Designing a Large Language Model Based Open Data Assistant for Effective Use

Till Carlo Schelhorn[1(✉)], Ulrich Gnewuch[2], and Alexander Maedche[1]

[1] Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
{till.schelhorn,alexander.maedche}@kit.edu
[2] University of Passau, Passau, Germany
ulrich.gnewuch@uni-passau.de

Abstract. Open data is widely recognized for its potential positive impact on society and economy. However, many open data sets remain underutilized because users, such as civil servants and citizens, lack the necessary technical and analytical skills. Additionally, existing open data portals often fall short of providing user-friendly access to data. Although conversational agents equipped with Large Language Models have emerged as a promising solution to address these challenges, it is unclear how to design Large Language Model based open data assistants that allow users to formulate their information needs in natural language and ultimately use open data effectively. To address this gap, we undertake a Design Science Research project guided by the theory of effective use. In this first cycle of the project, we present meta-requirements and propose initial design principles on how to design a Large Language Model based open data assistant for effective use. Subsequently, we instantiate our principles in a prototype and evaluate it in a focus group with experts from a medium-sized German city. Our results contribute design knowledge in the form of design principles for open data assistants and inform future design cycles of our Design Science Research project.

Keywords: Open Data · Conversational Agents · Large Language Models · Theory of Effective Use · Design Science Research

1 Introduction

”Everyone has the right to [...] seek, receive and impart information” [34, Art. 19]. Freedom of information is recognized as a fundamental right in democratic states. Therefore, legislative frameworks have been established in many countries to speed up the opening of data to the general public, e.g. in the EU [8] or the US [25]. These open data initiatives are motivated by the objective of increasing openness and transparency, enabling participation in the democratic process, and thus strengthening democracy [22]. Beyond its positive impact on society, open data can also benefit the economy by laying the foundation for innovative digital services and the development of new business models [29]. The European Commission estimates that open data could boost the European economy by €40 billion annually [7].

Open data is typically published on open data portals on different levels, e.g. on the level of cities, municipalities, states, or state unions such as the EU [22]. As an example, the open data portal of the EU contains over 1,600,000 data sets [9]. Yet, open data is not limited to government data but includes all data "that can be used, studied, and modified without restriction" [24, p. 1]. In order to leverage the potential of open data, it must also be used. Open data usage refers to "the activity that a person or an organization conducts to view, understand, analyze, visualize or in other ways use a dataset that has been provided to the public" [44, p. 429].

Although the benefits of open data are widely acknowledged, various challenges impede users from using and leveraging its potential [1, 5, 26, 43]. First, the user base of open data portals is diverse, encompassing citizens, journalists, activists, researchers, employees of private companies, and civil servants [32]. This diversity results in a wide range of skill sets and domain knowledge among users. Thus, many users lack the technical and analytical skills or the domain knowledge required to effectively utilize open data portals [38]. Second, amplifying these issues, open data portals often have complex user interfaces with challenging navigation structures and inadequate search capabilities, hindering the identification of relevant datasets [26]. Consequently, many datasets available on open data portals remain unused by the public [27]. Nevertheless, there is a demand for public information captured by open data. Public offices increasingly receive requests for public information leading civil servants to retrieve relevant open data [11]. However, this creates additional workload and leads to a bottleneck. For example, FragDenStaat, a German website facilitating citizens' requests for public information, reports an average response time of 43 days for these requests [10].

Amidst these challenges, conversational agents (CAs) emerge as a potential solution providing accessible interfaces through natural language [23]. According to the theory of effective use [3], transparent interaction is a key dimension for effectively using Information Systems (IS). Indeed, Ruoff et al. [30] demonstrate that conversational interfaces can significantly enhance the effective use of IS, particularly in the context of dashboards, by achieving heightened levels of transparent interaction. Moreover, Burton-Jones and Grange [3] identify representational fidelity, and informed action as additional dimensions influencing the efficiency and effectiveness of IS usage.

In addition, prototypes of "open data assistants" have been introduced in literature to facilitate access and utilization of open data [13, 18, 37]. The ascent of Large Language Models (LLMs) has further expanded possibilities, as their extensive capabilities in natural language comprehension can be leveraged to create powerful CAs [39]. Recent literature has showcased the potential of LLMs for reasoning and solving complex tasks through self-generated chain-of-thought [40]. Additionally, attention has been given to enhancing LLMs' capabilities for answering domain-specific questions by utilizing various external knowledge sources [21]. This effort extends to structured data through the generation and execution of SQL queries [16, 28]. These advancements demonstrate promise for the development of open data assistants that not only assist users in understanding the data but also provide insight into how answers to their questions are generated. However, there is limited research on the application of LLM-powered CAs for open data. This raises the following research question: *How to design an LLM-based open data assistant for effective use?*

