



# Designing for AI Transparency in Public Services: A User-Centred Study of Citizens' Preferences

Stefan Schmager<sup>1</sup>, Samrat Gupta<sup>1,2</sup>(✉), Ilias Pappas<sup>1,3</sup>,  
and Polyxeni Vassilakopoulou<sup>1</sup>

<sup>1</sup> Department of Information Systems, University of Agder, Kristiansand, Norway  
samrat.gupta@uia.no

<sup>2</sup> Information Systems Area, Indian Institute of Management Ahmedabad, Ahmedabad, India

<sup>3</sup> Department of Computer Science, Norwegian University of Science and Technology,  
Trondheim, Norway

**Abstract.** Enhancing transparency in AI enabled public services has the potential to improve their adoption and service delivery. Hence, it is important to identify effective design strategies for AI transparency in public services. To this end, we conduct this empirical qualitative study providing insights for responsible deployment of AI in practice by public organizations. We design an interactive prototype for a Norwegian public welfare service organization which aims to use AI to support sick leaves related services. Qualitative analysis of citizens' data collected through survey, think-aloud interactions with the prototype, and open-ended questions revealed three key themes related to: articulating information in written form, representing information in graphical form, and establishing the appropriate level of information detail for improving AI transparency in public service delivery. This study advances research pertaining to design of public service portals and has implications for AI implementation in the public sector.

**Keywords:** Transparency · Human-Centered AI · Action Design Research · Design Preferences

## 1 Introduction

Artificial intelligence (AI) has advanced rapidly since its inception in the mid-1940s [1, 2]. In the public sector context, there are several examples of successful AI deployments. For instance, AI helps tax authorities in classifying individuals and corporate taxpayers to customize services and prevent fraud, predicting the arrival, departure and delays in public transport for commuters, pinpointing locations for police department in which they should concentrate their efforts in to deter crime, and immigration authorities to identify potential red flags and recommend immigration applications that should be approved [3–6]. There has also been some criticism of the use of AI applications in public services. For instance, concerns regarding prejudice and discrimination were raised when the Austrian labor administration established a system to classify job searchers according

to their chance of landing a position [7]. AI is envisaged to change society in both favorable and unfavorable ways, whether it be through bettering healthcare outcomes or streamlining supply chains [1, 8]. Globally, public service organizations are competing to be the first to capitalize on the AI hype and, in the process, are attempting to establish themselves as trustworthy, transparent and open AI stakeholders [1, 9–11].

Transparency is one of the important cultural values, particularly for Nordic countries [12, 13]. Transparency in AI caters to the right and ability to obtain information about the use of AI in organizations [1]. Transparency can take many forms, but among its many elements is the availability of data regarding internal operations and performance while improving the ability of outside entities to monitor decisions or actions occurring within that organization [14]. AI transparency in public service organizations is deemed highly important by numerous actors engaged with elucidating AI processes [1, 15]. Lack of transparency in the public services is often linked with dysfunctional government, poor decision-making, lack of accountability of public officials, social media induced polarization and spread of misinformation [10, 16–18]. It is generally believed that transparency is vital for maintaining a balance of power between the public service authorities and citizens thereby increasing the likelihood of exposing wrongdoings or abuse of power [19].

A common mechanism to promote transparency is providing access to data through Application Programming Interfaces (APIs), dedicated apps and portals [20]. Strictly regulated operations on such data are enabled through an interface for analyzing, visualizing, and exploring by using data-driven algorithms and big data technologies [10]. However, achieving full AI transparency in public services and its practical realization is challenging due to several reasons [19]. First, creating transparency from public services data is hampered by numerous socio-technical constraints, so simply opening the data is insufficient [21, 22]. Second, public services data portals may reinforce the opinions of their designers through the presentation of selected and aggregated data, thereby stifling the range of perspectives of a diverse society [10, 23]. Given the challenges, research advancements can be helpful for establishing transparency in rendering public services. To this end, we seek to address the following research question:

*What are the effective design strategies for AI transparency in public services?*

In doing so, we consider the context of a Norwegian public services organization that plays a key role in Norwegian government's public administration in managing different types of citizens' benefits. This study specifically focusses on one of the organization's several AI initiatives which is to develop a model to predict the length of sick leaves. Such a model can serve as an additional resource for case handlers assisting them in concentrating their efforts towards the goal of providing effective services intended for "user-adapted follow-up" [3]. Such a context is especially befitting to our study because Nordic countries are known for their innovative approaches to ensuring the transparent functioning of government, and they are highly regarded within the European Union for their high standards of transparency [1, 11]. This study contributes to the evolving discourse on human AI interaction and citizens' needs for AI transparency in the context of public services [8, 24, 25].

This article is structured as follows. Section 2 covers the related literature and background of this study while elaborating more on the role and importance of AI

transparency in rendering public services. The next section discusses the methodology employed to conduct this research thereby explaining the details pertaining to action design research and data collection. Section 4 presents an overview of the design of AI prototype used for this study. Subsequently, we discuss the findings of this study. A discussion of implications and limitations of this study is carried out in Sect. 6. Finally, Sect. 7 concludes this study.

