Analyzing Data from Urban Citizen Participation by applying the Retrieval Augmented Generation Architecture

Marten Borchers
University of Hamburg
marten.borchers@uni-hamburg.de

Enrico Milutzki
University of Hamburg
enrico.milutzki@uni-hamburg.de

Simon Oelgeschläger
University of Hamburg
simon.oelgeschlaeger@studium.uni-hamburg.de

Valeria Magdych
University of Hamburg
valeriia.magdych@studium.uni-hamburg.de

Martin Semmann
University of Hamburg
martin.semmann@uni-hamburg.de

Eva Bittner
University of Hamburg
eva.bittner@uni-hamburg.de

Abstract

This study explores the application of the retrieval-augmented generation architecture for large language models in analyzing citizens' contributions from urban participation. Existing literature highlights the potential of large language models to streamline analytical processes. However, challenges regarding required functions, domain expertise, and transparency remain underexplored. This research addresses these issues through a design science research approach. We identified eleven issues with a systematic literature review and twelve expert interviews, formulated twelve meta-requirements, and derived four design principles on which we developed a web prototype. We evaluated it with 42 experts from a crowdsourcing platform. Our findings demonstrate that retrieval-augmented generation models can enhance efficiency in automated categorization, sentiment analysis, and summarization by focusing the model's attention. However, transparency limitations persist as an ongoing challenge. Our findings contribute to existing knowledge by illustrating how hybrid intelligent systems can improve urban experts' ability to analyze and interpret participation data in smart cities.

Keywords: Citizen Participation, Data Analysis, Retrieval Augmented Generation, Large Language Model, and Design Science Research.

1. Introduction

The ongoing digitalization of urban spaces presents new opportunities for enhancing citizen participation in urban development (Smith & Martín, 2021). Citizen participation is often conducted in response to local tensions and community conflicts due to urbanization (Caragliu & Del Bo, 2022). Digital tools like participation platforms (Royo et al., 2020) and mobile applications are increasingly utilized to expand citizen participation and support their scalability, creating public value (Borchers et al., 2023). However, as these approaches expand, so does the volume of collected citizen contributions. This growth poses significant challenges for local authorities, urban planners, and decision-makers, who face difficulties in analyzing and evaluating large and heterogeneous amounts of textual data (Royo et al., 2020). The literature highlights the critical need for effective methods to process and utilize these vast quantities, as manual evaluation methods are increasingly impractical given the scale of citizen inputs, prompting the exploration of advanced computational approaches (Arana-Catania et al., 2021). One promising approach involves the deployment of Artificial Intelligence (AI), particularly Large Language Models (LLMs), and the newly emerging Retrieval Augmented Generation (RAG) techniques (Shi et al., 2024). RAG extends LLMs with additional contextual data and has shown potential for automating the analysis of domain-specific textual data (Nikishina et al., 2025). While previous studies have explored Machine Learning (ML) methods for analyzing citizen contributions (Cai, 2021; Romberg & Escher, 2023), applying general-purpose LLMs remains underexplored (Williams et al., 2024), which is why we propose the following research question (**RQ**).

RQ: How should RAG-based IT artifacts be designed to support urban experts' analysis of participation data?

We answer the **RQ** with a design-oriented approach (Peffers et al., 2007) and continue with the Related Work, Problem Identification, and Method sections.

URI: https://hdl.handle.net/10125/111679
978-0-9981331-9-5

Afterward, the issues are described in the Identification of the Problem section. Meta-requirements and design principles are elaborated in the Objectives of a Solution section. The development of the IT artifact is specified in the Design and Development section, and in Demonstration and Evaluation, we describe our evaluation. Afterward, we present our findings, discuss them, and then summarize the paper.

2. Related Work

Urban development is shaped by numerous factors that increasingly require integration into city planning. As urban areas and cities grow more complex, decision-making becomes more challenging (Fu et al., 2023). To face that, active citizen engagement, through mechanisms known as citizen participation, is vital and today requires digital tools to increase effectiveness and efficiency (Cai, 2021).

2.1 Citizen Design Science Model

The Citizen Design Science Model integrates the three dimensions of Citizen Science, Citizen Design, and Design Science, serving as a tool to support citizen participation (Mueller et al., 2017). Each dimension is a combination of two areas of citizens, design, and science, as these are crucial and induce each other, as shown in Figure 1 (Torrecilla, 2019).

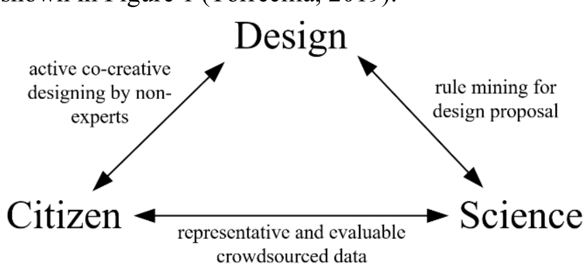

Figure 1. Citizens Design Science Model

Citizen Science involves eliciting citizens' requirements, needs, and ideas, considering the existing contextual and local conditions of projects and the methods of participation employed (Nicolas et al., 2021). This encompasses on-site (analog), digital, and hybrid participation modalities, such as interviews, workshops, and surveys (Mueller et al., 2017). Therefore, Citizen Science addresses how to include and enable participation in a representative and, depending on the objective, valid approach to ensure the elicitation of citizens' contributions (Mueller et al., 2017). Citizen Design determines how citizens actively contribute to the design of a project's vision during participation. This can address the development of buildings, parks, districts, and communities where social interactions, places, and factors like safety, openness, and new living projects are relevant. The possibilities for citizens to contribute their ideas and requirements vary from simple text to 2D and 3D approaches (Munz et al., 2023). In terms of scalability, Citizen Design is influenced by Citizen Science, which defines how citizens can express themselves and impacts Design Science through the analysis of collected data (Mueller et al., 2017). Design Science delineates the methodology of analyzing citizen contributions into actionable insights to support architects and urban planners in developing projects (Altrock, 2022). Empirical evidence from prior research shows that converting these contributions into specific project goals is very complex (Cai, 2021). It underscores the necessity of establishing the requirements for urban initiatives and the technical frameworks for participatory engagement and data analysis before initiating participation activities to ensure the process's feasibility (Repette et al., 2021). Typical data formats are Excel, CSV, JSON, and GEOJSON. Depending on the amount, AI or ML-based approaches are required to support feasible, cost-efficient, and comprehensible data analysis (Borchers et al., 2024), which directly relates to our **RQ**.