To address this question, our research follows the Design Science Research (DSR) paradigm proposed by Hevner et al. [12]. DSR aims to solve a real-world problem by suggesting an innovative solution. We conduct our DSR project in a real-world environment in cooperation with a medium-sized German city that already runs an Open Data portal and is faced with the challenges introduced above. Thus, from a stakeholder perspective we include both civil servants and citizens. This paper encompasses the first design cycle of our larger DSR project. Our research is grounded in the theory of effective use (TEU) as kernel theory [3]. Guided by interviews, focus groups, and a review of existing literature on the barriers of open data, we identify several meta-requirements for an LLM-based open data assistant. In response to these requirements, we formulate three initial design principles to guide the implementation of our artifact. The artifact, along with its underlying design principles, undergoes evaluation in a focus group comprising civil servants from our partner city. Through this, we successfully demonstrate how an LLM-based open data assistant can support the utilization of open data portals by users with varying technical and analytical skills. This contribution extends design knowledge in the form of design principles for open data assistants, providing actionable guidance for the creation of such assistants and offering valuable insights for future design cycles.

2 Foundations and Related Work

2.1 Conversational Agents and Large Language Models

Conversational User Interfaces (CUIs) enable users to interact with IS in written or spoken natural language [23]. “Conversational” refers to all types of spoken interaction supporting the use of spontaneous language and often displaying human-like characteristics [23]. Research on CUIs has a long history with ELIZA being the first chatbot developed in the 1960s [41]. Since then there has been a large interest in research on CUIs in the form of chatbots and CAs in different contexts [6]. Recent literature has investigated the influence of CAs on effective use in the context of dashboards showing that CAs can complement traditional graphical user interfaces [30].

Large Language Models (LLMs) have rapidly reshaped the landscape of natural language processing (NLP) and CAs [39]. The self-attention mechanism of the underlying transformer architecture enables the models to capture and understand the relationship between different words of an input sequence [35]. ChatGPT made LLMs available to the general public and became the fastest growing consumer application in history [14]. Since then, LLMs have shown remarkable capabilities in code generation tasks, such as generating SQL queries, showcasing their versatility beyond conventional language understanding [28].

To address challenges like hallucinations, research has extended LLM functionality with Retrieval-Augmented Generation (RAG) integrating the models with external knowledge sources and enabling them to answer domain-specific questions [21]. Additionally, the explicit prompting of chain-of-thought has emerged as a strategy to guide LLMs in solving complex tasks [40]. These developments have facilitated the implementation of LLM-powered agents with extensive capabilities consisting of a reasoning engine and access to different external tools and knowledge bases [17, 42].

2.2 Theory of Effective Use

To achieve maximum benefits from IS, they need to be used effectively [3]. Burton-Jones and Grange define effective use as “using a system in a way that helps attain the goals for using the system” [3, p. 633]. They define a hierarchy of three dimensions that influence the effective and efficient use of IS: (1) transparent interaction, (2) representational fidelity, and (3) informed action [3]. Each dimension is a requirement for the higher-level dimension. A user needs to be able to transparently interact with an IS to obtain a faithful representation of the underlying domain (representational fidelity) in order to be able to take an informed action [3]. Transparent interaction is defined as “accessing the system’s representations unimpeded by its surface and physical structures” [3, p. 642]. The surface structure refers to the user interface of the IS and the physical structure to devices in the physical world that are used to interact with the IS (e.g. mouse and keyboard).

Burton-Jones and Grange state two major types of actions to improve effective use: adaption and learning. Adaption is “any action a user takes to improve (1) a system’s representation of the domain of interest; or (2) his or her access to them” [3, p. 644]. They further define learning actions as “any action a user takes to learn (1) the system (its representations, or its surface or physical structure); (2) the domain it represents; (3) the extent to which it faithfully represents the domain (i.e., its fidelity); or (4) how to leverage representations obtained from the system” [3, p. 644].

2.3 Related Work

Several conversational assistants for open data have been proposed in the literature, including for the city of Aragon [13], Austria [18], and Shanghai [37]. However, del Hoyo-Alonso et al. [13] conclude that traditional intent-based CAs can assist users interacting with open data, but require extensive implementation effort to satisfy all possible user questions. They suggest extending the capabilities of open data assistants with LLMs. Yet, they also raise concerns regarding the generation of hallucinations by these models and stress the need to ensure the validity of the information provided to the user [13]. In addition, initial practical solutions are also offered: ZurichGPT, an LLM-based agent for exploring open data of the city of Zurich, was published recently [4]. They focus on providing the sources of retrieved information with the generated answer. However, they also explicitly state that users should always check information on the official website of the city [4].