## 2 Background

Transparency in public services is holistically defined as “the availability of information about an organization or actor that allows external actors to monitor the internal workings or performance of that organization” [14]. As such, ensuring transparency in rendering public services through AI has the potential to improve quality, enable personalization, ensure informed decision-making and more efficient use of resources in public service delivery [9, 26, 27]. Although, the awareness about ethical importance of algorithms and AI has been growing, the general public is unaware of even the most fundamental aspects of AI [28]. There are initiatives in place to make AI more understandable, such as a national online AI course of Finland which inspired a global course [1]. However, a lot needs to be done to overcome the misconceptions and reservations about the operation of AI. In the past, citizens’ concerns have led to termination of several public service AI initiatives after their inception [29]. One such example pertains to System Risk Indication (SyRI) which was a Dutch initiative for profiling citizens to detect fraud with social services [30]. The Court found that the SyRI Act is in violation of the European Convention on Human Rights (ECHR) thus calling for ceasing the further use of SyRI [31]. The plaintiffs contended that because of a lack of transparency, the citizens are unaware of the risk model or algorithm that was employed and therefore unable to defend themselves against decisions made using SyRI [30]. Another such example relates to a lawsuit filed by a number of Brazilian legal organizations demanding a halt on the use of face recognition technology for public safety<sup>1</sup>. They claimed that the technology violated citizens’ fundamental rights to free speech, assembly, and association, and would worsen issues such as inequality and structural racism. Therefore, it is crucial for citizens to know that AI is used in a transparent and accountable manner upholding all established rules for public service delivery [9, 32]. National strategic policies for AI strengthen notions of openness and trust while fostering sociotechnical understanding and resulting in a deeper comprehension of AI transparency [1].

### 2.1 Role of National Culture in AI Transparency

Though several cultures share the same values of transparency, trust, openness, accountability, and fairness, they do not give those values the same priority [33]. These variations may cause AI policies from various nations to enshrine the same values in different ways [1]. Moreover, with inherent economic, moral, and socio-cultural differences between

---

<sup>1</sup> <https://www.aljazeera.com/economy/2023/7/13/facial-recognition-surveillance-in-sao-paulo-could-worsen-racism>.

countries and regions, perceptions on how human-AI interactions should operate do not follow a common mindset [34]. For example, preference disparities in interface design were found when older adult East Asians and Caucasians were compared to two distinct interface designs [35]. As such, various governments (primarily in Nordic countries) have started to create national strategic policies for AI. It is essential to understand the socio-cultural shifts and how they should be maintained in a nation's strategic AI policy [36]. However, researchers and practitioners who can inform respective governments about the national outlook on AI are themselves grappling with its understanding due to scarcity of appropriate research studies in this field.

The concept of transparency has been frequently debated in the field of AI and Human Computer Interaction (HCI) [14, 37]. Proponents of transparency stress that transparency fosters understanding, a tighter bond between citizens, and a better understanding of public services [12]. Thus, a number of authors have contended that the public's lack of faith in public services is mainly because they are not provided with real information regarding the operations and efficacy of public services [38]. On the flip side, critics of transparency argue that showing the outcomes of public services may not actually boost citizens' trust [39]. They contend that transparency could possibly "delegitimize" the administration [14]. They assert that, since transparency requires users who can process it to function and most of the material released by public service portals is too difficult for even professionals to understand, transparency in AI just increases ambiguity and lack of trust [40, 41].

Despite contrasting viewpoints about AI transparency in public services, governments across the world are jockeying to make their public services more transparent. However, they come across different types of barriers in the design of open data applications and portals [42]. These barriers can be classified as data quality barriers, technical barriers, economic barriers, organizational barriers, ethical barriers, human barriers, political barriers, and usage barriers [10].

## 2.2 Norwegian Guidelines on AI Transparency

Transparency is highlighted as a key value in AI development in the Norwegian policy guide, which describes the country's national strategies for AI [34]. This guide also makes serious efforts to promote and practice transparency in its discussions of AI [43]. As compared to other national policy documents which provide truncated technical overviews, the first few pages of Norwegian policy guide explain AI technologies in detail thus providing any stakeholder fundamental AI literacy [1]. Furthermore, the Norwegian guide recognizes that "lack of transparency" is an issue when citizens attempt to access AI-based public services [1]. The policy states that in order to minimize burdens on citizens, improve access to services, and eventually enhance perceptions of transparency, "the Government has established a 'once only' principle to ensure that citizens and businesses do not have to provide the same information to multiple public bodies" [44].

This study responds to recent calls for research citizens' preferences and AI adoption in public services [26, 45–47].

## 3 Methodology

### 3.1 Research Context

This study is conducted in collaboration with a large Norwegian public welfare service organization responsible for different kinds of social benefits. The organization has established a specialized team to explore the potential of AI and machine learning technology in improving the efficiency, robustness, and quality of public services. One direction for the use of AI systems the team explores is to reduce administrative burdens for citizens and employees by improving the use of time and resources. The organization has developed a predictive model for estimating the probable duration of people being absent from their work due to sickness. The model is envisioned to be used as a decision support system for case handlers to decide whether a legally mandated meeting to discuss additional support for the person on sick leave will be necessary. These meetings can be resource-heavy and involve many actors, but Norwegian law allows for them to be omitted if deemed not to be necessary.

### 3.2 Action Design Research

The study employs the Action Design Research (ADR) methodology, which is a collaborative research approach that allows researchers to work closely with organizations to solve real-world problems [48]. The ADR methodology involves multiple iterative “build-intervene-evaluate” (BIE) cycles, where both researchers and stakeholders jointly shape the research artefact. We held bi-weekly collaboration sessions with the organizational partner to discuss design choices, provide feedback, and plan for the next iterations of the prototype. The emerging artefact then served as a tool to explore and refine the practical application of AI in public service, ensuring that the resulting systems are not only technologically advanced but also ethically aligned and human-centric [49]. Consistent with ADR, we reflected on the problem understanding, theories, and emerging designs, evaluating potential knowledge contributions to the field.