2.2 LLMs and RAG for Citizen Participation

The increased application of LLMs promise a paradigm shift in supporting citizen participation and data analysis in urban planning (Gao et al., 2024). LLMs are trained on vast natural language corpora and can be used to perform tasks such as sentiment analysis, topic detection, summarization, and translation with increasing accuracy (Fu et al., 2023). In the context of citizen engagement, LLMs, especially RAG, could support the automatic evaluation of thousands of contributions, assisting experts in grouping and prioritizing issues, identifying key themes, and assessing the overall sentiment of public opinion (Cranefield & Pries-Heje, 2023).

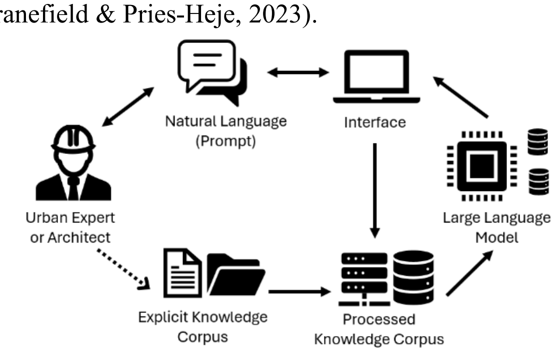

Figure 2. Basic Retrieval Augmented Generation (RAG) Architecture

The RAG architecture may facilitate information control by integrating a domain-specific corpus, such as participatory processes, as part of a processing pipeline, as shown in Figure 2 (Li et al., 2022). Therefore, RAG systems combine the generative power of LLMs with retrieval mechanisms by including a domain-specific knowledge corpus. This must be predefined and could include expert knowledge about analyzing participation data. Additionally, it is possible to incorporate knowledge about the project, such as local names for various places, restaurants, museums, points of interest, and geographic characteristics (Fu et al., 2023). This hybrid approach could enable more effective and efficient data analysis compared to pure LLM-based analysis tools (Kandpal et al., 2023).

3. Method

As we follow a design-oriented approach, we conduct a Design Science Research (DSR) project. DSR is a well-known paradigm in systems science, which can be used to gain knowledge by developing justified IT artifacts based on derived design principles (Hevner et al., 2004). DSR is an iterative approach that leads to new design theories and practical solutions addressing real-world problems (Kuechler & Vaishnavi, 2008). We implemented the DSR activities for this work according to Peffers et al. (2007), as shown in Figure 3.

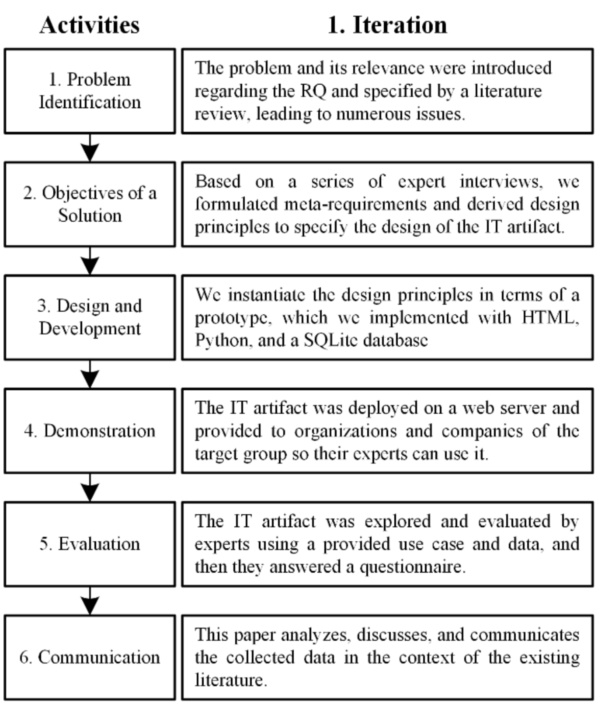

Figure 3. Applied Design Science Research Activities

The introduction section explains and argues the problem and its relevance, leading to the **RQ**. To explicitly define issues, we conducted a systematic literature review (Cooper, 1988). In the second activity, we conducted expert interviews to formulate meta-requirements and derive design principles to specify a solution (Peffers et al., 2007). Afterward, the core artifact is developed and integrated into a web prototype (Shi et al., 2024). We hosted the prototype on a web server to enable the demonstration and the systematic evaluation, which was conducted by experts acquired through a crowd-sourcing platform. The evaluation process adheres to the guidelines by Venable et al. (2016), combining both practical tests and theoretical validation. All collected results were analyzed inductively-deductively, according to Mayring & Fenzl (2019), to validate and elaborate on the derived design principles. Finally, all results are communicated in this paper.

4. Problem Identification

For the identification of the issues (**ISS**), we conducted a literature review according to Cooper (1988) and defined the (1) focus, which the **RQ** and the domain of citizen participation define. The (2) goal is to identify issues. The (3) kind of representation, which is neutral and well-founded, and the (4) coverage, which lies on the pivotal central, as LLM and especially RAG systems, are quite new. Regarding the (5) organization of the review, we focus on conceptual works, and the (6) audience consists of specialized scholars and practitioners. For the literature review, we used the following search string and focused on the databases of the AIS and ScienceDirect.

