including ideas, questions and criticism (cf. figure 2, accessible online https://geoportal-hamburg.de/beteiligung_grasbrook/mapview-beitraege).

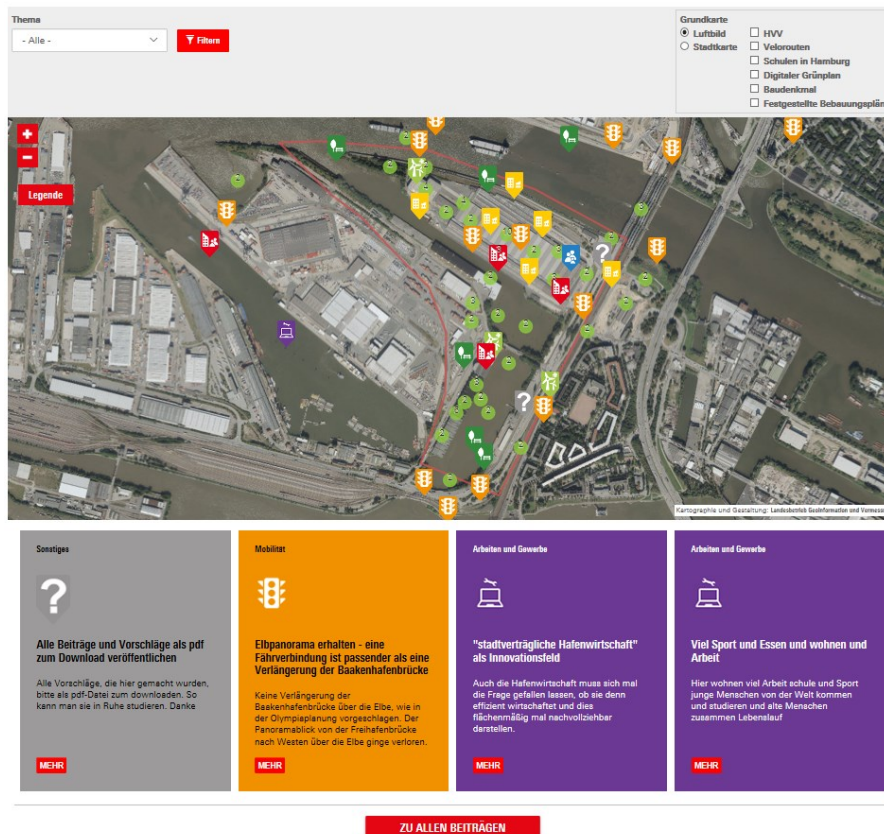


Fig. 2. An exemplary application of the online participation tool

The online participation tool is a basic online service that allows citizens to participate at any time and from any location. For the city administration and those officials responsible for project planning, the focus is particularly on greater reach and inclusion of social groups that are not able to participate at in-person meetings and workshops. For citizens who would like to participate, the tool provides an overview over the discussed topics, a thematic filter function and a city map visualization.

Since 2016, the city workshop offers a tool for digital participation as part of its participation platform. The open source online participation tool was developed in cooperation with the city's Agency Geoinformation and Surveying (LGV) and since then has been used in more than 30 participation processes, with a total of over 10,000 contributions created by users. This geodata-based web application allows citizens to gather information about urban development projects submit contributions, including

ideas, questions and criticism (cf. figure 2, accessible online https://geoportal-hamburg.de/beteiligung_grasbrook/mapview-beitraege).

The online participation is a basic online service that allows citizens to participate at any time and from any location. For the city administration and those responsible for project planning, the focus is particularly on greater reach and inclusion of social groups not able to participate at in-person meetings and workshops. For citizens who would like to participate, the tool provides an overview of which topics are discussed, a thematic filter function and a city map visualization.

A greater reach provides a higher number of citizens' ideas and, in consequence, a significantly bigger effort for the evaluation and analysis of the ideas by the public administration employees is required. Currently, contributions from digital and analog participation need to be merged in one digital file and are evaluated manually. The entire process is quite time-consuming. In several work steps, the contribution data are viewed, checked with regard to content and topic and, if necessary, differentiated into further categories or subcategories. At the end, the results are summarized, and potential courses of action are formulated.