Nevertheless, current research rapidly develops the capabilities of LLMs further. The explicit prompting of chain-of-thought [40] and provisioning of external knowledge [21] enables the development of LLM-powered agents with extensive problem-solving capabilities [17, 42]. These could answer questions more reliably and even answer questions regarding complex tasks. For example, Bran et al. developed a chemistry agent by augmenting an LLM with chemistry tools enabling it to automate several chemistry-related tasks [2]. However, limited research exists on theory-guided design for LLM-based agents. Additionally, the works on open data assistants we found either excluded users from their evaluation or the assistants lacked LLM capabilities. We therefore argue that there is limited knowledge of what design principles should guide the development of

LLM-based agents. We address this gap using a DSR approach to design and evaluate an LLM-based open data assistant for effective use in a real-world environment.

3 Methodology

Our research is guided by the DSR paradigm [12] and follows the DSR method described by Kuechler and Vaishnavi [19]. This paper encompasses the first design cycle of our DSR project. To understand the challenges faced with open data we undertook an initial literature review, conducted two focus groups with citizens, and interviewed civil servants responsible for the open data portals of five different German cities. Based on the identified issues related to the use of open data we derived meta-requirements and proposed three design principles. Our research is grounded in the TEU as the kernel theory [3]. To complement this, we draw prescriptive knowledge from existing literature. Specifically, Ruoff et al. [30] propose the adaption of dashboards with conversational interfaces achieving higher levels of transparent interaction leading to enhanced efficiency and effectiveness. Furthermore, we rely on existing knowledge on the creation of LLM-based agents [17, 42] and providing them with access to external knowledge sources [21]. We instantiated our design principles in a prototype, an LLM-based open data assistant, and equipped it with access to open data from our partner city. The city is a medium-sized city in the south of Germany. It inhabits around 250.000 people and employs over 4.000 people in its city administration. Furthermore, in 2020 the city started a large smart city project sponsored by the German Federal Ministry for Housing, Urban Development and Building. This showcases its commitment to innovation and technological advancements, making it a promising candidate for collaboration on our DSR project.

Venable et al. [36] suggest multiple formative and summative evaluation episodes for DSR. In our first cycle, we undertook a formative and qualitative evaluation to evaluate our artifact. This evaluation occurred within an expert focus group comprised of eight civil servants (ranging from 28 to 57 years old, with 2 females and 6 males). The participants are all stakeholders in the open data portal of the city with diverse technical skills and domain knowledge. They serve various roles, either as active users of the portal or as contributors to its development and content.

Moving forward, our research plan includes another design cycle informed by the insights gained from the initial evaluation. This iterative process aims to refine both the design principles and the implemented artifact. In the subsequent evaluation, we intend to conduct a summative quantitative evaluation episode to shed light on the impact of our refined design principles on the effective use of our open data assistant.

4 Designing the LLM-Based Open Data Assistant

4.1 Awareness of Problem

In the upcoming section, we will address the issues identified in existing literature on open data usage, as well as insights obtained from our focus groups with citizens and interviews with civil servants. Our approach involves aligning the findings from literature with the statements gathered during the focus groups and interviews. Subsequently, we derive three meta-requirements informed by our kernel theory TEU (see Sect. 2.2).

First of all, open data portals serve various users with a wide range of different skill sets [32]. Many lack the technical and analytical skills or the domain knowledge to effectively use open data [38]. Furthermore, open data portals often come with poor usability and do not "take the user's perspective into account" [15, p. 256]. The participants of our focus groups criticized the insufficient search capabilities of open data portals and struggled to identify "what data is available and where to find it". This issue is reinforced by the large amount of heterogeneous open data initiatives and portals available [1]. Additionally, open data portals often contain raw data and information is fragmented across different files [43]. The participants found that "downloading and analyzing CSV files is not feasible" for them. Even though some data was presented with visualizations and reports, finding relevant information was still challenging for the participants due to multiple layers of menus and sub-menus. Lastly, poor (meta-) data quality further impedes users trying to find relevant information [31]. The participants of our focus group complained about the "confusing and inconsistent labeling of data" that makes the identification of relevant information unnecessarily difficult. Our findings were confirmed through the conducted interviews. According to insights gathered from civil servants, open data portals primarily serve as a resource for city internal sources. Despite the availability of visualizations and reporting tools on numerous open data portals, these features often go underutilized due to their perceived complexity. Consequently, many users, encompassing citizens, researchers, and journalists, prefer to directly request information from the respective departments within the city.

To derive meta-requirements (MRs) for a system solving these identified issues we draw existing knowledge from the theory of effective use (TEU) [3]. To effectively use an IS, the user must have unimpeded access to the system's representations [3]. We showed that users of open portals often struggle to identify relevant data and extract information thus not being able to interact transparently with the system. One approach to improve the transparent interaction is to adapt the system's surface structure (e.g. its user interface) [3]. Natural language can be used to interact more naturally with open data portals independent of the user's skills and the interface of the portal [30]. Therefore, we propose our first MR:

MR1: The system should allow users to ask questions about and interact with open data using natural language to help them achieve higher levels of transparent interaction.