- *„large language model“ AND urban planning AND citizen participation AND data analysis*

The initial search yielded 171 results, from which we excluded 29 due to a lack of access, 126 due to a lack of relevance based on the abstract, and three due to quality issues, so that 19 papers remained, which revealed several issues. One of these issues (**ISS1**) is the lack of intuitive analytical tools (Bono Rossello et al., 2024; Fürstenau et al., 2021). The literature highlights that current solutions do not offer user-friendly interfaces that enable non-technical users to interpret and explore large datasets. Automated and efficient analysis tools also have a significant deficiency (**ISS2**). Tasks such as categorizing citizen contributions or filtering large volumes of responses remain predominantly manual, laborious, and time-consuming, which hampers scalability (Bhattarai et al., 2023; Hao et al., 2024). Furthermore, the capacity for real-time data assessment (**ISS3**) is absent (Huang et al., 2025; Ludzay & Leible, 2022). The literature emphasizes that timely evaluation of ongoing participation, such as early detection of misinformation or emerging themes, is crucial for adaptive and responsive processes.

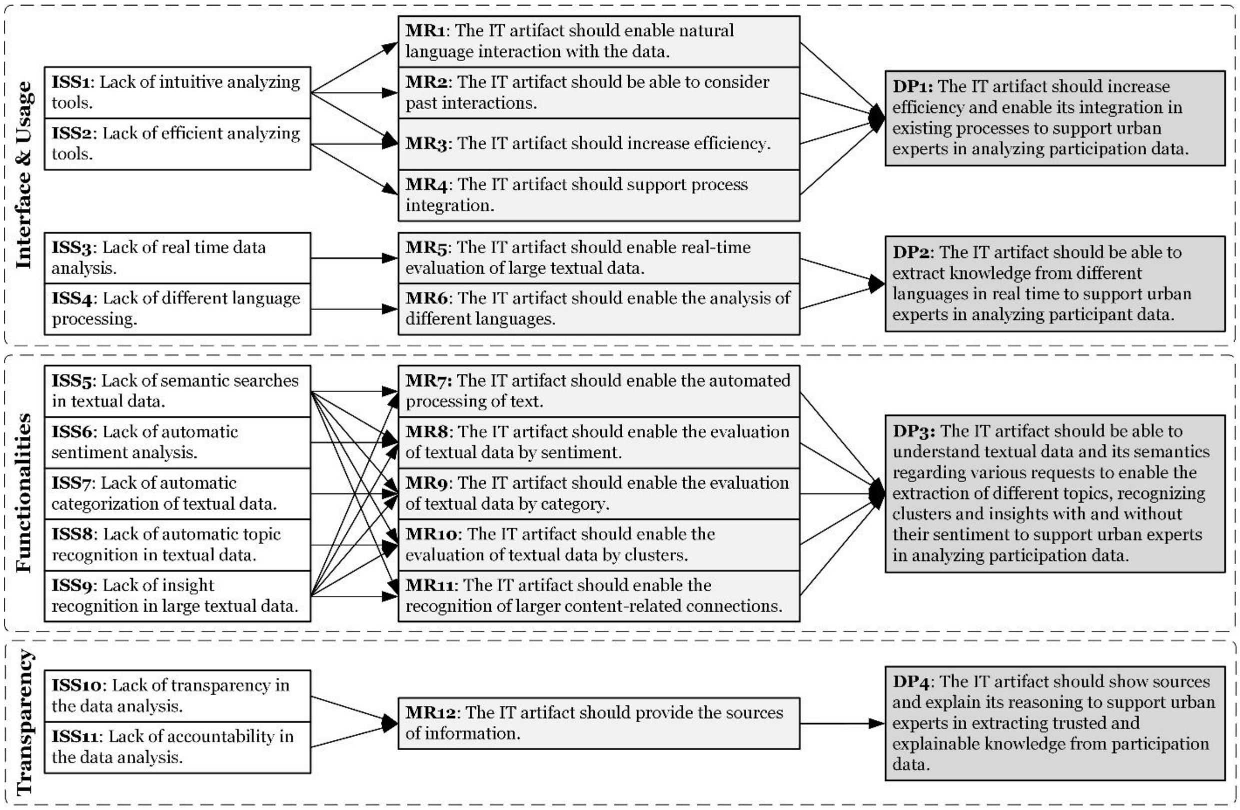

Figure 4. Overview of Issues (ISS), formulated Meta-requirements (MR), and derived Design Principles (DP)

Multilingual processing (**ISS4**) challenges these issues, as urban populations often speak multiple languages (Moats & Ganguly, 2025). There are also significant challenges related to the analysis of textual data, specifically in semantic search and understanding (**ISS5**). Conventional keyword searches prove inadequate when dealing with large datasets filled with indirect descriptions (Cai, 2021). The lack of automatic sentiment analysis (**ISS6**) further hampers understanding of public mood (Cranefield & Pries-Heje, 2023). Additionally, the automatic categorization of textual contributions (**ISS7**) remains an unmet need (Borchers et al., 2024; Moats & Ganguly, 2025). Manual classification is inefficient and inconsistent, but LLM automation could enhance scalability and consistency. Recognizing key themes and patterns (**ISS8**) also remains challenging without automated summaries (Borchers et al., 2024; Graf-Drasch et al., 2023; Hughes et al., 2023). This also applies to extracting meaningful insights from vast textual information (**ISS9**), which is hindered by the volume and complexity of citizen contributions (Becker et al., 2022; De Vreede et al., 2021; Rehm et al., 2022). The issues of transparency and accountability (**ISS10** and **ISS11**) are equally critical. The literature expresses concerns about the opacity, the so-called "black box" nature, of LLMs and other AI models, which complicates the validation and verification of results (Díaz-Rodríguez et al., 2023; Dwivedi et al., 2023; Huang et al., 2025). This lack of explainability diminishes trust, especially when the outcomes influence public or political decisions. Many urban experts worry that the inability to trace how models reach conclusions undermines legitimacy and hampers acceptance by political stakeholders and citizens (Mohammed et al., 2024).

