Identifying Algorithmic Decision Subjects' Needs for Meaningful Contestability

Yurrita, Mireia; Verma, Himanshu; Balayn, Agathe; Alfrink, Kars; Gadiraju, Ujwal; Bozzon, Alessandro

DOI
10.1145/3757415

Publication date
2025

Document Version
Final published version

Published in
Proceedings of the ACM on Human-Computer Interaction

Check for updates

Identifying Algorithmic Decision Subjects' Needs for Meaningful Contestability

MIREIA YURRITA, Delft University of Technology, Utrecht University, The Netherlands
HIMANSHU VERMA, Delft University of Technology, The Netherlands
AGATHE BALAYN, Delft University of Technology, The Netherlands
KARS ALFRINK, Delft University of Technology, The Netherlands
UJWAL GADIRAJU, Delft University of Technology, The Netherlands
ALESSANDRO BOZZON, Delft University of Technology, The Netherlands

Contestability has been proposed as a key element in designing algorithmic decision-making processes that safeguard decision subjects' rights to dignity and autonomy. However, little is known about how contestability can be operationalized based on decision subjects' needs and preferences. We address this research gap by identifying decision subjects' information and procedural needs for enacting meaningful contestability. To this end, we chose an illegal holiday rental detection scenario as our case; a high-risk decision-making process in the public sector. We conducted 21 semi-structured interviews with citizens with experience renting their homes out and different levels of AI literacy. We found that decision subjects request interventions that facilitate (1) cooperation in sense-making, (2) support in contestation acts, and (3) appropriate responsibility attribution. Our results highlight the cooperative work behind contestability, and motivate future efforts to structure individual and collective action, to personalize explanations for contestability, and to open up sites of contestation in AI pipelines.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; **Collaborative and social computing**; • **Computing methodologies → Machine learning**.

Additional Key Words and Phrases: contestability, decision subjects, information needs, procedural needs, public AI

ACM Reference Format:
Mireia Yurrita, Himanshu Verma, Agathe Balayn, Kars Alfrink, Ujwal Gadiraju, and Alessandro Bozzon. 2025. Identifying Algorithmic Decision Subjects' Needs for Meaningful Contestability. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW234 (November 2025), 29 pages. https://doi.org/10.1145/3757415

Authors' Contact Information: Mireia Yurrita, Delft University of Technology, Utrecht University, Delft, Utrecht, The Netherlands, M.YurritaSemperena@uu.nl; Himanshu Verma, Delft University of Technology, Delft, The Netherlands, H.Verma@tudelft.nl; Agathe Balayn, Delft University of Technology, Delft, The Netherlands, agatheb.research@gmail.com; Kars Alfrink, Delft University of Technology, Delft, The Netherlands, c.p.alfrink@tudelft.nl; Ujwal Gadiraju, Delft University of Technology, Delft, The Netherlands, U.K.Gadiraju@tudelft.nl; Alessandro Bozzon, Delft University of Technology, Delft, The Netherlands, A.Bozzon@tudelft.nl.

ACM 2573-0142/2025/11-ARTCSCW234
https://doi.org/10.1145/3757415

1 Introduction

Several artificial intelligence (AI)[1] systems employed for decision-making in the public sector (e.g., AI for policy enforcement or for essential public services) [37] can negatively affect decision subjects' safety and fundamental rights, and are, therefore, considered *high-risk* by the European Union's Artificial Intelligence Act (EU AI Act) [36]. In order to safeguard decision subjects' rights to dignity and autonomy in high-risk algorithmic decision-making, an increasing number of scholars in the HCI community (e.g., [2, 100, 101, 114]) have claimed that AI systems should be *contestable* i.e., open and responsive to human intervention throughout their lifecycles [3]. Despite recent interest in making AI systems–and the decision-making processes where these are embedded–contestable, most prior work is theoretical, and has rarely accounted for the perspective of decision subjects when suggesting contestable AI design guidelines [3, 57, 82]. Failing to generate empirical insights into decision subjects' needs for contestability might, in turn, lead to designs that do not contribute to decision subjects' perceptions of control and voice [100, 114].