The second level in the hierarchy of TEU is representational fidelity, the extent to which a user can obtain a faithful representation of the underlying domain by the system [3]. Learning the fidelity of the system's representation is one way to improve this. Enabling the user to comprehend how the system concludes its answer enables them to recognize if the system faithfully represents the underlying domain, e.g. gives a correct answer. Therefore, when obtaining an answer from the system based on open data, being presented with the steps of how the system came up with this answer helps the user to achieve higher levels of representational fidelity. We propose our second MR:

MR2: The system should augment its responses with information about its internal reasoning process to help users achieve higher levels of representational fidelity.

The final dimension of TEU is the informed action the user performs based on the information obtained through the system's representations [3]. Users accessing open data usually have some information needs they want to satisfy. This information is then used to make data-driven decisions, may it be in their private or professional life. Providing false information would hinder the user from taking such an informed action. Therefore, we propose our third MR as follows:

MR3: The system should avoid provisioning responses that contain false or fabricated information in order to enable users to take informed action.

4.2 Suggestion

To address these MRs we formulate three initial DPs. We draw from existing prescriptive knowledge regarding the design of CAs for effective use and LLM-based agents. To enable the user to use natural language when interacting with the assistant (MR1) requires a CA [30]. Serving a large number of different requests by users not knowing the domain-specific terminology the CA should have extensive language comprehension capabilities. Del Hoyo-Alonso et al. [13] therefore suggest providing the assistant with an LLM-powered agent. Thus, to address our first MR we formulate our first DP:

DP1: Provide the open data assistant with an LLM-based agent to increase its language comprehension and generation capabilities.

Enabling the user to understand the internal reasoning process of the assistant is supposed to achieve higher levels of representational fidelity (MR2). LLM-powered agents generate a sequence of actions. Based on the observation resulting from the previous action the agent decides which action to take next until it solves their task. This sequence of actions is often called chain-of-thought [40]. Providing the user with this chain-of-thought could enable them to understand the agent's reasoning process. We therefore address our second MR by formulating the following DP:

DP2: Enable the open data assistant to present the chain-of-thought the conversational agent generates to increase transparency of the system for the user. To reduce hallucinations and provisioning of false answers (MR3) by LLMs, RAG proved as a viable strategy [21]. Providing the LLM-based agent with access to data available on the open data portal and explicitly prompting it to base its answer on this data achieves this. Therefore, we formulate our last DP:

DP3: Provide the open data assistant with the capabilities to access available open data and restrict it to only use this data for generating responses.

4.3 Development

As the next step of our design process, we instantiated the design principles in a prototype to evaluate the proposed design. We developed an LLM-based open data assistant using two high-level Python frameworks: Langchain [20] for the implementation of the LLM agent and Streamlit [33] for the web-based interface of the application. As LLM we relied on the OpenAI API using the GPT-3.5 Turbo foundational model. Figure 1 shows the architecture of our prototype system.

The prototype offers a CA that enables the user to chat with the open data assistant through natural language (DP1). The users' input is forwarded to the Langchain agent

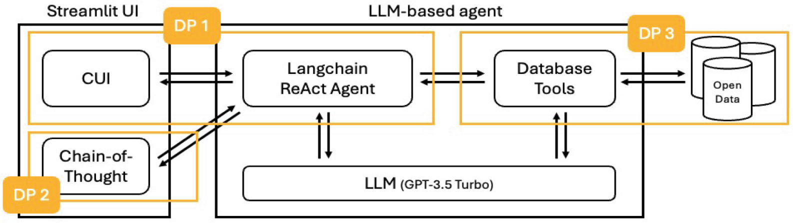
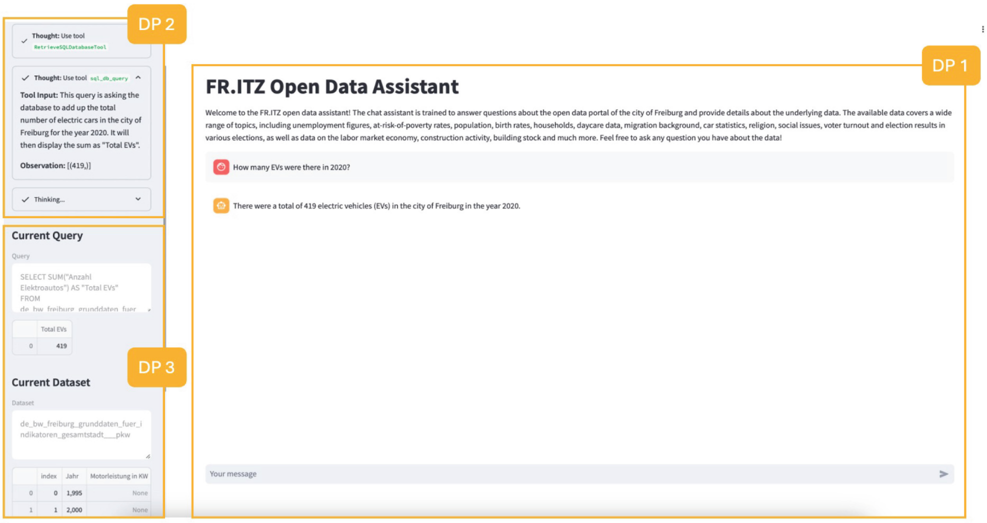