5. Objectives of a Solution

For the formulation of the meta-requirements (**MR**), which are also based on the issues and derivation of design principles (**DP**), we conducted twelve expert interviews, as shown in Table 1, with an average experience (Exp.) of 11.3 years. The phrasing of the design principles was conducted according to Gregor et al. (2020), who define that these should include the recipient, context, aim, and a short rationale. **ISS1** highlights a deficiency in intuitive analysis tools, which necessitates **MR1** to ensure natural interaction, enabling experts to communicate with the system using human language. This aims to mitigate usability barriers and promote an intuitive analysis (E1, E2, and E5). **MR2** is formulated to ensure the IT artifact can consider past interactions to support a continuous conversational flow (E1, E2, and E5). Thus, **MR3** explicitly states that the artifact should enhance operational efficiency by reducing manual effort and streamlining workflows for urban experts, as based on **ISS2**, E2, E3, E4, E7, E8, and E11. **ISS2** also supports **MR4**, which demands process integration in existing workflows (E1, E4, E6, E9, E10, and E11). Together, **MR1** to **MR4** lead to **DP1**: "The IT artifact should increase efficiency and enable integration in existing processes to support urban experts in analyzing participation data."

No.	Exp.	Role	Type
E1	18	Employee	Company
E2	7	Urban Planning Expert	Company
E3	4	Junior Data Analyst	Company
E4	17	Urban Planning Expert	Public Service
E5	21	Communication Expert	Public Service
E6	10	Urban Planning Expert	Public Service
E7	2	Participation Expert	Company
E8	18	Participation Expert	Public Service
E9	20	Participation Expert	Company
E10	4	Participation Expert	Company
E11	6	Developer	Public Service
E12	9	Urban Project Manager	Public Service

Table 1. Overview of Expert Interviews

ISS3 highlights the absence of real-time data analysis capabilities, posing a significant obstacle to making timely decisions during participation. To address this challenge, **MR5** is formulated to facilitate the evaluation of large textual datasets, providing urban experts with immediate insights (E1, E2, E3, E4, E5, E7, E10, E11, and E12). Furthermore, **ISS4** points to the deficiency in processing multiple languages, which restricts the application of IT artifacts in diverse urban contexts. In response, **MR6** specifies linguistic diversity (E3, E6, E9, and E10) and, together with **MR5**, leads to **DP2**: "The IT artifact should be able to extract knowledge from different languages in real-time to support urban experts in analyzing participant data." **ISS5** highlights the lack of semantic searches, which is why we formulate **MR7** to ensure that the IT artifact supports the automated processing of textual data (E1, E2, E9, E10, E11, and E12). Building on this, **ISS6** highlights the lack of automatic sentiment analysis. Therefore, **MR8** emphasizes that the system should enable the evaluation of textual data based on sentiment (E2, E3, E4, and E11). **ISS7** identifies the absence of automated categorization, supporting **MR9** (E1, E2, E3, E7, E8, E9, E10, E11, and E12). **ISS8** points to a deficiency in recognizing underlying topics within texts, which is essential for inductive topic recognition and clustering (**MR10**) to identify key content areas and conceptual groupings (E1, E2, E3, E7, E8, E9, E10, E11, and E12). **ISS9** lacks insight recognition, uncovering latent connections and overarching narratives (**MR11**), and providing context-aware analysis (E1, E3, E5, E9, E10, and E11). In combination, **MR7** to **MR11** led to the derivation of **DP3**: "The IT artifact should be able to understand textual data and its semantics regarding various requests to enable the extraction of different topics, recognizing clusters and insights with and without their sentiment to support urban experts in analyzing participation data." **MR12** development is directly informed by **ISS10** and **ISS11**, which identify critical shortcomings related to transparency and accountability. **ISS10** highlights the lack of clear information regarding the origins of data and analytical outputs. Similarly, **MR12** is formulated based on all experts' opinions, except E3 and E9, to ensure that the IT artifact provides clear sources of information for all data and analyses. This requirement mandates that the system explicitly indicate where data originates, leading to **DP4**.

6. Design and Development

The web prototype was developed using Python, employing the Flask web framework to facilitate its deployment as a web application. The front-end leverages web technologies including HTML, JavaScript, and CSS. The design of the prototype is based on the design principles (cf. Section 5) and the RAG architecture (cf. Section 2.2), utilizing OpenAI's GPT-4 model as its foundation. We chose that model because it is already quite capable and can understand text on a high level. Thus, the explicit knowledge corpus can be free from semantic limitations that a less capable model would have, negatively affecting the usage experience. The prototype considers **DP1**, opting for an intuitive design and chatbot solution, recognizing that this could also be integrated into existing solutions, which we do not have access to. The interactions with data are ensured, as the user can upload one or more datasets, such as Excel, CSV, or txt files. Using RAG ensures real-time evaluation, which, depending on the amount of data, can still require a few seconds (**DP2**). Two chat modes are revealed. Completions API and Assistant API (cf. Figure 5, top left). The first supports the interaction in which text is uploaded and analyzed based on submitted requests, enabling semantic searches and more general interaction (**DP3**). The Assistant API enables specific tasks and is based on an explicit knowledge corpus to support classification, clustering, and sentiment requests (**DP3**). For both approaches, we also extend the prompt with a role description for the model to ensure that it behaves in a professional, objective, and understandable way. Accordingly, the role description, selected parts of the knowledge corpus, and the uploaded data are processed for each prompt to obtain a comprehensible and accurate response (cf. Figure 5). This also considers **DP4** as the role description, which strongly emphasizes the