In this context, the research project Civitas Digitalis (<https://civitas-digitalis.informatik.uni-hamburg.de/en/about-the-project/>) was initiated with Stadtwerkstatt as partner from practice. A project goal is to develop and evaluate a toolset for supporting a more efficient and effective online citizen participation.

4 Research Approach

We follow a qualitative analysis approach to explorative research. We aim at developing descriptive artefacts that can be categorized as a theory for analyzing [27]. Our research approach is rooted in the paradigm of pragmatism [28]. We studied the findings through an argumentative-deductive analysis [29].

4.1 Data Collection

For the analysis of this paper multiple types of data have been collected from three different sources. The first source of data are online citizen participation projects realized by the city of Hamburg in the last couple of years. We collected the ideas in German language from citizens in nine participation projects, cleaned the data by removing double entries or insufficient details entries. The total number of ideas was 3,859. The number of ideas per project vary from 95 to 1,689 ideas. The ideas, in turn, differ significantly in length, with the longest counting 1,079 words and the shortest 1 word. The median of words per idea is 15. Furthermore, the ideas have different properties. While most ideas consist of both the title and the description, some have only one of them. For instance, 1,018 ideas have no title. For every project, the ideas have been assigned by the citizen to a given category and after an idea has been submitted, the assignment has been analyzed, evaluated and potentially corrected by public administration employees. For the nine projects there is a total of 50 used categories—from now on called

subcategories—that for the purpose of the analysis have been combined into eight groups—from now on called categories.

The second source of data are interviews conducted in the course of this research that lasted between 35 and 60 minutes. One interview was undertaken with a citizen participation expert from a service provider that operates participation processes for municipalities. The second interview partner was an expert from the Stadtwerkstatt Hamburg which supports offices of the city with participation processes. Both interviews were conducted via telephone and based on a semi-structured questionnaires. The third interview has been conducted with an AI expert in order to evaluate the feasibility of the planned applications.

A workshop with 11 officials from the city of Hamburg and experts from service providers was the third source of data. In this workshop, the participants were asked to categorize 25 ideas from a recent citizen participation process (which is not part of the previously mentioned nine processes) into five predefined categories. The exercise was concluded by three groups of 3-5 experts in parallel, resulting in three classifications. Participatory observations as well as the allocation of ideas by the participants have been documented in writing.

4.2 Data Analysis

The analysis of the collected data included manually conducted reviews of the summarized interview reports, workshop results and reports as well as an automated NLP. The manual analysis was conducted by one researcher in our research group and was reviewed by a second researcher to assure consistency and to correct potential errors.

With respect to methodological development and practical tool availability of NLP (cf. e.g., [10, 11, 30–33]), we decided to analyze data based on a traditional machine learning approach as well as on a neural network and deep learning approach (cf. our open source implementation <https://civitasdigitalis.fortiss.org/> with data sets available upon request). Prior to applying each NLP approach, stratified sampling has been applied to the collected data set, splitting it into a training (80% - 3,087 ideas) and an evaluation (20% - 772 ideas) set, i.e. the resulting training and evaluation sets have the same distribution over the classes of ideas.

The first approach was based on the tool LingPipe [34] given its suitability for the analysis tasks (cf. e.g. [35, 36]). We implemented a character-level language-model based classifier. The classifier trains a model based on the occurrence count of characters and their combinations as well as the probability of both. The classifier predicts the assignment to a class based on the multivariate distribution of characters and their combinations to that class.

With regard to recent developments in neuronal network and deep learning approaches to NLP (cf. e.g. [37–39]), the second approach applied in our data analysis is a classifier based on BERT [40]. BERT implements a model architecture that includes a multi-layer bidirectional transformer encoder that can be configured to apply up to 1,024 hidden layers as well as 340 million parameters. Consequently, we decided to use the largest available BERT transformer that is pre-trained in a multilingual text

setup and recommended for German language. We experimented with different characters sequence length, batch sizes, learning rates and number of epochs during the customization of the model, in order to evaluate accuracy.