### 3.3 Data Collection and Analysis

We recruited 28 individuals from the age range of 18–65 to mirror the demographic distribution of Norwegians on sick leave, as reported by the central statistics office in Norway (2022). Because the study focused on understanding specific user perspectives in a given context through rich qualitative data, the insights of 28 participants were found to be sufficient in providing diverse and in-depth empirical data [50–52]. The data collection included three consecutive parts. First, participants provided general demographic information and self-assessed their prior AI knowledge and technology use frequency. Next, a moderated task-based user study was conducted, where participants interacted with the interactive prototype and made decisions based on a scenario. The scenario led to a decision point at which they were asked whether they would consent to the use of the AI-supported sick leave duration prediction system. Participants were asked to think aloud while interacting with the prototype, which helped to capture thoughts and feelings better and provided insights into reasoning [53] Finally, participants rated

their agreement with a set of statements related to the scenario and answered open-ended questions about their decision-making processes. The study followed ethical standards and legal regulations and was approved by the Norwegian Center for Research Data (NSD).

The collected data was then categorized into three types: scale ratings, verbal feedback, and interactions with the interactive prototype. Initially, each data type was examined independently before being integrated with the other two. In the first step, we analyzed participants' responses to predefined questions. We compiled the scale ratings into a spreadsheet, which provided a comprehensive view of the data, and revealed diversity, trends, and allowed us to track any shifts in opinions throughout the study. We then undertook an exhaustive analysis of the qualitative data, including verbal feedback from the think-aloud recordings and responses to open-ended questions to different explanation types. This involved multiple reviews of the study recordings and a detailed examination of interview transcripts. During this process, we highlighted notable incidents, quotes, and expressions from participants, and engaged in in-depth discussions within our research team to collaboratively interpret these insights. Finally, we analyzed how participants interacted with the prototype, drawing from session recordings and notes. We focused on identifying which information elements were accessed, how participants engaged with these elements and the time spent on them. By integrating this analysis with the insights from the earlier stages, we aimed to enhance our overall examination and identify emerging concepts and recurrent themes, cross-referencing observations and findings across the different data categories.

## 4 Prototype Development

To evaluate the different types of explanations and information in a naturalistic context, we developed an ensemble artefact, embodied in an interactive prototype [54]. The prototype mimics a public service portal, depicting a specific interaction sequence of users being asked whether they would consent to the optional use of an AI-based prediction system. The development of our prototype was informed by the Google PAIR framework, which is a comprehensive collection of guidelines formalizing Google's insights from industry and academia<sup>2</sup>. Google's PAIR framework is the most comprehensive and balanced when comparing different Human-Centered AI frameworks [55]. Auernhammer, 2020 emphasizes that it is important to embed technologies within a user-centric interface that aligns with users' workflows and cognitive models, beyond just crafting a robust algorithm or machine-learning model [56].

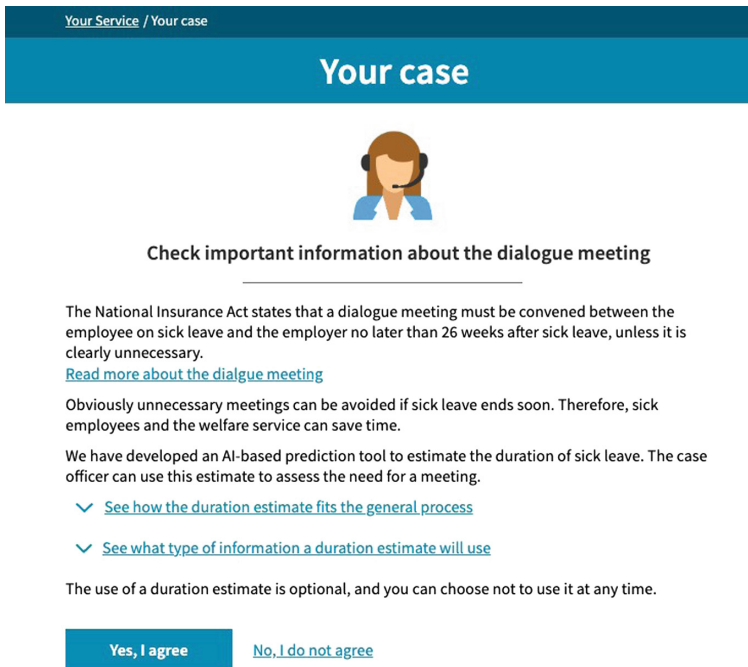
### 4.1 Information Elements

The prototype replicates a public service portal, presenting a scenario where users are asked about their willingness to consent to the use of an AI-based prediction system. The scenario starts with a standard portal notification that directs users to the introduction page of the AI prediction tool. Consistent with the PAIR framework's first chapter (User

---

<sup>2</sup> <https://pair.withgoogle.com/guidebook/chapters>.

Needs + Defining Success), it's crucial to identify areas where AI's capabilities align with user requirements. Ehsan & Riedl, 2020 emphasize that for AI systems to be considered trustworthy and ethically sound, explanations must be clear enough to enable users to make informed judgments based on these understandings [57]. In the specific public service context, it's also vital to consider and convey the rights and responsibilities of all parties involved [58]. For our case, this includes explaining the government's mandate on mandatory dialogue meetings, the possibility of waiving these meetings, and the citizen's right to dissent to the use of the AI system. This information is provided in a textual format (Fig. 1). Furthermore, the information includes the collective benefit derived from consenting to use AI in the particular context. This entails addressing considerations of social ethics rather than imposing moral obligations on citizens.