Fig. 1. Overview of the schematic architecture of our prototypical implementation

running in the back. We followed the ReAct approach [42] for the design of the agent. It is equipped with access to an SQL database comprising 42 diverse datasets sourced from the open data portal of our partner city (DP3). Four different tools enable the agent to interact with the database: (1) retrieving the five most relevant datasets for the users' input, (2) querying the SQL schema for a specific dataset, (3) assessing the syntactical and semantic correctness of a provided SQL query, and (4) executing a SQL query and retrieving the resulting data. Upon receiving user input, the agent dynamically selects the appropriate tool and input. Subsequently, based on the observed response from the chosen tool, the agent determines whether to provide an answer directly or invoke another tool. This chain-of-thought (tool, input, and observation) is transparently displayed in the Streamlit UI throughout the interaction (DP2). Upon completion of the sequence of actions, the UI provides a comprehensive display of executed SQL queries and relevant tables employed during the process. This ensures transparency and traceability of the agent's decision-making process. Figure 2 shows the web-based Streamlit UI with an exemplary user question.



✓ **Thought:** Use tool `RetrieveSQLDatabaseTool`

✓ **Thought:** Use tool `sql_db_query`

Tool Input: This query is asking the database to add up the total number of electric cars in the city of Freiburg for the year 2020. It will then display the sum as "Total EVs".

Observation: [(419,)]

Current Query

```
SELECT SUM("Anzahl Elektroautos") AS "Total EVs"
FROM
de_bw_freiburg_grunddaten_fuer
```

	Total EVs
0	419

Current Dataset

de_bw_freiburg_grunddaten_fuer_indikatoren_gesamtstadt___pkw

	index	Jahr	Motorleistung in KW
0	0	1,995	None
1	1	2,000	None

FR.ITZ Open Data Assistant

Welcome to the FR.ITZ open data assistant! The chat assistant is trained to answer questions about the open data portal of the city of Freiburg and provide details about the underlying data. The available data covers a wide range of topics, including unemployment figures, at-risk-of-poverty rates, population, birth rates, households, daycare data, migration background, car statistics, religion, social issues, voter turnout and election results in various elections, as well as data on the labor market economy, construction activity, building stock and much more. Feel free to ask any question you have about the data!

How many EVs were there in 2020?

There were a total of 419 electric vehicles (EVs) in the city of Freiburg in the year 2020.

Fig. 2. Screenshot of the prototypical implementation of our LLM-based open data assistant answering an exemplary user question

The UI features a chat window that provides a brief description of the open data assistant. Upon sending a message, the agent's reasoning process is revealed on the left, showcasing the chain of thought. Further details about specific actions can be accessed by expanding the individual thoughts, revealing natural language descriptions and the observed return values. When data is queried from the knowledge base, both the query and results are presented below. The complete dataset is also visible, with an option for the user to expand the table view for more in-depth examination.

5 Evaluation

5.1 Method

In collaboration with eight experts from our partner city, we organized a focus group to evaluate our prototype and conduct a SWOT analysis. The expert group comprised individuals from various departments, namely (1) two IT department employees responsible for developing the city's open data portal, (2) two employees from different specialized city administration departments, and (3) four employees from the statistics department, tasked with providing reports and addressing data-specific inquiries from both internal city departments and external requests. These participants, being stakeholders of the open data portal, represent diverse perspectives and skill levels.

Our primary objective was to gather feedback from both technical and nontechnical users of the open data portal. The IT department, characterized by high technical proficiency, primarily focuses on data provisioning through the portal. Contrarily, the specialized departments represent the data consumer side, possessing significant domain knowledge but limited technical and analytical capabilities. Acting as intermediaries, the statistics department combines technical and analytical skills with domain knowledge, preparing reports, data, and responses for various city administration departments and external entities, including citizens, journalists, and researchers.

Following a brief introduction, we presented our prototype to the experts, showcasing examples and explaining distinct design features. Subsequently, each expert accessed the system from their respective devices, enabling them to explore the prototype. We then gathered feedback in a discussion to understand their opinions on the prototype. To further analyze the gathered insights, we conducted a SWOT analysis wherein the participants documented their comments, later categorized into the SWOT dimensions. In the next section, we will discuss the results of this evaluation, highlighting the identified strengths, weaknesses, opportunities, and threats.