From a procedural perspective, the few empirical studies conducted to date have either (1) considered the standpoint of *human controllers* (i.e., domain experts who interact with the algorithmic system [3]) for identifying the challenges of implementing contestable AI systems in the public sector or (2) have focused on designing for contestability in contexts other than the public sector (e.g., content moderation [101]) [2]. The extent to which those findings are aligned with decision subjects' procedural needs for contestability in the public sector is unknown. From an information[2] perspective, recent work has explored the interplay between output explanations and *recourse* (i.e., operationalization of contestability that allows decision subjects to change the decision output by acting on input variables [98]). However, decision subjects might want to contest not only the decision output, but also more fundamental issues regarding the system (e.g., goal of the system, the idea of automation itself [100], or data sources [7]). It is still unclear which information enables decision subjects to engage in such contestation acts.

In this paper, we aim to generate *empirical* insights into the procedural and informational means that *decision subjects* need to meaningfully contest *high-risk* public decision-making processes. Consequently, we seek to answer the following research question:

> **RQ: What are decision subjects' information and procedural needs to meaningfully contest algorithmic decision-making processes?**

To address this research question, we opted for a scenario in public decision-making; more specifically, a risk scoring system for the detection of illegal holiday rentals[3] (Section 3). We

[1] Throughout this paper we will use the terms *Artificial Intelligence (AI)* and *algorithmic system* interchangeably to refer to systems that are designed to interpret external data, and to learn from that data to perform specific tasks [112]. Due to the "demand for data, technical complexity, and unpredictable interactions" [112] of such systems, human-AI interactions are uniquely difficult to design for. This same nature of AI makes contestability uniquely difficult to design for [3].

[2] Throughout the paper, we will use the term (1) *information* to refer to a set of facts that describes a decision or a decision-making process, (2) *information item* to refer to a unit of relevant information [76], and (3) *explanation* to refer to tools or processes that an agent (explainer) uses to describe the decision (or the decision-making process) to another agent (explainee) [73]. An explanation involves a communicative effort for making the information that composes the explanation *understandable*. *Information needs* can, therefore, involve both *information items* —relevant content— or *explanations* —information (items) presented as part of an interaction.

[3] https://algoritmeregister.amsterdam.nl/en/illegal-holiday-rental-housing-risk/ (last accessed 14.01.2024). Note: the entry of this algorithmic system in the algorithm register is from 2020. Due to delays in data collection as a consequence of the COVID-19 pandemic, the system has not been deployed. See Section 3.1 for information about the status of the system and the rationales behind choosing this case.

conducted 21 semi-structured interviews with participants who have personal experience renting out their homes as short-term rentals. We presented a scenario to our participants where they were detected by the algorithmic system, and asked questions on *what* they would like to contest in the decision-making process, *how* they would formulate their contestation, and the information they would need for it. Given the effect of AI literacy on users' information needs [53, 114], we ensured diversity in levels of AI literacy among participants. Our study was preregistered before data collection.[4]

Our results indicate that contestability in algorithmic decision-making is not limited to individual appeal processes, and requires a cooperative effort between *civil servants* (in roles that go from policy-making to AI development or street-level bureaucracy), *citizens*, and *third parties* (e.g., legal counsellors). As far as *information needs* are concerned, participants sought information that could help them make sense of algorithmic decisions and that would enable them take action to remedy the situation (Section 4.1). Participants expressed their willingness to engage in communication with human controllers and external parties to make sense of the provided information. When it comes to *procedural needs*, our participants expressed the need for support mechanisms (Section 4.2), i.e., they sought support both from the decision-making organization and from fellow decision subjects. Participants additionally highlighted the need for interventions that would ensure accountability in the decision-making process and social transparency (i.e., visibility of the complex socio-organizational context [33]) in public administration (Section 4.3).

In this paper, we make two main contributions to the CSCW community:

(1) We adopt an empirical approach to contestability and generate insights into decision subjects' information and procedural needs for contestability in the public sector.
(2) We draw implications for practice and for research. These implications encourage public agencies and the research community to account for the cooperative work behind contestability.