**Fig. 1.** Introductory interface of the AI prediction tool

To enhance transparency regarding the process, data usage, and the model's functionality, participants were allowed to explore different information elements and descriptions of the AI prediction tool. One such element is a process diagram, designed to illustrate the AI's role within the overall process and emphasize the case handler's critical involvement (Fig. 2). As per Chapter 3 of the PAIR guidelines (Mental Models), such visual explanations are vital in setting realistic expectations about the role of an AI system, its capabilities, limitations, and the value it offers (Google PAIR, 2022). The diagram shows the AI system within the context of the full process flow and delineates areas where AI supplements human decision-making. It also depicts the procedure in

the absence of the AI-enhanced prediction tool, in case a citizen would not consent to its use. The chart aims to provide a comprehensive view of AI’s role in the process, thereby fostering transparency and a deeper understanding, and enabling citizens to also better understand their rights and role in the process.

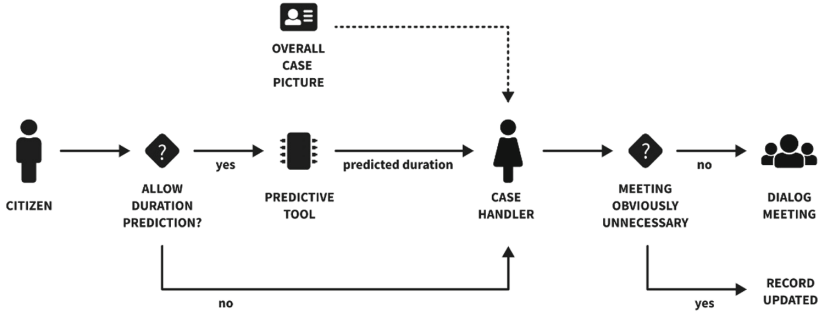


Fig. 2. Process visualization

The prototype also includes a data table to provide information about the data used by the AI model and the rationale (Fig. 3). This element is based on the PAIR recommendations in Chapter 2, which emphasizes the importance of explaining the usage and benefits of specific data to users. The data table consists of an introductory paragraph clarifying that no new data will be collected, but existing data reused. The table itself consists of two columns: one for the type of data and the other for why it’s necessary. This transparent approach to data aims to enhance participants’ understanding of the system and provide insights into the personal data used.

INFORMATION ELEMENT	NEED
Start date of sick leave	Calculation of duration of the sick leave.
Change in degree in the sickness	Identify degree changes in sickness to detect improvements or deterioration
Diagnosis and co-diagnosis, and change in diagnose / co-diagnose	Different types of diagnoses / co-diagnoses are associated with short / long sick leave durations.
Gender and age of sick leaver	Individual age and gender are associated with sick leave duration
Sick leave history	Relationship between previous sickness absence and current sickness absence.
Unique ID of attesting doctor	Variations in sick leave practice affects sick leave.

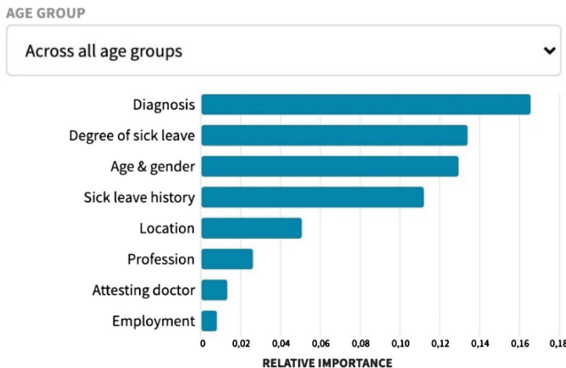
Fig. 3. Information about the data used by the AI prediction tool

Lastly, we integrated an interactive chart showing Shapley Values for feature importance visualization (Fig. 4) [59]. This chart highlights the importance of various variables across different age groups, providing a visual representation of the model’s logic. This enhances exploratory transparency and helps citizens adjust their trust in the system



[58]. This integration aligns with Chapter 4 of the PAIR framework (Explainability + Trust), which states that while providing detailed explanations of the model's logic may not always be possible, the reasoning behind predictions can sometimes be made accessible. The interactive feature importance chart aims to fill this gap by presenting users with a visually engaging way to comprehend the model's inner workings.

A higher relative importance means that the specific group of information elements has a greater effect on the duration estimate.



**Fig. 4.** Feature importance chart

At the end of the interaction sequence, citizens reach a consent decision point where they must decide whether to agree to the use of AI or to opt-out (Fig. 5). This emphasizes the voluntary nature of their agreement and the option to not provide consent. Consistent with Chapter 5 (Feedback + Control) of the PAIR framework, it's essential to convey citizens' control over their data usage and the decisions they make.

The use of a sick leave duration prediction is completely optional and you can opt out of the process at any time.




**Fig. 5.** Consent decision point

The chapter also stresses the importance of feedback mechanisms, which can help refine the AI model's performance and user experience. While opting out can be considered implicit feedback, this stage also presents an opportunity to request more detailed feedback from citizens. Within the prototype, we encouraged users to express their opinions and objections (Fig. 6). This not only facilitates the direct participation of citizens but also upholds their right to object to personal data usage without negative consequences or justification requirements. This approach ensures transparency and upholds the principle that the right to object is an inherent right of citizens within the process.

YourService / Your sick leave

## Your sick leave



**You did not agree to the use of a sick leave duration prediction**

Please note, the use of a sick leave duration prediction is completely optional and there are no negative consequences for you.

But to be able to further improve our services, it would be great if you could share your reasons why you did not agree to an automated assessment. This will help us to better understand and incorporate citizens values and understandings.