5.2 Results

The results of our focus group evaluation indicate an overall positive reception of our prototype. The participants acknowledged the usefulness and utility of the system for both the internal use of the city's civil servants and citizens. However, there are still some issues where data is misinterpreted by the open data assistant. Nevertheless, the experts in our focus group praised our approach and gave valuable feedback on further improvements.

Strengths. The experts liked the robust natural language understanding of the prototype. The system demonstrated a capacity to accurately comprehend and interpret queries, even when questions were imprecise or contained erroneous terminology. One participant emphasized this point, stating, "knowing the exact keyword that is used for the relevant dataset is typically necessary for finding specific data", but not with our system. The system's capability to allow users to articulate their information needs in natural language was highlighted, particularly for non-technical users. A participant noted that many of their colleagues are unaccustomed to handling data and found that "formulating a question in natural language eases the access to the data." Furthermore, commendation was given to the system's rapid accessibility of information. A participant remarked that under normal circumstances, it would take several minutes to locate the correct dataset and subsequently analyze the data for the required information. However, they identified our system "as definitely the fastest way to extract information." Lastly, participants expressed appreciation for the transparent display of the chain-of-thought. This feature was acknowledged for aiding the understanding of the data utilized in generating responses, making it simple to determine the correctness of dataset usage and the accuracy of data queries.

Weaknesses. Despite the notable strengths observed in our system, the expert identified several weaknesses, particularly in the prototype's occasional misinterpretation of data. Instances were noted where the agent erroneously presented values from incorrect columns or aggregated columns when such computation was unnecessary. Consequently, these misinterpretations led to incorrect answers, posing a potential challenge to the system's reliability. While displaying the chain-of-thought enables users to verify the relevance of queried data the participants expressed their concerns. Firstly, they highlighted the practicality issue, as most users may not undertake the time-consuming task of double-checking every response. Moreover, participants expressed worries about the potential harm to the assistant's credibility and the decline in user trust. One participant emphasized the current perception of the system as more of an "expert tool," primarily due to the display of executed SQL queries and data tables. Despite acknowledging the transparency inherent in this design principle, participants desired a more direct approach to displaying the chain-of-thought. A suggestion was made to "supply the relevant actions for retrieving the information directly with the answer through natural language," enhancing the user's understanding of the system's decision-making process. Furthermore, experts articulated the need for additional information about the data. For instance, inquiring about the number of employees might not encompass civil servants, self-employed individuals, and marginal employees in the count of socially insured people. Supplying such context was considered essential to assist users in verifying agent's responses, particularly for users lacking domain-specific knowledge. Lastly, concerns were raised regarding the sustainability of utilizing OpenAI's language model, encompassing environmental impact, future operational costs, and the security of users' data.

Opportunities. Amidst these identified weaknesses, experts highlighted valuable opportunities to enhance the prototype. First and foremost, they recommended incorporating a broader range of open data sources from the city into the assistant. One participant suggested, "Enriching the answers with background information and metadata" as a means to mitigate data misinterpretation by both the agent and the user, consequently reducing the potential for misinformation. Another proposed enhancement involves the inclusion of existing analyses and reports conducted by civil servants.

This addition could not only improve the agent's overall performance but also contribute to decreased query times, particularly when data is already preprocessed and evaluated. Furthermore, experts advised simplifying the technical aspects of the explanation of the chain-of-thought. This is seen as a means to enhance system accessibility, particularly for non-technical users, thereby making the system accessible to a broader audience. The final suggestion put forth was exploring on-premise open-source LLMs as a potential alternative to the OpenAI models. While this could reduce dependence on OpenAI, it could also decrease system performance. These opportunities present avenues for refining the prototype and addressing its identified shortcomings.

Threats. The experts also named their concerns about potential threats to our system. The primary concern expressed by the participants centered around the system providing inaccurate answers. One participant emphasized the difficulty in detecting misinterpretations of data by the system, noting that "supplying incomplete or false information could be very harmful to users." Consequently, the participants agreed that it is crucial to prioritize the evaluation of erroneous results in the future to prevent the spread of false information. In addition to this, participants recommended enabling users to more easily detect false answers by simplifying the presentation of reasoning steps and underlying data. This approach is seen as essential in enhancing the system's transparency and facilitating user verification of responses. Furthermore, the dependence on OpenAI, previously identified as a weakness, was acknowledged as a potential threat. Concerns were raised about the variability in model performance and the associated costs of maintaining the service, which could pose obstacles to the sustained effectiveness of our system.