2 Related work

This section summarizes previous work on algorithmic decision-making in the public sector and contestable AI. In Section 2.1, we summarize prominent work on public AI. In Section 2.2, we include papers that have theoretically defined procedural means for contestability. In Section 2.3, we include literature concerning information needs for meaningful contestability.

2.1 Algorithmic Decision-Making in the Public Sector

In an environment like the public administration where decisions are often made based on incomplete, contradictory, and changing information [23, 64, 79, 90], the usage of AI has the potential to improve both the efficiency and quality of decision-making processes [116]. However, the development and use of AI for public decision-making has also been claimed to be uniquely challenging [84] mainly because street-level bureaucrats [5] need to be able to effectively apply human discretion while navigating bureaucratic processes in a resource-deficient context [84].

In addition to challenges in the development and effective use of algorithmic systems in the public sector, public AI systems face issues of perceived legitimacy [19]. Perceived legitimacy of public decision-making processes not only depends on the quality of the decision-making. According to the process-based model suggested by Tyler [97], the public's behaviour is "powerfully influenced by people's subjective judgments about the fairness of the procedure" through which decisions are

[4] https://osf.io/ejyt5 While more widespread in quantitative studies, by pre-registering our qualitative study we aim to (1) describe the original aims of the study, (2) register the assumptions that underlie the collection and analysis of the data, and (3) enable the scientific community to monitor the evolution of the study [42].

made. Previous work has shown that communities impacted by algorithmic decisions in the public sector have concerns about the way in which data and algorithms are used [19]. While citizens are not opposed to delegating to fully autonomous systems, they do want to engage in a dialectical exchange with system controllers [4].

Contestability in algorithmic decision-making processes has, indeed, been defined as a dialectical exchange between decision-makers and decision subjects [82], a form of procedural justice that gives voice to decision subjects [3], and increases perceptions of legitimacy [75]. Perceived legitimacy of public decision-making processes, in turn, has been claimed to contribute to compliance, cooperation, and empowerment of citizens [97]. Given (1) the rapid adoption of AI systems in the public sector, (2) the potential (harmful) impacts of their widespread use, and (3) the relevance that *contestability* bears for procedural fairness and legitimacy perceptions in such high-stakes algorithmic decision-making processes, we decided to examine decision subjects' contestability needs in a public decision-making scenario.

2.2 Procedural Means for Contestability

Contestability refers to the quality that enables different actors (e.g., human controllers, decision subjects) to "*understand, construct, shape and challenge*" algorithmic decision-making processes [56]. Since algorithmic decision-making processes rely on interconnectivity [80] (i.e., the score that an individual gets is dependent on the scores of other individuals), designing ways in which decision subjects can meaningfully ensure a correct decision output and fair process is of paramount importance. Contestability has been conceptualized as *recourse* (i.e., the act of changing the output of an algorithmic system by altering input variables [98]), *appeal* (i.e., the act of opposing an algorithmic decision because it is considered to be faulty [106]) and as a design goal, *contestability by design* (i.e., AI systems that are open and responsive to human intervention throughout their lifecycles [3, 6, 82]). Both recourse and appeal are limited to acting on the decision output and are reactive in nature, whereas contestability by design allows measures to be taken *ex-ante* [3, 6]. Due to algorithmic systems' demand for data, technical complexity, and unpredictable interactions [112], contestability in algorithmic decision-making presents additional challenges compared to human-led decision-making [80]. Recent prominent work (e.g., [3, 6, 68, 82]) have set the grounds for conceptualizing contestability in algorithmic decision-making and have *theoretically* defined some procedural means that would enable algorithmic systems to be contestable by design.

Through a literature review, Alfrink et al. [3] synthesized five system features (e.g., built-in safeguards) and six development practices (e.g., agonistic development approaches) that contribute to contestable AI. Alfrink et al. [2] then used this framework to design a conceptual contestable AI system and identify the challenges of implementing contestable AI in the public sector. Similarly, Lyons et al. [68] analyzed responses to the Australian "AI Ethics Framework" which includes contestability as a key ethical principle and conceptualized how contestability could operate in relation to AI. Both frameworks were created based on theoretical claims without empirical grounding. There is, therefore, little insight into which of those elements decision subjects need to shape and challenge algorithmic decision-making. While acknowledging the importance of setting a normative framework that legally constrains the scope of contestability, in this paper we argue that the lack of guidelines on decision subjects' procedural needs for contestability might result in contestation processes that either do not improve perceptions of legitimacy, in general [100] or that do not improve perceptions of *procedural voice* (i.e., ability to share one's views during a procedure [95]) and *influence* [95] in particular [114].