REASONS FOR OPTING OUT OF THE AUTOMATED ASSESSMENT

✓ Please select

I do not want the model the access my data

I do not think that the model can predict my sick leave duration

I do not think the expected duration of my sick-leave is relevant when deciding about the dialogue meeting.

[Send feedback](#) | [I prefer to not provide feedback](#)

**Fig. 6.** Feedback form in case of opting out

## 5 Findings

The participants of the study were generally positive about the use of AI in public service, as indicated by their responses when asked about their level of comfort with the use of AI. Before interacting with the prototype, around 53% of the participants reported feeling comfortable with the general use of AI in public services, while 30% were unsure or declined to say, and 17% felt uneasy about the idea. These initial responses suggest a mix of cautiousness and openness towards AI in the public sector, with some participants expressing concerns about the potential risks and uncertainties associated with AI use. However, after interacting with the prototype, the participants' outlook became more positive overall. 65% stated feeling at ease, 23% remained uncertain or preferred not to answer, and only 12% maintained reservations. This slight shift in the level of comfort suggests that the interactive experience provided by the prototype helped address some of the concerns and uncertainties that participants had about AI in public services. To further unpack how these interactions might have influenced participants' stances on the use of AI in public service, we also analyzed the verbalized comments from the think-aloud protocol as well as the actual interactions with the prototype. This analysis revealed three key themes: a) Articulating Information in Written Form, b) Representing Information in Graphical Form, and c) Establishing the Appropriate Level of Information Detail.

## 5.1 Articulating Information in Written Form

The first theme emphasizes the importance of copywriting in AI interfaces. Participants carefully evaluated the textual explanation used in the system. While the feedback on the provided explanations was generally positive, several participants expressed comments suggesting that the explanations as well as other interface text were still unclear. There were multiple remarks about challenges in comprehending the textual content, with some participants suggesting that the language used was too technical or complex. Participants stated: *“There were some words I didn’t understand”* and *“I think that here it needs to be made much easier. Sit down really. It’s not designed for a simple user”*. Some participants connected their comments about the language to the organization. One participant expressed: *“It’s their terminology that I think is... is a bit wrong... or I think a lot of people will stop here and really need help.”* Another participant remarked: *“That’s kind of a typical [name of organization] description. [...] They can write it a little easier, but I understand what’s there.”* This underscores the critical role of clear and understandable communication in AI-driven public services and highlights the nuanced role of language in shaping user experiences and trust in AI systems.

## 5.2 Representing Information in Graphical Form

The second theme identified in our data analysis is centered around the visual elements used in the prototype, specifically focusing on the process visualization (Fig. 2) and the feature importance charts (Fig. 4). Our analysis delved into which graphical representations captured the participants’ attention and their subsequent feedback on these visual elements. Upon examining the engagement with this visualization, we found that different graphical representations resonated differently with users, influenced significantly by their backgrounds. For instance, the process visualization was particularly well-received by participants who had familiarity with such representations in their professional contexts. These participants stated: *“I really love this”* and *“I love to see how the process is with and without the tool”*. At the same time, this type of visualization was difficult to comprehend and puzzled users who had no prior familiarity with such representations. Some users expressed their general favor for visual explanations but still had difficulties understanding the visualization: *“I found it exciting that it was explained with pictures or icons, but I also found it a bit confusing”*. Others stated that the provided visualization did not match their personal Mental Model of the process: *“I don’t quite feel my understanding of the situation is similar to that on the process diagram then.”* However, some participants stated that they did not understand the process visualization at all: *“Hard to figure out the chart”* and *“I don’t understand anything about this”*.

The second graphical information in the prototype was an interactive feature importance chart (Fig. 4). Despite aiming to provide more transparency into the inner workings of the sick-leave prediction model, the feature importance chart was found to be too difficult to comprehend by participants, describing it as *“too academic”*, or it was unclear what was the underlying reason for including it: *“Yes, I understand what it shows. But I don’t understand why it shows it really.”* The findings hint towards the necessity of considering the diversity of potential prior experience with certain forms of visualizations,

as well as incorporating diverse graphical representations to accommodate different citizen groups' needs. The right combination of visual approaches tailored to diverse users is key to building transparent, intuitive interfaces that help citizens of all backgrounds interact successfully with AI systems.

### 5.3 Establishing the Appropriate Level of Information Detail

The third theme identified in our analysis explores the optimal level of detail for the information presented to citizens, particularly concerning the data table (Fig. 3) as well as textual information. The data table provided a comprehensive overview of the data utilized by the sick leave prediction model, including explanations for the necessity of each data type. Several participants perceived the level of detail as overly extensive, describing it as *“too much to read”* or *“Way too much text”*. However, a few participants appreciated the availability of detailed information, delving into it to thoroughly grasp the scope and reason for the data repurposing. One participant noted, *“I think it is useful that it says what the purpose of collecting information is, and what kind of information is collected.”* However, despite not fully reading the details, many participants acknowledged the value of having this information available. As one participant expressed: *“I thought it was really nice. Even though I haven't read it very thoroughly right now, I thought it was really good to know exactly why they're using it and that information and what it's being used for: [...] And then I get some confidence in understanding that that's why and that they're not just gathering any information and kind of that it's to give rise to something.”*

The level of detail in the provided information was also brought up in other places. One other instance was the explanation of the governmental mandate for dialogue meetings. One participant remarked that: *“I would rather have said that it is required by law, not necessarily the National Insurance Act. In other words, for someone who has nothing to do with the National Insurance Act, it is not so important which law it is.”* But the same participant also appreciated the availability of more information if desired, stating that: *“[...] if you don't know what it entails then it's nice that you can read a little more about what it entails.”* The overarching insight from this theme is that citizens appreciate the presence of detailed information, even if they do not engage with it extensively. The availability of thorough explanations contributed to a sense of ease among participants, illustrating that the mere provision of detailed explanations can foster trust in AI systems. Leveraging progressive disclosure, which presents essential information first and then elaborates on more detailed aspects, can help maintain focus on the primary objectives. This approach is about empowering citizens and fostering confidence through a commitment to openness and accountability.