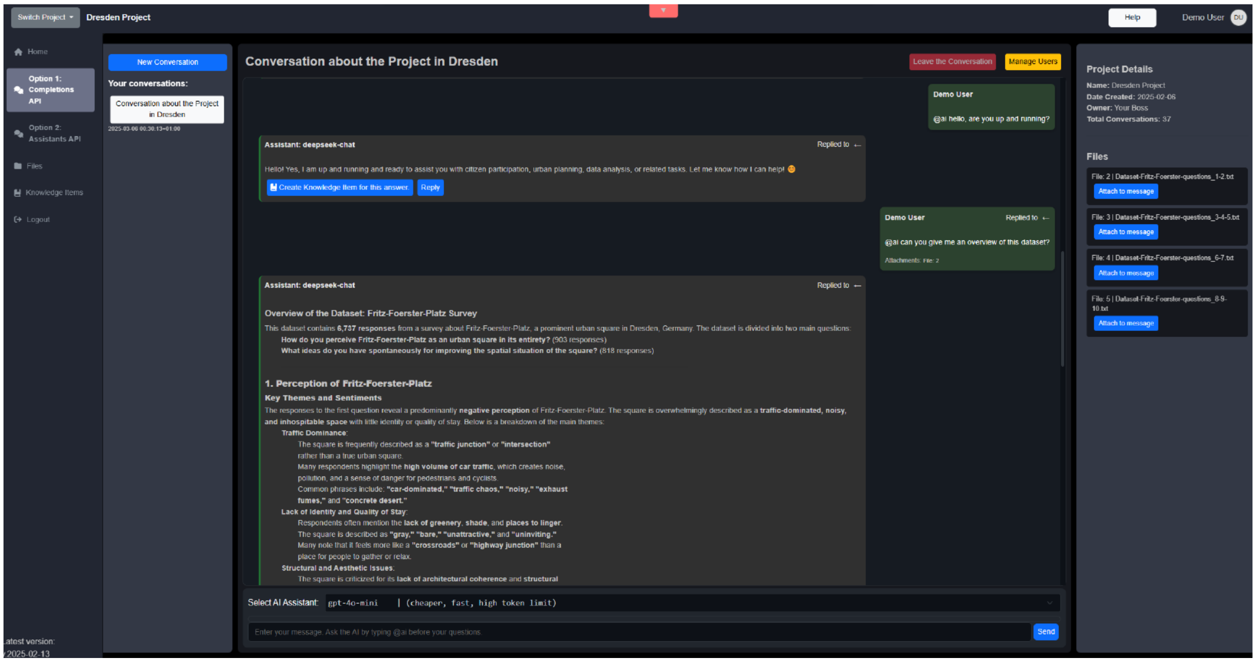

Figure 5. Overview of the Prototype and its Functions

transparency of the statements by communicating examples, the file names, and the lines on which they are based. In cases where queries lead to highly explicit information, this can be implemented in a technically convincing manner. However, this is more difficult for requests that lead to summaries because the sources are more semantic and extensive, which can aggravate transparency, unless reduced to specific parts of citizens' data.

7. Demonstration and Evaluation

For the prototype's demonstration and evaluation, we hosted it on a web server, allowing external experts to inspect and explore it. To receive feedback, we recruited urban experts, data analysts, and architects from the Prolific crowdsourcing platform and compensated them for their participation.

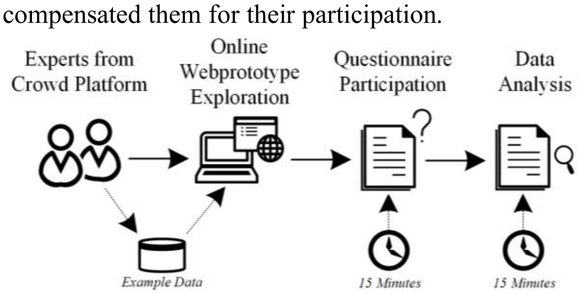

Figure 6. Evaluation Procedure

The demonstration and evaluation were two-fold. All participants received a weblink, a description of a task, and two example data sets from past citizen participation from a larger city, which they should use to achieve and explore the prototype and its function (Venable et al., 2016). All participants had 15 minutes for that. In the second part, all participants completed a survey that required an additional 15 minutes. The questionnaire included statements that participants had to answer according to a 6-point Likert scale (Hartley, 2014) and an open-ended question, where they were asked to describe their experience and provide feedback on the prototype. We acquired 47 participants from Prolific, who successfully explored and answered all questions. We excluded five, as they failed the attention checks that were part of the questionnaire (n = 42). The data analysis was conducted by paraphrasing similar textual statements regarding the design principles (Mayring & Fenzl, 2019).

8. Findings

All participants (n = 42) rated their expertise in citizen participation, urban planning, data analysis, and AI using a 6-point Likert scale, ranging from 1 (lowest) to 6 (highest). The average (AV) expertise score for citizen participation is 4.02 with a standard deviation (SD) of 1.62, 3.90 for urban planning (SD = 1.69), 4.79 for data analysis (SD = 1.35), and 5.05 for AI (SD = 1.15). The above-average values indicate that participants generally considered themselves knowledgeable, particularly in AI, which supports the credibility of the subsequent findings. Regarding **DP1**, the answers show that the prototype was perceived as very efficient overall. The tool was rated as timesaving compared to traditional citizen contribution analysis methods (AV = 4.90, SD = 1.39). Its learning curve was seen as moderate (AV = 4.55, SD = 1.52), though with relatively high variance. Integration into existing workflows (AV = 4.71, SD = 1.11) was approved positively, even if it was evaluated as a standalone prototype. Overall, these results indicate high practical usability for participation experts and urban planners. The idea of combining automated analysis speed with human interpretation to enhance overall workflow received strong agreement (AV = 5.10, SD = 1.14). The application's assistance was also rated positively (AV = 4.62, SD = 1.15). Regarding whether the application caused additional work, confusion, or errors, it received a moderate rating (AV = 4.05, SD = 1.65), indicating that while some disruptions occurred, they were relatively limited overall, approving **DP1**. **DP2** was only noticed to a limited extent and was quickly taken for granted, as translations are widely known and a common functionality of LLMs. With the implementation and evaluation, the technical feasibility was proven, as the sample data also included a small number of other languages. However, no expert explicitly highlights this demand or complains about it.