One of the very few *empirical* studies on decision subjects' needs for contestability was grounded in a context other than the public sector (i.e., content moderation [101]). The extent to which decision subjects' procedural needs for contestability in contexts such as content moderation can be extrapolated to contestability needs for contesting algorithmic decision-making in the public sector is not clear.

2.3 Information for Enacting Meaningful Contestability

For decision subjects to build arguments as part of their contestation process, they need knowledge, which, in turn, requires information [82]. This information needs to be meaningful for decision subjects to be able to engage in a rational and fruitful discussion [82], i.e., functional information that empowers decision subjects to exercise their *right to contest* algorithmic decisions as defined in Article 22(3) of the EU's *General Data Protection Regulation* (GDPR) [87]. Such information can be provided in the form of *explanations* [68] or *justifications* [44]. The goal of *justifications* is to demonstrate the appropriateness of the decision with respect to a norm (i.e., these are normative and extrinsic), whereas *explanations* aim at generating understanding about how a decision was made (i.e., these are intrinsic and factual). Information for meaningfully enacting contestability has been claimed to include the *why* behind the decision, as well as, *how* the decision-making process took place [3, 46, 82]. Despite the importance given to the topic, there is no empirical insights into the content and form of the justifications or explanations that decision subjects deem necessary. Determining what should or should not go into explanations is not trivial. Some decision subjects might want to "know everything" about *how* the system works, as it is the case for human-AI collaboration and for recommender systems [53, 88].

For contestability, previous work has mostly looked into generating decision output explanations for enabling decision subjects to engage in acts of *recourse* (e.g., [48, 78]). To this end, decision subjects need to *understand* [48, 107] and *act* [52] on an unfavorable decision through a set of *actionable* factors (i.e., factors that can be acted upon so as to change the decision output [51, 92]) or counterfactual explanations [48, 104]. Previous work indicates that when engaging in contestation processes, decision subjects might not only want to contest the decision output itself (scope of recourse) but also issues concerning the goals of the system or the idea of automation [100, 114]. Limiting information to output explanations might, therefore, hinder decision subjects' ability to question structural aspects (e.g., data sources [7]) of the decision-making process [41]. Current knowledge around *what* decision subjects would like to contest and *how* they would like to formulate their contestations might, therefore, be subject to blind spots resulting from limiting information to output explanations [68, 114].

2.4 Positioning Our Work

In this paper, we aim to generate in-depth empirical insights into decision subjects' procedural and information needs for meaningful contestability that is not limited to algorithmic outputs. To this end, we conduct semi-structured interviews with potential decision subjects in a decision-making process in the public sector.

Our work builds on prior work and further informs it by:

(1) **Adopting an *empirical* approach to identify needs for meaningful contestability**. Our results will provide insights into how decision subjects' needs align or differ from the claims made in theoretical frameworks for contestability summarized in Sections 2.2 and 2.3.
(2) **Focusing on *decision subjects'* information and procedural needs for contestability**. Our results will provide a needs-based perspective that can further inform the organizational challenges for contestability identified by Alfrink et al. [2] in public administration.
(3) **Focusing on a *public* decision-making context**. To this end, we choose a case in which risk scoring is used for fraud detection. Contestability needs that we identify might complement the ones identified by Vaccaro et al. [101] on content moderation processes.