## 6 Discussion and Contributions

As evident from our findings, in order to enhance transparency and drive citizens' engagement with AI enabled public services, there is a need to restructure information elements and realign content according to preferences of citizens. To generate trust, citizens must play active role in using public services and provide actionable feedback about their

perceptions of the use of AI in these services. AI capabilities in public services can advance the creation of public value in the form of trust, flexibility, adaptability, and responsiveness [26]. One of the most significant obstacles that public services must overcome is driving engagement with their stakeholders amidst growing speculations about responsible use of AI [10]. In order to ensure that AI policy initiatives are accurate and pertinent, it is imperative that public offices increase the level of awareness and involvement with citizens. Thus, evaluating the preferences of the public can promote continued cooperation between all stakeholders, accelerating the flow of information and the ensuing chances for further development of public AI based interfaces.

### **6.1 Contribution to Research**

The findings of this study underscore the link between the government and the citizens thereby broadening the scope of current research on adoption of AI in the public sector. The three themes identified in this study emphasize the critical role of clear and comprehensible communication in AI-enabled public services. Educating the public about the features of AI systems and the role that humans play in decision-making significantly improves public perceptions of AI, which is crucial for its adoption. Our research goes beyond earlier conceptual and review-based studies [1, 10, 15, 19, 60] by being able to assess individual preferences and perceptions about AI transparency. The information citizens receive strengthens their faith in the public service portals, and having a human in the loop gives them confidence that they can clarify their unique situation in an open, accountable and trustworthy manner.

### **6.2 Contribution to Practice**

This study sheds light on how the interactive engagement with the AI prototype enhances its positive perception, suggesting that structured information and hands-on engagement with AI can positively shape public opinion about transparency and trust in public services. The right combination of comprehensible communication, visual approaches, and requisite information availability is the key to building transparent, intuitive interfaces that help citizens of all backgrounds interact successfully with AI enabled public services. Public entities may find it easier to provide their services efficiently if citizen's trust is enhanced and they derive a value from usage of AI in public services. Through the identification of characteristics that contribute to a positive attitude toward AI, we offer practical guidance to practitioners about the transparent design of AI systems within a public service setting.

### **6.3 Limitations and Future Research**

Citizens' attitudes toward AI are influenced by their socio-cultural identity and, in particular, their general level of trust in the government, which is notably high in Norway (OECD 2022). Consequently, it is critical to carry out more targeted research with larger samples as well as studies in other sociocultural contexts. Secondly, in this study, we focus on citizens' preferences of transparency in AI based public services. Further

research may delve into the interplay of key characteristics such as transparency, trust and openness of AI based public services. Trust is systemic in nature, embedded in the broader system of public and private entities connected to AI [1, 61]. As such, there is a fascinating opportunity for research partners from around the world for investigating and creating a framework for human-centered in public services.

## 7 Conclusion

This study provided valuable insights into how citizens perceive and interact with AI systems introduced in public services. The findings demonstrate that directly engaging users helps improve generally positive views by addressing concerns through informed experience. The three key themes that emerged from our analysis showed citizens have high standards for transparent communication. The implications of this research can help guide development of future citizen-facing AI to maximize benefits and allay worries through responsible, user-centered practices. Hence, it aims to contribute with practical insights in the extant body of research on normative guidelines for the responsible design and development of AI with the goal of human benefit [9, 26, 55, 62, 63]. Our research contributes to the broader understanding of public sentiment towards AI in public services and offers practical implications for the design of AI based systems in the public sector.

## References

1. Robinson, S.C.: Trust, transparency, and openness: how inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technol. Soc.* **63**, 101421 (2020)
2. Ulnicane, I., Knight, W., Leach, T., Stahl, B.C., Wanjiku, W.G.: *Governance of Artificial Intelligence: Emerging International Trends and Policy Frames: The Global Politics of Artificial Intelligence*. Taylor & Francis, Milton Park, Oxfordshire (2022)
3. Schmager, S., Grøder, C.H., Parmiggiani, E., Pappas, I., Vassilakopoulou, P.: What do citizens think of AI adoption in public services? Exploratory research on citizen attitudes through a social contract lens. In: *Proceedings of the 56th Hawaii International Conference on System Sciences* (2023)
4. Kuziemski, M., Misuraca, G.: AI governance in the public sector: three tales from the frontiers of automated decision-making in democratic settings. *Telecommun. Policy* **44**(6), 101976 (2020)
5. Höchtl, J., Parycek, P., Schöllhammer, R.: Big data in the policy cycle: policy decision making in the digital era. *J. Organ. Comput. Electron. Commer.* **26**(1–2), 147–169 (2016)
6. Molnar, P.: Technology on the margins: AI and global migration management from a human rights perspective. *Camb. Int. Law J.* **8**(2), 305–330 (2019)
7. Lopez, P.: Bias does not equal bias: a socio-technical typology of bias in data-based algorithmic systems. *Internet Policy Rev.* **10**(4), 1–29 (2021)
8. Zhang, P., Nah, F.F.H., Benbasat, I.: Human-computer interaction research in management information systems. *J. Manag. Inf. Syst.* **22**(3), 9–14 (2005)
9. Tiwari, A.A., Gupta, S., Zamani, E.D., Mittal, N., Agarwal, R.: An overarching conceptual framework for ICT-enabled responsive governance. *Inf. Syst. Front.* (2023). <https://doi.org/10.1007/s10796-023-10415-4>