6 Discussion and Conclusion

This paper presents the first design cycle of our Design Science Research project, dedicated to the design of a Large Language Model based open data assistant for effective use. Based on interviews, focus groups, and a literature review we derived meta-requirements and subsequently proposed three design principles grounded in the theory of effective use. We instantiated the principles in a prototype and evaluated the artifact in a focus group consisting of eight civil servants from a medium-sized German city. The overall feedback was very positive underscoring the potential of open data assistants. This research contributes design knowledge in the form of three design principles and an artifact. Our results provide valuable insights for future research on the design of Large Language Model based open data assistants for effective use. Moreover, our work offers practical guidance for open data providers and city administrations seeking to enhance data accessibility.

It is crucial to acknowledge limitations of our study. This includes the small sample size and the absence of a quantitative evaluation, both of which constrain the generalizability and comprehensive understanding of the prototype's impact on effective use of open data. While recognizing these limitations, we believe that this initial design cycle lays a robust foundation for subsequent iterations. Our study has shown promising avenues for future research, such as enhancing the transparent representation of the system's reasoning process and advancing the interaction between the user and the assistant to support human-AI collaboration. Adapting the assistant to varying user skill levels and facilitating mutual learning between users and the assistant could be an interesting avenue for future research. Additionally, the development and evaluation of additional DPs, supported by empirical evidence, would contribute further to the knowledge base in this domain.

References

1. Attard, J., Orlandi, F., Scerri, S., Auer, S.: A systematic review of open government data initiatives. Gov. Inf. Q. **32**(4), 399–418 (2015). https://doi.org/10.1016/j.giq.2015.07.006
2. Bran, A.M., Cox, S., Schilter, O., Baldassari, C., White, A.D., Schwaller, P.: ChemCrow: augmenting large-language models with chemistry tools, October 2023. https://doi.org/10.48550/arXiv.2304.05376
3. Burton-Jones, A., Grange, C.: From use to effective use: a representation theory perspective. Inf. Syst. Res. **24**(3), 632–658 (2013). https://doi.org/10.1287/isre.1120.0444
4. Christian Stocker: Ask ZüriCityGPT anything about the government and administration of the City of Zurich, June 2023. https://www.liip.ch/en/blog/askzuricitygpt-anything-about-the-government-of-the-city-of-zurich
5. Conradie, P., Choenni, S.: On the barriers for local government releasing open data. Gov. Inf. Q. **31**, 10–17 (2014). https://doi.org/10.1016/j.giq.2014.01.003
6. Diederich, S., Brendel, A., Morana, S., Kolbe, L.: On the design of and interaction with conversational agents: an organizing and assessing review of human computer interaction research. J. Assoc. Inf. Syst. (2022). https://doi.org/10.17705/1jais.00724
7. European Commission: Riding the wave How Europe can gain from the rising tide of scientific data Final report of the High Level Expert Group on Scientific Data. European Commission, January 2010
8. European Parliament: Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast), June 2019. http://data.europa.eu/eli/dir/2019/1024/oj/eng
9. European Union: The official portal for European data. https://data.europa.eu/en
10. Frauenhofer DPS: FragDenStaat Analytics. https://publicanalytics.fokus.fraunhofer.de/fragdenstaat/dashboard
11. German Federal Ministry of the Interior and Community: Informationsfreiheitsgesetz. https://www.bmi.bund.de/DE/themen/moderne-verwaltung/opengovernment/informationsfreiheitsgesetz/informationsfreiheitsgesetz-artikel.html
12. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**(1), 75–105 (2004). https://doi.org/10.2307/25148625
13. del Hoyo-Alonso, R., Rodrigalvarez-Chamarro, V., Vea-Murgía, J., Zubizarreta, I., Moyano-Collado, J.: Aragón open data assistant, lesson learned of an intelligent assistant for open data access. In: Følstad, A., et al. (eds.) Chatbot Research and Design. CONVERSATIONS 2023. LNCS, vol. 14524, pp. 42–57. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-54975-5_3
14. Hu, K., Hu, K.: ChatGPT sets record for fastest-growing user base - analyst note. Reuters, February 2023. https://www.reuters.com/technology/chatgpt-setsrecord-fastest-growing-user-base-analyst-note-2023-02-01/
15. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, adoption barriers and myths of open data and open government. Inf. Syst. Manag. **29**(4), 258–268 (2012). https://doi.org/10.1080/10580530.2012.716740