3 Method

In this section, we introduce the case that we adopted for our study (Section 3.1) and summarize details about participant recruitment (Section 3.2) and interview design (Section 3.3). Supplementary materials associated with this paper include the pre-registration document, screening survey, interview protocol, prompts used during the interviews, and the codebook. These are all openly available in our OSF repository for the benefit of the community and in the spirit of Open Science.[5]

3.1 Case: Illegal Holiday Rental Detection

Algorithmic systems for law enforcement fall into the category of high-risk AI systems [36]. Within this category, we decided to select an algorithmic system suggested by the municipality of Amsterdam for accelerating the detection of illegal short-term rentals[6] as our case. The algorithmic system was designed to be used if a report on a particular address was received. After receiving the report, the algorithmic system (based on a random forest model) would compute the probability of a property being illegally rented for holiday purposes. It would do so by relying on data about the identity and housing rights of the decision subject, the building, and previous illegal housing cases. Based on the probability, civil servants would decide whether to further investigate the report. This system was suggested in November 2019 and expected to be pilot tested in 2020. However, due to the effect that the COVID-19 pandemic had on worldwide tourism, there were delays in data collection, which resulted in the system not being deployed to date (January 2024).[7]

Although the system has not been deployed, there are two main reasons why this represents a compelling case for identifying decision subjects' needs for contestability. First, this case deals with a timely and increasingly complex problem that impacts cities in several Western countries. Due to the issues that short-term rentals offered to tourists (e.g., Airbnb) have generated in the availability of long-term rentals for citizens [12], municipalities in several Western countries have started to regulate those rentals (e.g., Amsterdam, Barcelona) or even ban them (e.g., New York City) [74]. This last example is especially relevant. In September 2023, the municipality of New York City decided to ban short-term rentals that host more than two guests while the owner or tenants of the property are not present.[8] To enforce this policy, platforms like Airbnb are required to ensure their listings have pertinent licenses issued by the municipality certifying compliance with the regulation. This has led to the proliferation of a "black-market" where lessors use platforms such as Facebook or Craiglist to announce their short-term rentals and to avoid being policed by the platforms.[9] In response to this trend, many municipalities have put in place workflows where citizens can (anonymously) report an illegal holiday rental.[10] Algorithmic systems could, then, be seen as powerful tools to filter reports and help civil servants identify which reports they should investigate further. This is, precisely, the way in which the system suggested by the municipality of Amsterdam was designed to operate. It is, therefore, a realistic representation of what municipalities in other Western countries could end up implementing. In an anticipatory exercise, the insights we

[5] https://osf.io/c5x7e/

[6] https://algoritmeregister.amsterdam.nl/en/illegal-holiday-rental-housing-risk/(last accessed 14.01.2024)

[7] See the status of the project in the following official communication https://amsterdam.raadsinformatie.nl/document/12731876/2#search=%22Afhandeling%20toezegging%20pilot%20algoritme%20Alpha%20handhaving%20vakantieverhuur%22(last accessed 14.01.2024)

[8] https://www.nytimes.com/2023/09/05/nyregion/airbnb-regulations-nyc-housing.html (last accessed 14.01.2024)

[9] https://www.wired.com/story/airbnb-ban-new-york-illegal-listings/ (last accessed 14.01.2024)

[10] Barcelona: https://meet.barcelona.cat/habitatgesturistics/en; New York City: https://portal.311.nyc.gov/article/?kanumber=KA-02317; Berlin: https://ssl.stadtentwicklung.berlin.de/wohnen/zweckentfremdung_wohnraum/formular/adresswahl.shtml; Porto: https://www.asae.gov.pt/espaco-publico/formularios/queixas-e-denuncias.aspx (last accessed 14.01.2024)

get on decision subjects' contestability needs can be useful not only for the research community looking into contestability in algorithmic decision-making—*contestability by design*, which goes beyond post-hoc appeals, requires measures to be taken ex-ante [3, 6]—but also for municipalities thinking of implementing algorithmic systems to accelerate the detection of illegal holiday rentals.

Second, this case is part of the *algorithm register*[11] initiative launched by various European cities [38]. In an effort to ensure that algorithmic systems used for public services are "*responsible, transparent, and secure*", several cities (e.g., Amsterdam, Barcelona, Brussels) have put in place a register where information is provided about algorithmic systems used as decision support systems for public services. To this end, a short description about the system, information about mechanisms to ensure its responsible use, and technical information are openly shared. A form to provide feedback for continuous improvement is also included for each entry. The insights we get about decision subjects' *information* and *procedural needs* could, therefore, help improve a system that already advocates for transparency and contestability by design (e.g., by implementing mechanisms for quality assurance [3]).