10. Matheus, R., Janssen, M., Janowski, T.: Design principles for creating digital transparency in government. *Gov. Inf. Q.* **38**(1), 101550 (2021)
11. Persson, P., Zhang, Y., Asatiani, A., Lindman, J., Rudmark, D.: Toward citizen-centered digital government: design principles guided legacy system renewal in a Swedish municipality (2024)
12. Kassen, M.: Understanding transparency of government from a Nordic perspective: open government and open data movement as a multidimensional collaborative phenomenon in Sweden. *J. Glob. Inf. Technol. Manag.* **20**(4), 236–275 (2017)
13. Vidaver-Cohen, D., Brønn, P.S.: Reputation, responsibility, and stakeholder support in Scandinavian firms: a comparative analysis. *J. Bus. Ethics* **127**, 49–64 (2015)
14. Grimmelikhuijsen, S., Porumbescu, G., Hong, B., Im, T.: The effect of transparency on trust in government: a cross-national comparative experiment. *Public Adm. Rev.* **73**(4), 575–586 (2013)
15. Felzmann, H., Villaronga, E.F., Lutz, C., Tamò-Larrieux, A.: Transparency you can trust: transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data Soc.* **6**(1), 2053951719860542 (2019)
16. Gupta, S., Jain, G., Tiwari, A.A.: Polarised social media discourse during COVID-19 pandemic: evidence from YouTube. *Behav. Inf. Technol.* **42**(2), 227–248 (2023)
17. Kaur, K., Gupta, S.: Towards dissemination, detection and combating misinformation on social media: a literature review. *J. Bus. Ind. Mark.* **38**(8), 1656–1674 (2023)
18. Qureshi, I., Bhatt, B., Gupta, S., Tiwari, A.A.: Introduction to the role of information and communication technologies in polarization. In: Qureshi, I., Bhatt, B., Gupta, S., Tiwari, A.A. (eds.) *Causes and Symptoms of Socio-Cultural Polarization*, pp. 1–23. Springer, Singapore (2022). [https://doi.org/10.1007/978-981-16-5268-4\\_1](https://doi.org/10.1007/978-981-16-5268-4_1)
19. Fung, A.: Infotopia: unleashing the democratic power of transparency. *Polit. Soc.* **41**(2), 183–212 (2013)
20. Luna-Reyes, L.F., Bertot, J.C., Mellouli, S.: Open government, open data and digital government. *Gov. Inf. Q.* **31**(1), 4–5 (2014)
21. Conradie, P., Choenni, S.: On the barriers for local government releasing open data. *Gov. Inf. Q.* **31**, S10–S17 (2014)
22. 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)
23. Kitchin, R., Lauriault, T.P., McArdle, G.: Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Reg. Stud. Reg. Sci.* **2**(1), 6–28 (2015)
24. Saldanha, D.M.F., Dias, C.N., Guillaumon, S.: Transparency and accountability in digital public services: learning from the Brazilian cases. *Gov. Inf. Q.* **39**(2), 101680 (2022)
25. Vössing, M., Kühn, N., Lind, M., Satzger, G.: Designing transparency for effective human-AI collaboration. *Inf. Syst. Front.* **24**(3), 877–895 (2022)
26. Zamani, E.D., Smyth, C., Gupta, S., Dennehy, D.: Artificial intelligence and big data analytics for supply chain resilience: a systematic literature review. *Ann. Oper. Res.* **327**(2), 605–632 (2023)
27. Vassilakopoulou, P., Haug, A., Salvesen, L.M., Pappas, I.O.: Developing human/AI interactions for chat-based customer services: lessons learned from the Norwegian government. *Eur. J. Inf. Syst.* **32**(1), 10–22 (2023)
28. Kieslich, K., Keller, B., Starke, C.: Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Soc.* **9**(1), 20539517221092956 (2022)
29. van Veenstra, A.F., Grommé, F., Djafari, S.: The use of public sector data analytics in the Netherlands. *Trans Gov. People Process Policy* **15**(4), 396–419 (2020)