5 Findings

5.1 Level of Standardization of AI Application in Online Citizen Participation

Based on the analysis of documents, interviews and workshop results, we apply the framework for standardization analysis as follows (cf. table 2) and note “NA” in each table cell, if the analyzed data did not allow to present a finding. On the organizational level, there have been only general standardization artefacts regarding security and privacy. In particular, the EU General Data Protection Regulation (GDPR) has been considered as a guide for conducting any type of data analysis. For instance, there have been concerns regarding any application using an API or 3rd party software that was not hosted on computers in Germany or in our research group. These concerns have also been linked to still missing standards for communication and interaction, since vendor or tool lock-in could result. In terms of particular business processes or models, there have been only initial considerations such as at which step of the participation process AI would be suitable and of use.

Table 2. Analysis of standardization level of AI application in the case study

| | Administration | Modeling | Processing | Communication & Interaction | Security & Privacy |
|-------------------------|--|------------------------|---------------------------------|---------------------------------|--------------------|
| Organizational | initial considerations | initial considerations | NA | vendor or tool lock-in concerns | GDPR |
| Semantic | shared terminology in similar participation projects | initial considerations | NA | NA | |
| Technical / Syntactical | initial considerations | NA | tools based on machine learning | model or technique specific | NA |

At the semantic level, there has been a shared terminology regarding the classification of citizen ideas (e.g. “transportation” or “public space”). Still, this terminology has emerged and has not been a product of particular coordination effort. From a modeling functional perspective, there have been some implicit initial considerations (e.g. particular sentiment in an idea), but these have not been further detailed.

The technical and syntactic level included some initial considerations on how to integrate NLP techniques in existing tools (e.g. host an own analysis services, integration with front-end of the participation tool). Since the initial task of the researchers involved in the project was to apply NLP techniques, we were able to analyze a set of tools that can be applied and reused as standardized artefacts (e.g. LingPipe and BERT). Further, we were able to analyze standardization of communication and interaction between different tools. Based on the particular technology and technique applied, there are particular data exchange formats (e.g. a vector based representation of text for feature extraction). Still, there is no standard available and interoperability between tools and techniques have to be fitted to context.

5.2 AI Based Analysis

In order to compare both NLP approaches and their practical applicability, we conducted a number of tests with the data available (cf. table 3). Consequently, we applied a statistical confusion matrix that summarizes true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) as well as the F1 score to measure the accuracy of the prediction. For each category (listed in German), we had a different number of ideas split into training data and prediction test data.

For the NLP approach implemented with LingPipe, an average of 66.84% accuracy was observed. The accuracy varied significantly between categories. Although no correlation was analyzed in detail, our initial findings suggest that categories with higher numbers of ideas have better accuracy results.

Table 3. Comparison of the applied NLP approaches

| Category | LingPipe tool | | | | | | | BERT tool | | | | | | | Test Data | Training Data |
|-----------------------|---------------|----|-----|-----|--------|-----------|---------------|-----------|----|----|-----|--------|-----------|---------------|-----------|---------------|
| | TP | FN | FP | TN | Recall | Precision | F1 | TP | FN | FP | TN | Recall | Precision | F1 | | |
| Verkehr & Mobilität | 326 | 12 | 120 | 314 | 96,45% | 73,09% | 83,16% | 302 | 36 | 66 | 368 | 89,35% | 82,07% | 85,55% | 185 | 1504 |
| Wohnen & Arbeiten | 24 | 24 | 21 | 703 | 50,00% | 53,33% | 51,61% | 24 | 24 | 19 | 705 | 50,00% | 55,81% | 52,75% | 15 | 225 |
| Grün & Erholung | 20 | 38 | 8 | 706 | 34,48% | 71,43% | 46,51% | 32 | 26 | 28 | 686 | 55,17% | 53,33% | 54,24% | 24 | 265 |
| Sport & Freizeit | 74 | 21 | 30 | 647 | 77,89% | 71,15% | 74,37% | 76 | 19 | 32 | 645 | 80,00% | 70,37% | 74,88% | 50 | 426 |
| Klima & Umweltschutz | 1 | 18 | 1 | 752 | 5,26% | 50,00% | 9,52% | 7 | 12 | 8 | 745 | 36,84% | 46,67% | 41,18% | 10 | 85 |
| Städtebau & Stadtraum | 20 | 50 | 10 | 692 | 28,57% | 66,67% | 40,00% | 23 | 47 | 36 | 666 | 32,86% | 38,98% | 35,66% | 33 | 317 |
| Soziales & Kultur | 25 | 28 | 11 | 708 | 47,17% | 69,44% | 56,18% | 34 | 19 | 20 | 699 | 64,15% | 62,96% | 63,55% | 25 | 239 |
| Sonstiges | 26 | 65 | 55 | 626 | 28,57% | 32,10% | 30,23% | 27 | 64 | 38 | 643 | 29,67% | 41,54% | 34,62% | 44 | 412 |