3.2 Participant Recruitment and Selection

Given the timely and widespread applicability of the case (i.e., concerning major cities in several Western countries), we recruited participants who have experience renting out their homes as short-term rentals. Participants are located in municipalities from Western countries where workflows for detecting illegal holiday rentals have been put in place. Although it is unknown whether *all* these municipalities use algorithmic systems as part of those workflows (i.e., transparency around algorithmic systems used for public services is still not common practice [38]), if an algorithmic system like the one suggested by the municipality of Amsterdam was implemented, our participants could become decision subjects of the system by being correctly or incorrectly flagged. We recruited 21 participants in total (demographics in Table 1). We stopped collecting data when additional interviews failed to generate significantly new information. According to Clarke and Braun [24], when using qualitative interviews to capture experiences, understandings, and perceptions, the recommended dataset size is moderate (i.e., 10-20 participants), which aligns with the number of participants we recruited.

Table 1. Summary of our participants' demographics

Feature	**Category (Number of participants)**
AI literacy	High (7), Medium (7), Low (7)
Background	Computer Science (5), Engineering (4), Law (4), Business (3)[12], Design (3), Architecture (2), Physics (1), Social Work (1)
Country[13]	Netherlands (9), Spain (7), US (2), Portugal (1), Germany (1), Canada (1)
Immigration status[14]	Native (12), Non native (9)

Since AI literacy has been shown to impact information needs [53, 114], we decided to ensure diversity in participants' AI literacy. We created a screening survey (cf. our repository) with questions about participants' literacy in and experience with AI. The screening survey comprised four items defined by Schoeffer et al. [86] (in a 5-point Likert scale). This way of operationalizing AI literacy has been used in prior studies and has been shown to be useful in capturing differences in informational fairness perceptions across individuals [4, 114]. We published the screening survey

[11] https://www.algorithmregister.org/ (last accessed 14.01.2024)

on online housing channels. We also put posters around our institution and reached out to personal contacts. We then selectively invited participants for our interview. To this end, we averaged the four items that define AI literacy and divided participants in low, medium, and high AI literacy [4, 53, 54]. We reached out to participants while ensuring AI literacy diversity. As done in previous work [53], we refined the boundaries that define what constitutes low, medium and high AI literacy based on the interview answers given by our participants. This allowed us to account for potential discrepancies between self-assessed and functional AI literacy. A summary of our participants' AI literacy is provided in Table 2. We refer to our participants as P_k, where k is the identifier of a specific participant.

Table 2. Overview of our participants' AI literacy

AI Literacy	**Specification**	**Participants**
Low: self-assessment [1,3]	Had not heard much about AI	P14, P17
	Could not understand what AI entailed	P9, P13
	Unconfident about technicalities of AI	P3, P18, P19
Medium: self-assessment (3,4]	Technical background; familiar with basic statistics	P1, P2, P5, P11, P16
	Working on concepts adjacent to AI	P12, P20
High: self-assessment (4,5]	Working with or on AI on a managerial level	P6, P7, P10
	Working with or on AI from an engineering perspective	P8, P21
	Working with or on AI from a fairness perspective	P4, P15

3.3 Design of Interview Protocol and Materials

For our study, we opted to conduct qualitative interviews prompted by vignettes (i.e., written fictitious descriptions of events related to a topic of study [11, 81]). Choosing to run qualitative interviews allowed us to get rich and detailed insights into participants' needs for contestability [24]. As suggested by Clarke and Braun [24] when using qualitative interviews to capture participants' perceptions and needs, participants had a hypothetical personal stake in the selected case (i.e., they were renting properties as short-term rentals). The usage of vignettes has been claimed to be appropriate to capture perceptions, beliefs, and attitudes in social research, as well as to identify participants' reactions and needs in a particular situation [11, 47]. Scenario- or vignette-based techniques have previously been used in qualitative AI research; for instance, in public AI research for exploring the perspectives of decision subjects in the early stages of AI system usage in child welfare services [19], or in explainable AI research for advancing the conceptual development of *social transparency* [34]. Our interviews comprised four main sections and three prompts to encourage our participants to expand on their answers [24]. The interview protocol and prompts can be found in our repository.