30. Bekker, S.: Fundamental rights in digital welfare states: The case of SyRI in the Netherlands. In: Spijkers, O., Werner, W.G., Wessel, R.A. (eds.) *Netherlands Yearbook of International Law* 2019. NYIL, vol. 50, pp. 289–307. T.M.C. Asser Press, The Hague (2021). [https://doi.org/10.1007/978-94-6265-403-7\\_24](https://doi.org/10.1007/978-94-6265-403-7_24)
31. Van Bekkum, M., Borgesius, F.Z.: Digital welfare fraud detection and the Dutch SyRI judgment. *Eur. J. Soc. Secur.* **23**(4), 323–340 (2021)
32. Henman, P.: Improving public services using artificial intelligence: possibilities, pitfalls, governance. *Asia Pac. J Public Adm.* **42**(4), 209–221 (2020)
33. Schwartz, S.: A theory of cultural value orientations: explication and applications. *Comp. Sociol.* **5**(2–3), 137–182 (2006)
34. van Berkel, N., Papachristos, E., Giachanou, A., Hosio, S., Skov, M.B.: A systematic assessment of national artificial intelligence policies: perspectives from the Nordics and beyond. In: *Proceedings of the 11th Nordic conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pp. 1–12 (2020)
35. Haddad, S., McGrenere, J., Jacova, C.: Interface design for older adults with varying cultural attitudes toward uncertainty. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1913–1922 (2014)
36. Kulesz, O.: Culture, platforms and machines: the impact of artificial intelligence on the diversity of cultural expressions. *Intergovernmental Committee Prot. Promot. Diversity Cult. Expr.* **12** (2018)
37. Abdul, A., Vermeulen, J., Wang, D., Lim, B.Y., Kankanhalli, M.: Trends and trajectories for explainable, accountable and intelligible systems: an HCI research agenda. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article 582 (2018)
38. Cook, T.E., Gronke, P.: The skeptical American: revisiting the meanings of trust in government and confidence in institutions. *J. Politics* **67**(3), 784–803 (2005)
39. Bannister, F., Connolly, R.: The trouble with transparency: a critical review of openness in e-government. *Policy Internet* **3**(1), 1–30 (2011)
40. Heald, D.: Varieties of transparency. In: Heald, D. (ed.) *Transparency: the key to better governance?* British Academy (2006). <https://doi.org/10.5871/bacad/9780197263839.003.0002>
41. Im, T., Cho, W., Porumbescu, G., Park, J.: Internet, trust in government, and citizen compliance. *J. Public Adm. Res. Theory* **24**(3), 741–763 (2014)
42. Chen, C.P., Zhang, C.Y.: Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. *Inf. Sci.* **275**, 314–347 (2014)
43. Dexe, J., Franke, U.: Nordic lights? National AI policies for doing well by doing good. *J. Cyber Policy* **5**(3), 332–349 (2020)
44. Stani, E., Barthélemy, F., Raes, K., Pittomvils, M., Rodriguez, M.A.: How data vocabulary standards enhance the exchange of information exposed through APIs: the case of public service descriptions. In: *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, pp. 807–810 (2020)
45. Asatiani, A., Malo, P., Nagbøl, P.R., Penttinen, E., Rinta-Kahila, T., Salovaara, A.: Sociotechnical envelopment of artificial intelligence: an approach to organizational deployment of inscrutable artificial intelligence systems. *J. Assoc. Inf. Syst. (JAIS)* **22**(2), 325–252 (2021)
46. Wirtz, B.W., Langer, P.F., Fenner, C.: Artificial intelligence in the public sector—a research agenda. *Int. J. Public Adm.* **44**(13), 1103–1128 (2021)
47. Mikalef, P., et al.: Enabling AI capabilities in government agencies: a study of determinants for European municipalities. *Gov. Inf. Q.* **39**(4), 101596 (2022)
48. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action design research. *MIS Q.*, 37–56 (2011). <https://doi.org/10.2307/23043488>



49. Schmaget, S., Pappas, I., Vassilakopoulou, P.: Defining human-centered AI: a comprehensive review of HCAI literature. In: *Proceedings of the 2023 Mediterranean Conference on Information Systems* (2023)
50. Creswell, J.W., Poth, C.N.: *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage publications. Thousand Oaks, California (2016)
51. Marshall, B., Cardon, P., Poddar, A., Fontenot, R.: Does sample size matter in qualitative research? A review of qualitative interviews in IS research. *J. Comput. Inf. Syst.* **54**(1), 11–22 (2013)
52. Faulkner, L.: Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behav. Res. Methods Instrum. Comput.* **35**, 379–383 (2003)
53. Ericsson, K.A., Simon, H.A.: Verbal reports as data. *Psychol. Rev.* **87**(3), 215 (1980)
54. Miah, S.J., Gammack, J.G.: Ensemble artifact design for context sensitive decision support. *Australas. J. Inf. Syst.* **18**(2), 5–20 (2014)
55. Shneiderman, B.: Human-centered artificial intelligence: reliable, safe & trustworthy. *Int. J. Hum. Comput. Interact.* **36**(6), 495–504 (2020)
56. Auernhammer, J.: Human-centered AI: the role of Human-centered design research. In: Boess, S., Cheung, M. Cain, R. (eds.), *The Development of AI Synergy - DRS International Conference* (2020)
57. Ehsan, U., Riedl, M.O.: Human-centered explainable AI: towards a reflective sociotechnical approach. In: Stephanidis, C., Kurosu, M., Degen, H., Reinerman-Jones, L. (eds.) *HCI 2020. LNCS*, vol. 12424, pp. 449–466. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-60117-1\\_33](https://doi.org/10.1007/978-3-030-60117-1_33)
58. Schmaget, S.: From commercial agreements to the social contract: human-centered AI guidelines for public services. In: *The 14th Mediterranean Conference on Information Systems (MCIS)*, Catanzaro, Italy (2022)
59. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30** (2017)
60. Vassilakopoulou, P.: Sociotechnical approach for accountability by design in AI systems. In: *2020 European Conference on Information Systems, ECIS* (2020)
61. Thiebes, S., Lins, S., Sunyaev, A.: Trustworthy artificial intelligence. *Electron. Mark.* **31**, 447–464 (2021)
62. Vassilakopoulou, P., Parmiggiani, E., Shollo, A., Grisot, M.: Responsible AI: concepts, critical perspectives and an information systems research agenda. *Scand. J. Inf. Syst.* **34**(2) (2022)
63. Akbarighatar, P., Pappas, I., Vassilakopoulou, P.: A sociotechnical perspective for responsible AI maturity models: findings from a mixed-method literature review. *Int. J. Inf. Manage. Data Insights* **3**(2), 100193 (2023)