The NLP approach implemented with BERT achieved an improvement of 1.56% and resulted in a total of 68.4% accuracy. In summary, both approaches showed promising results, the data available is of reasonable quality and there is a large enough data set. For instance, the best results were generated for the category “Verkehr & Mobilität” (transportation & mobility). This category contains 1.689 ideas in the data set and, arguably, the data quality is consistent among all ideas. On the contrary, the category “Sonstiges” (miscellaneous) contains 456 ideas in the data set that appear to be of varying quality, i.e. ideas were put in this category only if not matching any other. Apparently, there were no further semantic or any other classification rules applied.

5.3 Comparing AI and Human based Analysis

The collected human-based classification obtained in the workshop and the above described analysis allow for a comparison of human-based and AI-based categorization of participation ideas. The baseline for this comparison is the original category assignment of the contributor which we will consider as the correct category.

The groups of experts in the workshop categorized 14, 15 and 16 out of 25 categories correctly. This corresponds with a success rate of 60% in average. While some categories seemed to be easier to categorize for the experts, others appear to be less so – potentially, due to semantic heterogeneity. This was also reflected in the discussions among the workshop participants.

For the AI-based categorization, we applied both NLP approaches and selected the results with highest accuracy. Moreover, we developed an analysis improvement that delivered not only the first best fitting category, but also the second best fitting category. Considering, with the first best guesses considered, 8 out of 25 categories have been classified correctly which corresponds with a 32% success rate. Including also the second-best guesses, 13 out of 25 ideas have been assigned to their correct category, resulting in a success rate of 52%. As for the human-based results, the success rate by category varied also for the AI-based assignment.

Table 4. Human and AI based analysis

| | Human based | | | | AI based | |
|---------------------------------|-------------|---------|---------|------------|------------------|-----------------------------|
| | Group 1 | Group 2 | Group 3 | average | first-best guess | first and second-best guess |
| Correctly assigned ideas | 14 | 16 | 15 | 15 | 8 | 13 |
| Success Rate | 56% | 64% | 60% | 60% | 32% | 52% |

In comparison, the AI-based categorization is not as effective as the human-based one (cf. table 4). Considering the average of the workshop groups and the more favorable AI-based counting (including the second-best guess), there still remains a difference of two ideas less assigned correctly.

6 Discussion and Conclusion

The objective of this paper was to address the question of how to streamline AI application in government and, in particular, in the online citizen participation. We presented theoretical background on standardization, online citizen participation processes and application of AI and developed an argument that a standardization of AI artefacts at different levels and from different functional perspectives is required towards streamlining AI application. Based on a case study in Germany, we presented findings on levels of standardization, results from applying two different AI techniques for natural language processing as well as a comparison between human and AI performance.

Our findings show the following implications regarding AI application in government and, in particular, in online citizen participation. First, there are already NLP tools and pre-trained models available that can provide efficient support along the steps of the participation process. Quality, amount and availability of data seem to be of high importance for sufficient prediction, though. Second, human based analysis still has a number of advantages. As the results of our workshop show, humans are capable—in the course of intensive discussion and collaboration—to outperform AI and NLP. Additionally, the arguments provided by the workshop participants explained why a particular idea was assigned to a particular category. In the case with the implemented NLP techniques and tools, we were not capable of providing these insights.