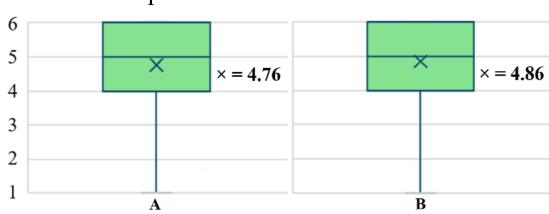

Figure 7. Fulfillment of the functional Requirements (A) and Accuracy of the generated Insights (B)

Regarding **DP3**, the prototype's ability to fulfill the functional requirements for evaluating citizen contributions was rated quite positively (AV = 4.76, SD = 1.25), as shown in Figure 7. Additionally, the AI's responses to the queries received favorable evaluations in terms of summaries, insights, and overall content (AV = 5.10, SD = 1.21). The coherence and structure of these responses were also viewed positively (AV = 4.86, SD = 1.20), indicating that the system provides well-organized and plausible outputs that support meaningful analysis (cf. Figure 7). The final design principle, **DP4**, focuses on providing sources and explanations to ensure trustworthy and explainable insights. Correspondingly, participants agreed that the concept of "Talking to Data," implemented with RAG, is very useful for understanding large datasets (AV = 5.02, SD = 1.14). Additionally, participants rated their trust in the accuracy of the AI's answers as moderately high (AV = 4.64, SD = 1.16). They also found that the application provided a reasonable level of transparency in how outputs were generated, such as explaining conclusions or showing the analysis process (AV = 4.43, SD = 1.30). Nevertheless, just over half of the participants (57%) did not ask the RAG any questions specifically about the origin of its responses. Regarding potential improvements, three participants wanted faster responses from the RAG, while four emphasized the desire for more personalized responses. Overall, the desire for a better user interface and usability was mentioned 19 times, indicating a need for improvement. About 98% of participants agreed that human-AI collaboration brings more benefits than drawbacks. The statement that the application improves collaboration between citizen participation experts and urban planners was also rated positively on average (AV = 4.90, SD = 1.21). Participants also expressed confidence in the application's potential for real-world integration (AV = 5.05, SD = 0.94).

9. Discussion

To answer the **RQ**, "How should LLM-based IT artifacts be designed to support urban experts' analysis of participation data?" we identified 11 issues, formulated twelve meta-requirements, and derived four design principles. Our findings highlight the approval of the design principles and the implemented prototype. The implemented prototype is supportive and can also be integrated into existing processes, as many experts currently evaluate data manually using Excel and do not utilize existing tools, as the expert interviews at the beginning revealed (**DP1**). Besides that, three experts stated that faster AI responses are required, as well as increased usability. **DP2** is feasible and was appreciated, as supported by the literature (Cai, 2021). **DP3** specifies the functions of the prototype and was approved. However, the implementation was also criticized as experts expressed a desire to individualize responses further, not just based on the explicit knowledge base but also on personal preferences. In addition, experts suggested the possibility of giving and collecting feedback to further adjust and personalize the RAG, enabling it to act as a personal agent (Fu et al., 2023). Both lead to the following new design principles.

DP5: The IT artifact should offer settings to urban experts, allowing them to individualize responses and terminology to increase interaction and usage explicitly.

DP6: The IT artifact should provide feedback options for urban experts to collect behavioral data, supporting the development of personalized assistance agents.

Experts also approved **DP4**, but the ideas of how to reach transparency differ, as pure textual descriptions of data sources are limited, which is why visual representations should be examined. In summary, the prototype achieved good results, as reflected in general satisfaction, evaluated quality and performance, as well as usability and interaction design. Additionally, we want to emphasize that 83% of participants preferred collaboration with AI over fully automated analysis, as indicated by a very high level of agreement among participants that AI will play a central role in data-intensive tasks in the future (AV = 5.64, SD = 0.58). The strong consensus, reflected in both the high mean and the low standard deviation, underscores a consistently high perception of AI's potential.

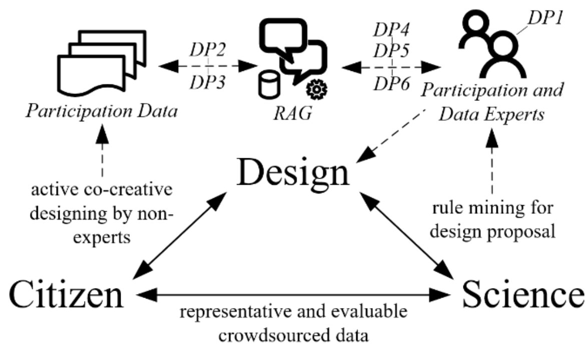

Figure 8. Positioning the Design Principles in the Citizen Design Science Model

We also want to elaborate on our results regarding the citizen design science model (cf. section 2.1). The six design principles are arranged as shown in Figure 8. While **DP2** and **DP3** focus on analyzing data from citizen design, **DP4** to **DP6** specify the interaction between RAG and urban experts, as well as how information can be queried, read, and utilized to extract design for urban projects as part of the model. Rules from design science can be considered, as the interaction with the RAG is up to the experts (Mueller et al., 2017), and operationalizing these rules is only necessary for automating the evaluation, which the experts have spoken out against. With the derived principles and implemented prototypes, we contribute to theory and practice. While the prototype can serve as an explicit orientation for practitioners, including the meta-requirements and issues, the design principles extend the knowledge bases and shift the urban informatics literature in the direction of individualized LLMs, e.g., RAG and personal agents (Cai, 2021). The arrangement of design principles in the context of citizen design science extends their applicability and enhances operational usage and framework development. However, our findings have some limitations. In **DP1**, it remains unclear to what degree efficiency is increasing, which would require comparative studies. These should be considered as an outlook to quantify this in percentage terms in the context of governmental activities, as well as companies and business models that offer data analysis as a service. Additionally, **DP5** and **DP6** should be further investigated in the context of the human-in-the-loop model and participation processes (Wiethof & Bittner, 2022). Moreover, both could also aggravate **DP4** as the RAG responses are becoming increasingly difficult to compare when they are individualized, which should be considered in the second DSR cycle.