16. Jiang, J., Zhou, K., Dong, Z., Ye, K., Zhao, W.X., Wen, J.R.: StructGPT: a general framework for large language model to reason over structured data, October 2023. https://doi.org/10.48550/arXiv.2305.09645
17. Karpas, E., et al.: MRKL systems: a modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning, May 2022. https://doi.org/10.48550/arXiv.2205.00445
18. Keyner, S., Savenkov, V., Vakulenko, S.: Open data Chatbot. In: Hitzler, P., et al. (eds.) The Semantic Web: ESWC 2019 Satellite Events. ESWC 2019. LNCS, vol. 11762, pp. 111–115. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32327-1_22
19. Kuechler, W., Vaishnavi, V.: On theory development in design science research: anatomy of a research project. EJIS **17**, 489–504 (2008)
20. LangChain Inc: LangChain Docs. https://python.langchain.com/docs
21. Lewis, P., et al.: Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474 (2020)
22. Lourenco, R.P.: An analysis of open government portals: a perspective of transparency for accountability. Gov. Inf. Q. **32**(3), 323–332 (2015). https://doi.org/10.1016/j.giq.2015.05.006
23. McTear, M., Callejas, Z., Griol, D.: The Conversational Interface. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32967-3
24. Murray-Rust, P.: Open Data in Science. Nature Precedings, p. 1, January 2008. https://doi.org/10.1038/npre.2008.1526.1, publisher: Nature Publishing Group
25. Orszag, P.: Open Government Directive (2009). http://www.whitehouse.gov/open/documents/opengovernment-directive
26. Purwanto, A., Zuiderwijk, A., Janssen, M.: Citizen engagement with open government data: a systematic literature review of drivers and inhibitors. Int. J. Electron. Gov. Res. **16**(3), 1–25 (2020). https://doi.org/10.4018/IJEGR.2020070101
27. Quarati, A., De Martino, M.: Open government data usage: a brief overview. In: Proceedings of the 23rd International Database Applications & Engineering Symposium. pp. 1–8. IDEAS '19, June 2019. https://doi.org/10.1145/3331076.3331115
28. Rajkumar, N., Li, R., Bahdanau, D.: Evaluating the Text-to-SQL Capabilities of Large Language Models, March 2022. https://doi.org/10.48550/arXiv.2204.00498
29. Ruijer, E., Grimmelikhuijsen, S., Meijer, A.: Open data for democracy: developinga theoretical framework for open data use. Gov. Inf. Q. **34**(1), 45–52 (2017). https://doi.org/10.1016/j.giq.2017.01.001
30. Ruoff, M., Gnewuch, U., Maedche, A., Scheibehenne, B.: Designing conversational dashboards for effective use in crisis response. J. Assoc. Inf. Syst. **24**(6), 1500–1526 (2023). https://doi.org/10.17705/1jais.00801
31. Sadiq, S., Indulska, M.: Open data: quality over quantity. Int. J. Inf. Manag. **37**(3), 150–154 (2017). https://doi.org/10.1016/j.ijinfomgt.2017.01.003
32. Safarov, I., Meijer, A., Grimmelikhuijsen, S.: Utilization of open government data: a systematic literature review of types, conditions, effects and users. Inf. Polity **22**, 1–24 (2017). https://doi.org/10.3233/IP-160012
33. Streamlit Inc.: Streamlit Docs. https://docs.streamlit.io/
34. United Nations General Assembly: Universal Declaration of Human Rights (1948). https://www.un.org/en/about-us/universal-declaration-of-human-rights
35. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30. Curran Associates, Inc. (2017)
36. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a framework for evaluation in design science research. Eur. J. Inf. Syst. **25**(1), 77–89 (2016). https://doi.org/10.1057/ejis.2014.36
37. Wang, D., Richards, D., Bilgin, A.A., Chen, C.: Implementation of a conversational virtual assistant for open government data portal: effects on citizens. J. Inf. Sci. (2023).https://doi.org/10.1177/01655515221151140, publisher:SAGEPublicationsLtd

38. Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., Dwivedi, Y.K.: Open data and its usability: an empirical view from the Citizen's perspective. Inf. Syst. Front. **19**(2), 285–300 (2017). https://doi.org/10.1007/s10796-0169679-1
39. Wei, J., et al.: Emergent Abilities of Large Language Models, October 2022. https://doi.org/10.48550/arXiv.2206.07682
40. Wei, J., et al.: Chain-of-thought prompting elicits reasoning in large language models. Adv. Neural. Inf. Process. Syst. **35**, 24824–24837 (2022)
41. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM **9**(1), 36–45 (1966). https://doi.org/10.1145/365153.365168
42. Yao, S., et al.: ReAct: Synergizing Reasoning and Acting in Language Models, March 2023. https://doi.org/10.48550/arXiv.2210.03629
43. Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Sheikh Alibaks, R.: Socio technical impediments of open data. Electron. J. eGov. **10**, 156–172 (2012)
44. Zuiderwijk, A., Janssen, M., Dwivedi, Y.K.: Acceptance and use predictors of open data technologies: drawing upon the unified theory of acceptance and use of technology. Gov. Inf. Q. **32**(4), 429–440 (2015). https://doi.org/10.1016/j.giq.2015.09.005