(1) *Participants' background and experience.* First, we asked a series of questions to capture our participants' experiences with contestation processes. The objective was to identify their motivations for deciding whether or not to contest an unfair decision. We also asked them about their experience and motivation for renting out their homes.

[12] Two of our participants have a joint background in Business and Law

[13] It refers to the country where the rented property is located. Our participants need to deal with that country's public administration for managing their property's rental.

[14] It refers to the mismatch between the home country of our participants and the country where the rental is located. Our participants are native or non native in the eyes of the public administration of the country where the property is located.

(2) *Perceptions around the use of AI*. We then introduced the first prompt (i.e., a fictional *piece of news* introducing the case; Figure 1), and asked our participants about the appropriateness and benefits of using algorithmic systems for detecting possible illegal holiday rentals. The fictional piece of news included real information about the system summarized from the introductory text in the algorithm register entry. The piece of news was tailored to the city where the rental was located to make the scenario more believable and for participants to feel they had a personal stake in the topic [24]. The objective of this section was to get a sense of how our participants perceived algorithmic systems (e.g., its perceived capabilities [49]) as a way to get context to their motivation for contesting (or not) the algorithmic decision-making process.

> "**Amsterdam has limited living space**; both for citizens and visitors. If a citizen wants to rent out their home to tourists, they need to meet certain requirements. **They must also report it to the municipality.**
>
> Not everyone adheres to those conditions. The municipality sometimes receives **reports**, for instance **from neighbors or rental platforms,** who suspect that a home has been rented out without meeting those requirements. If such a report is filed, employees of the department of Surveillance & Enforcement can start an investigation.
>
> **The municipality of Amsterdam has adopted an Artificial Intelligence system** that supports the employees of the department of Surveillance & Enforcement in their investigation of the reports made concerning **possible illegal holiday rentals**."

Fig. 1. Example of the piece of news shown to participants to introduce our case. The material used with each participant included the name of the city where their short-term rental was located.

(3) *Object of contestation (what to contest) and means for contesting (how to contest)*. Next, we introduced the second prompt (i.e., a *letter*; Figure 2). The letter was divided into three main sections. These included (a) first warning and future penalty (i.e., giving notice [50]), (b) right to present arguments against the decision by calling the municipality (i.e., right to be heard [50]) and (c) right to know about the decision and the decision-making process (i.e., reason giving [50]). The amount of the penalty[15] and the timeframes for contesting[16] are informed by the contestation procedures available in the municipality of Amsterdam, within the Dutch public administration context. The letter was tailored to the city where the rental was located. For this interview, we deliberately designed the letter using accessible language (i.e., avoiding legal jargon), following the guidelines on accessibility of (digital) communications of public authorities.[17] We asked our participants how they would react to this letter and how appropriate they considered the contestation means (i.e., a phone call) suggested by the municipality (i.e., perceived voice and influence [95], expected treatment [15]). Through this section we aimed to capture *what* our participants would like to contest and *how* they would ideally like to proceed [67].

[15] https://www.amsterdam.nl/wonen-leefomgeving/wonen/boetes-overtredingen-vakantieverhuur-bed/ (accessed 14.01.2024)
[16] https://www.cjib.nl/direct-regelen/ik-ben-het-niet-eens-met-mijn-boete (accessed 14.01.2024)
[17] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32016L2102 (accessed 14.01.2024)

Gemeente Amsterdam

Housing Department, Surveillance and Enforcement Division
Amsterdam City Hall
Bijlmerdreef 1005C, 1103 TW
+31 20 555 5837

July 15, 2023

Re: Illegal Holiday Housing Rental

Dear ...,

This letter has been issued because the Housing Department, Surveillance and Enforcement Division of the Municipality of Amsterdam has found the house you own at the address ... to be illegally rented as holiday housing without due notice to the Municipality.

This formal letter constitutes a **FIRST WARNING** and it is a request to strictly adhere to the private vacation rental policies of the Municipality. After this, we will be forced to take stronger action. The Municipality of Amsterdam may request a payment of up to 21,750 EUR penalty.