10. Conclusion

In this paper, we presented the first cycle of our DSR project, exploring the development of LLM-based IT artifacts by following the RAG architecture to support urban experts in analyzing citizen participation data. This is necessary due to the increasing complexity and volume of citizen contributions that challenge manual analysis. Following the DSR methodology outlined by Peffers et al. (2007), we systematically identified eleven key issues through a structured literature review and expert interviews, formulated twelve meta-requirements, and derived four initial design principles. These were instantiated in the development of a web-prototype, enabling intuitive, real-time, multilingual, and semantically rich data analysis (Shi et al., 2024). The prototype was evaluated in a two-step process by 42 experts in urban planning, AI, and data analysis, whose feedback validated the fulfillment of the core design principles, particularly regarding efficiency, usability, and AI-based functions. Based on their inputs, two additional design principles regarding individualization and feedback mechanisms were derived. Our findings extend both theory and practice in the fields of urban informatics, citizen design science, and AI-driven participatory platforms (Borchers et al., 2024). The proposed design principles not only provide orientation for practitioners developing similar tools but also contribute to the theoretical foundation of AI in citizen participation (Cai, 2021). At the same time, the results highlight ongoing challenges such as measuring efficiency gains quantitatively, advancing individualization without compromising transparency, and integrating human feedback (Romberg & Escher, 2023). Looking ahead, future research should include comparative efficiency studies, further investigation into personalization and explainability trade-offs, and the operationalization of human-in-the-loop models for participatory processes. With this and the findings presented in this paper, an important step toward enabling more effective, efficient, and transparent data analysis through citizen participation is possible.

11. Acknowledgment

We want to express our sincere gratitude to the Federal Ministry of Research, Technology and Space

(Bundesministerium für Forschung, Technologie und Raumfahrt) for supporting the project RESCUE-MATE. The project examines innovative approaches to support civic engagement in urban crisis scenarios. The funding of RESCUE-MATE, with grant number 13N16836, enabled us to carry out this research.

12. References

Altrock, U. (2022). Urban livability in socially disadvantaged neighborhoods: The experience of the German program "socially integrative city." Frontiers of Architectural Research, 11(5), 783–794. https://doi.org/10.1016/j.foar.2021.12.006

Arana-Catania, M., Lier, F.-A. V., Procter, R., Tkachenko, N., He, Y., Zubiaga, A., & Liakata, M. (2021). Citizen Participation and Machine Learning for a Better Democracy. Digital Government: Research and Practice, 2(3), 27:1-27:22. https://doi.org/10.1145/3452118

Becker, F., Siemon, D., & Robra-Bissantz, S. (2022). Smart Participation Design: Prescriptive Knowledge for Bottom-Up Participation. Communications of the Association for Information Systems, 51(1). https://aisel.aisnet.org/cais/vol51/iss1/3

Bhattarai, R., Hassan, L., & Vesa, M. (2023). Participatory Budgeting Case Studies throughout the Ages: A Longitudinal, Thematic, Systematic Literature Study. Hawaii International Conference on System Sciences 2023 (HICSS-56). https://aisel.aisnet.org/hicss-56/dg/emerging_topics_in_e-gov/4

Bono Rossello, N., Simonofski, A., Clarinval, A., & Castiaux, A. (2024). A Typology for AI-enhanced Online Ideation: Application to Digital Participation Platforms. Hawaii International Conference on System Sciences 2024 (HICSS-57). https://aisel.aisnet.org/hicss-57/dg/ai/3

Borchers, M., Cao, T.-B., Tavanapour, N., & Bittner, E. (2024, June 14). Designing AI-based Systems to support the Analysis of Citizens' Inputs from E-Participation. ECIS 2024 Proceedings. 7. European Conference on Information Systems (ECIS). https://aisel.aisnet.org/ecis2024/track23_designresearch/track23_designresearch/7

Borchers, M., Tavanapour, N., & Bittner, E. (2023, January 3). Designing Mobile Applications for Citizen Participation in Urban Planning. Proceedings of the 56th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences, USA. https://hdl.handle.net/10125/102682

Cai, M. (2021). Natural language processing for urban research: A systematic review. Heliyon, 7(3), 1–8. https://doi.org/10.1016/j.heliyon.2021.e06322

Caragliu, A., & Del Bo, C. F. (2022). Smart cities and urban inequality. Regional Studies, 56(7), 1097–1112. https://doi.org/10.1080/00343404.2021.1984421

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. Knowledge in Society, 1(1), 104. https://doi.org/10.1007/BF03177550

Cranefield, J., & Pries-Heje, J. (2023). Boundary Management Strategies for Leading Digital Transformation in Smart Cities. MIS Quarterly Executive, 22(2). https://aisel.aisnet.org/misqe/vol22/iss2/4

De Vreede, T., Vreede, G.-J. D., & Alawi, N. (2021). Achieving Success in Community Crowdsourcing: Lessons from the Field. Hawaii International Conference on System Sciences 2021 (HICSS-54). https://aisel.aisnet.org/hicss-54/cl/social_and_psychological_perspectives/2

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. Information Fusion, 99, 101896. https://doi.org/10.1016/j.inffus.2023.101896

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fu, J., Han, H., Su, X., & Fan, C. (2023). Towards Human-AI Collaborative Urban Science Research Enabled by Pre-trained Large Language Models (No. arXiv:2305.11418). arXiv. http://arxiv.org/abs/2305.11418

Fürstenau, D., Morelli, F., Meindl, K., Schulte-Althoff, M., & Rabe, J. (2021). A Social Citizen Dashboard for Participatory Urban Planning in Berlin: Prototype and Evaluation. Hawaii International Conference on System Sciences 2021 (HICSS-54). https://aisel.aisnet.org/hicss-54/dg/governance/2

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval-Augmented Generation for Large Language Models: A Survey (No. arXiv:2312.10997). arXiv. https://doi.org/10.48550/arXiv.2312.10997

Graf-Drasch, V., Keller, R., Meindl, O., & Röhrich, F. (2023). The Design of Citizen-Centric Green IS in Sustainable Smart Districts. Business & Information Systems Engineering, 65(5), 521–538.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). The Anatomy of a Design Principle. Journal of the Association for Information Systems, 21, 1622–1652. https://doi.org/10.17705/1jais.00649

Hao, H., Wang, Y., & Chen, J. (2024). Empowering Scenario Planning with Artificial Intelligence: A Perspective on Building Smart and Resilient Cities. Engineering, 43, 272–283. https://doi.org/10.1016/j.eng.2024.06.012

Hartley, J. (2014). Some thoughts on Likert-type scales. International Journal of Clinical and Health Psychology, 14(1), 83–86. https://doi.org/10.1016/S1697-2600(14)70040-7

Hevner, A., R, A., March, S., T, S., Park, Park, J., Ram, & Sudha. (2004). Design Science in Information Systems Research. Management Information Systems Quarterly, 28, 75.