Table 5. Future research on streamlining AI application in online citizen participation

| | Administration | Modeling | Processing | Communication & Interaction | Security & Privacy |
|-----------------------|--|--|---|--|---|
| Organizational | How does an online participation process with AI look like? | What is a reference model or an architecture, based on e.g. [22, 41] | How to integrate AI in existing tools for the design of online participation? | How to assure a technology shift? | How can GDPR conformity be evaluated and assured? |
| Semantic | What are shared context-specific concepts that can be integrated using AI? | What is a suitable ontology, e.g. based on [17, 23] | What are available AI services that fit requirements of online citizen participation? | How to share language specific AI results? | |
| Technical / Syntactic | How to monitor and manage AI applications? | How to integrate AI applications in existing IT infrastructures? | How to improve data quality and customize models? | Which competencies are required for the application of AI? | Are there any certified tools available? |

Third, our findings suggest that a standardization of AI application in government and, in particular, in online citizen participation is still in its infancy. There were only a few standardized artefacts available, predominantly on the technical/ syntactic level. Due to this current status, available data in suitable quality for efficient and effective AI application is even more challenging.

This research presents a first glimpse of the potential of and barriers to standardized artefacts for streamlining AI application in government. Given the contextual limitations—a case study on online citizen participation in Germany, data set size and quality, available tools and techniques etc.—we would like to encourage researchers to dig deeper in the sketched challenges and derive potential remedies. Therefore, we have summarized a number of questions that emerged during our research and could be addressed in future (cf. table 5). Additionally, future research could focus on interactions between the different standardization levels and their implications to adoption of AI in government. Given the number of human languages available and the expected potential of AI in government, we hope that future research would allow for a more efficient development and sustainability of AI application.

Acknowledgements

This research was partially funded by the German Federal Ministry of Education and Research (BMBF) with the project lead partner PTKA (Projektträger Karlsruhe am Karlsruher Institut für Technologie/KIT) in the context of the project Civitas Digitalis (funding code '02K15A050').

We thank our reviewers for their careful reading and their constructive remarks.

References

1. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited, (2016).
2. Minsky, M.: Steps toward artificial intelligence. *Proceedings of the IRE*. 49, 8–30 (1961).
3. Nilsson, N.J.: *Principles of Artificial Intelligence*. (1982).
4. Agrawal, A., Gans, J., Goldfarb, A.: *Prediction Machines: The simple economics of artificial intelligence*. Harvard Business Press (2018).
5. Chen, H.: AI, E-government, and Politics 2.0. *IEEE Intelligent Systems*. 24, 64–86 (2009).
6. Sun, T.Q., Medaglia, R.: Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*. (2018).
7. Androutsopoulou, A., Karacapilidis, N., Loukis, E., Charalabidis, Y.: Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*. (2018).
8. Tambouris, E.: Using Chatbots and Semantics to Exploit Public Sector Information. *EGOV-CeDEM-ePart 2018*. 125 (2018).
9. Teufl, P., Payer, U., Parycek, P.: Automated analysis of e-participation data by utilizing associative networks, spreading activation and unsupervised learning. In: *International Conference on Electronic Participation*. pp. 139–150. Springer (2009).