Huang, J., Bibri, S. E., & Keel, P. (2025). Generative spatial artificial intelligence for sustainable smart cities: A pioneering large flow model for urban digital twin. Environmental Science and Ecotechnology, 24, 100526. https://doi.org/10.1016/j.ese.2025.100526

Hughes, M., Dimitrakopoulou, D., Rojas, M., & Diby, S. (2023). Digital Civic Sensemaking: Computer-Supported Participatory Sensemaking of Nuanced, Experience-Based Dialogue. Hawaii International Conference on System Sciences 2023 (HICSS-56). https://aisel.aisnet.org/hicss-56/dsm/mediated_conversation/4

Kandpal, N., Deng, H., Roberts, A., Wallace, E., & Raffel, C. (2023). Large Language Models Struggle to Learn Long-Tail Knowledge. Proceedings of the 40th International Conference on Machine Learning, 15696–15707. https://proceedings.mlr.press/v202/kandpal23a.html

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. European Journal of Information Systems, 17(5), 489–504. https://doi.org/10.1057/ejis.2008.40

Li, H., Su, Y., Cai, D., Wang, Y., & Liu, L. (2022). A Survey on Retrieval-Augmented Text Generation (No. arXiv:2202.01110). arXiv. https://doi.org/10.48550/arXiv.2202.01110

Ludzay, M., & Leible, S. (2022). A Bottom-up E-Participation Process: Empowering Citizens to Innovate the Public Administration and its Sphere of Influence. Wirtschaftsinformatik 2022 Proceedings. https://aisel.aisnet.org/wi2022/student_track/student_track/36

Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Eds.), Handbuch Methoden der empirischen Sozialforschung (pp. 633–648). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_42

Moats, D., & Ganguly, C. (2025). Bringing AI participation down to scale. Patterns, 6(5), 101241. https://doi.org/10.1016/j.patter.2025.101241

Mohammed, A., Kutar, M., & Albakri, M. (2024). Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda. UK Academy for Information Systems Conference Proceedings 2024. https://aisel.aisnet.org/ukais2024/14

Mueller, J., Lu, H., Chirkin, A., Klein, B., & Schmitt, G. (2017). Citizen Design Science: A strategy for crowd-creative urban design. Cities, 72, 181–188. https://doi.org/10.1016/j.cities.2017.08.018

Munz, L., Kauzlaric, M., Mosimann, M., Fehlmann, A., Martius, O., & Zischg, A. P. (2023). Participatory development of storymaps to visualize the spatiotemporal dynamics and impacts of extreme flood events for disaster preparedness. International Journal of Disaster Risk Reduction, 98, 104039. https://doi.org/10.1016/j.ijdrr.2023.104039

Nicolas, C., Kim, J., & Chi, S. (2021). Natural language processing-based characterization of top-down communication in smart cities for enhancing citizen alignment. Sustainable Cities and Society, 66, 102674. https://doi.org/10.1016/j.scs.2020.102674

Nikishina, I., Sevgili, Ö., Li, M. M., Biemann, C., & Semmann, M. (2025). Creating a Taxonomy for Retrieval Augmented Generation Applications (No. arXiv:2408.02854). arXiv. https://doi.org/10.48550/arXiv.2408.02854

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. https://doi.org/10.2753/MIS0742-1222240302

Rehm, S.-V., Mcloughlin, S., & Donnellan, B. (2022). How To Cope With The Dynamics Of Urban Sustainability: Urban Experimentation Platforms As Tools For Adaptive Policy-Making. ECIS 2022 Research-in-Progress Papers. https://aisel.aisnet.org/ecis2022_rip/61

Repette, P., Sabatini-Marques, J., Yigitcanlar, T., Sell, D., & Costa, E. (2021). The Evolution of City-as-a-Platform: Smart Urban Development Governance with Collective Knowledge-Based Platform Urbanism. Land, 10(1), Article 1. https://doi.org/10.3390/land10010033

Romberg, J., & Escher, T. (2023). Making Sense of Citizens' Input through Artificial Intelligence: A Review of Methods for Computational Text Analysis to Support the Evaluation of Contributions in Public Participation. Digital Government: Research and Practice. https://doi.org/10.1145/3603254

Royo, S., Pina, V., & Garcia-Rayado, J. (2020). Decide Madrid: A Critical Analysis of an Award-Winning e-Participation Initiative. Sustainability, 12(4), Article 4. https://doi.org/10.3390/su12041674

Shi, Y., Tan, Q., Wu, X., Zhong, S., Zhou, K., & Liu, N. (2024). Retrieval-enhanced Knowledge Editing in Language Models for Multi-Hop Question Answering. Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, 2056–2066. https://doi.org/10.1145/3627673.3679722

Smith, A., & Martín, P. P. (2021). Going Beyond the Smart City? Implementing Technopolitical Platforms for Urban Democracy in Madrid and Barcelona. Journal of Urban Technology, 28(1–2), 311–330. https://doi.org/10.1080/10630732.2020.1786337

Torrecilla, F. J. S. (2019). Citizen Science Design the Role of Product Designers: A Study Case, Insectivoros. DS 95: Proceedings of the 21st International Conference on Engineering and Product Design Education (E&PDE 2019), University of Strathclyde, Glasgow. 12th -13th September 2019. 21st International Conference on Engineering & Product Design Education (E&PDE 2019). https://doi.org/10.35199/epde2019.19

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. European Journal of Information Systems, 25(1), 77–89. https://doi.org/10.1057/ejis.2014.36

Wiethof, C., & Bittner, E. (2022). Toward a Hybrid Intelligence System in Customer Service: Collaborative Learning of Human and AI.

Williams, S., Beery, S., Conley, C., Evans, M. L., Garces, S., Gordon, E., Jacob, N., & Medina, E. (2024). People-Powered Gen AI: Collaborating with Generative AI for Civic Engagement. An MIT Exploration of Generative AI. https://doi.org/10.21428/e4baedd9.f78710e6