10. Maragoudakis, M., Loukis, E., Charalabidis, Y.: A review of opinion mining methods for analyzing citizens' contributions in public policy debate. In: International Conference on Electronic Participation. pp. 298–313. Springer (2011).
11. Rao, G.K., Dey, S.: Decision support for e-governance: a text mining approach. arXiv preprint arXiv:1108.6198. (2011).
12. Balta, D.: Effective Management of Standardizing in E-Government. Corporate Standardization Management and Innovation. 149–175 (2019). <https://doi.org/10.4018/978-1-5225-9008-8.ch008>.
13. Balta, D., Krcmar, H.: Managing Standardization in eGovernment: A Coordination Theory based Analysis Framework. In: Parycek, P., Glassey, O., Janssen, M., Scholl, H.J., Tambouris, E., Kalampokis, E., and Virkar, S. (eds.) Electronic Government. pp. 60–72. Springer International Publishing (2018).
14. Sæbø, Ø., Rose, J., Flak, L.S.: The shape of eParticipation: Characterizing an emerging research area. *Government information quarterly*. 25, 400–428 (2008).
15. Susha, I., Grönlund, A.: eParticipation research: Systematizing the field. *Government Information Quarterly*. 29, 373–382 (2012).
16. Al-Dalou, R., Abu-Shanab, E.: E-participation levels and technologies. In: The 6th International Conference on Information Technology (ICIT 2013). pp. 8–10 (2013).
17. Wimmer, M.A.: Ontology for an e-participation virtual resource centre. In: Proceedings of the 1st international conference on Theory and practice of electronic governance. pp. 89–98. ACM (2007).
18. Islam, M.S.: Towards a sustainable e-Participation implementation model. *European Journal of ePractice*. 5, (2008).
19. Tambouris, E., Liotas, N., Kaliviotis, D., Tarabanis, K.: A framework for scoping eParticipation. In: Proceedings of the 8th annual international conference on Digital government research: bridging disciplines & domains. pp. 288–289. Digital Government Society of North America (2007).
20. Tambouris, E., Liotas, N., Tarabanis, K.: A framework for assessing eParticipation projects and tools. In: 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07). pp. 90–90. IEEE (2007).
21. Youthpolicy: Participation Models: Citizens, Youth, Online. A chase through the maze, http://www.youthpolicy.org/wp-content/uploads/library/Participation_Models_20121118.pdf, (2012).
22. Scherer, S., Wimmer, M.A.: E-participation and enterprise architecture frameworks: An analysis. *Information Polity*. 17, 147–161 (2012).
23. Porwol, L., Ojo, A., Breslin, J.G.: An ontology for next generation e-Participation initiatives. *Government Information Quarterly*. 33, 583–594 (2016).
24. Nielsen, M.A.: Neural networks and deep learning. Determination press USA (2015).
25. Hagen, L., Harrison, T.M., Uzuner, Ö., Fake, T., Lamanna, D., Kotfila, C.: Introducing textual analysis tools for policy informatics: a case study of e-petitions. In: Proceedings of the 16th annual international conference on digital government research. pp. 10–19. ACM (2015).
26. Braun, D., Hernandez-Mendez, A., Matthes, F., Langen, M.: Evaluating natural language understanding services for conversational question answering systems. In: Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue. pp. 174–185 (2017).

27. Gregor, S.: The nature of theory in information systems. *MIS quarterly*. 611–642 (2006).
28. Goldkuhl, G.: Pragmatism vs interpretivism in qualitative information systems research. *European journal of information systems*. 21, 135–146 (2012).
29. Wilde, T., Hess, T.: Forschungsmethoden der Wirtschaftsinformatik. *Wirtsch. Inform.* 49, 280–287 (2007). <https://doi.org/10.1007/s11576-007-0064-z>.
30. Stumme, G., Hotho, A., Berendt, B.: Semantic web mining: State of the art and future directions. *Web semantics: Science, services and agents on the world wide web*. 4, 124–143 (2006).
31. Berry, M.W., Castellanos, M.: Survey of text mining. *Computing Reviews*. 45, 548 (2004).
32. Tan, A.-H.: Text mining: The state of the art and the challenges. In: *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*. pp. 65–70. *sn* (1999).
33. Mirończuk, M.M., Protasiewicz, J.: A recent overview of the state-of-the-art elements of text classification. *Expert Systems with Applications*. (2018).
34. Carpenter, B., Baldwin, B.: Text analysis with LingPipe 4. LingPipe Inc. (2011).
35. Carpenter, B.: LingPipe for 99.99% recall of gene mentions. In: *Proceedings of the Second BioCreative Challenge Evaluation Workshop*. pp. 307–309. BioCreative (2007).
36. Carpenter, B.: Phrasal queries with LingPipe and Lucene: ad hoc genomics text retrieval. In: *TREC*. pp. 1–10 (2004).
37. Alberti, C., Lee, K., Collins, M.: A BERT Baseline for the Natural Questions. *arXiv preprint arXiv:1901.08634*. (2019).
38. Gao, J., Galley, M., Li, L.: Neural approaches to conversational AI. *Foundations and Trends® in Information Retrieval*. 13, 127–298 (2019).
39. Hu, D.: An Introductory Survey on Attention Mechanisms in NLP Problems. *arXiv preprint arXiv:1811.05544*. (2018).
40. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. (2018).
41. Santana, E.F.Z., Chaves, A.P., Gerosa, M.A., Kon, F., Milojcic, D.S.: Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture. *ACM Computing Surveys (CSUR)*. 50, 78 (2018).