Disagree with the warning you received?

Then you can file an objection by calling the Housing Department, Surveillance and Enforcement Division at +31 20 555 5837. You can do so within 6 weeks since the day you receive the letter. They will ask for more information from you and will offer you the possibility to provide an explanation for this violation.

Additional information:

The Housing Department, Surveillance and Enforcement Division uses an Artificial Intelligence (AI) system as part of their workflow to detect and investigate potential illegal holiday rentals. If you would like more information about the system and how it has been used in your case, please check algoritmeregister.amsterdam.nl and introduce your case number 185274.

Sincerely,

Housing Department, Surveillance and Enforcement Division

Bijlmerdreef 1005C, 1103 TW, Amsterdam

Fig. 2. Example of the letter shown to participants. The letter used with each participant was tailored to include their name, address, the logo, name and contact details of the municipality where their short-term rental was located.

(4) *Information needs.* Finally, we introduced the third prompt (i.e., the *information sheet*). The information available in the algorithm register was summarized in three categories [55] and

organized through a color code (Figure 3): (a) green for information related to the scope of the system (i.e., reasons for system conception, role of the system and potential harms [71]), (b) orange for the decision rules of the process (i.e., information about training data [62, 71], system architecture [71]) and (c) blue for information related to the outputs (i.e., rationales behind instance-level decisions and model performance [62]). For the decision explanation, we simulated a SHAP explanation [65] (i.e., feature-based explanation [86]).[18] We indicated data features that contributed to the decision. We used positive (+) signs to indicate that a data feature contributed to high fraud risk [16, 31]. We avoided to include data features that are explicitly protected by law (e.g., gender [13]) in the decision explanation. Through this prompt, we asked our participants what they would like to know more about. The objective of this section was to identify if they would use this information for building their arguments as part of the contestation process [82, 107]. We decided to introduce information about the system after the letter to see if there were any differences between the object of contestation before and after being given information about the system [41].

Ecological validity: We (the authors) designed the interview material so that (1) it would illustrate a decision-making scenario where the illegal holiday rental detection system could be embedded, and (2) it would be sufficiently believable for our participants, i.e., it would not be considered science fiction [2, 10]. To improve the ecological validity of our work, we ensured participants were coming from cities where illegal holiday rental detection efforts are already in place and we tailored the materials (e.g., logos, address, recipient name) to the city where each participants' property was located. We additionally pilot tested the interview protocol and the prompts with 2 experts in human-computer interaction (different from the authors) from our institution. For each interview question, we evaluated whether it helped answer our research question, we looked for problematic assumptions, and we reflected on how meaningful participants would find it [24]. For each prompt, we checked the wording and layout. Based on the insights we got from the pilot test, we modified the layout of the *Information Sheet* to make it more engaging. We decided to change the decision explanation to textual form [86, 103], rather than a visual to avoid saliency bias and halo effect [32, 35].

3.4 Data Collection and Analysis

Data Collection. We conducted the interviews between July and August 2023. All interviews were conducted online, using the Zoom video conferencing tool, and lasted 1 hour on average. Participants were offered 25 EUR (or equivalent) as compensation for their time. Our study was approved by a research ethics committee at our institution. All our participants signed an informed consent form. After each interview, we acquired the transcription of the recording through the videoconferencing platform if the interview was conducted in English. We then anonymized the transcription. If the interview was conducted in a language other than English, we had the recording transcribed in the original language through a third-party transcription software, and then locally translated the transcription using *DeepL*.[19] After obtaining the transcriptions in English, we reviewed and corrected them.

Data Analysis. A critical realist [39, 72] and contextualist [72, 94] approach underpins our analysis. We acknowledge that although a reality exists and informs our findings, we, as researchers, play a

[18] The numbers in the presented SHAP value (first blue box in Figure 3) aim at representing the effect of each feature on the output risk value, rather than their effect on the final risk probability. That is why some of these features present an effect > 1. We left it up to the participants to ask for clarifications about the scale if they considered this information item to be important for contestability.

[19] DeepL Translator: https://www.deepl.com/en/translator.

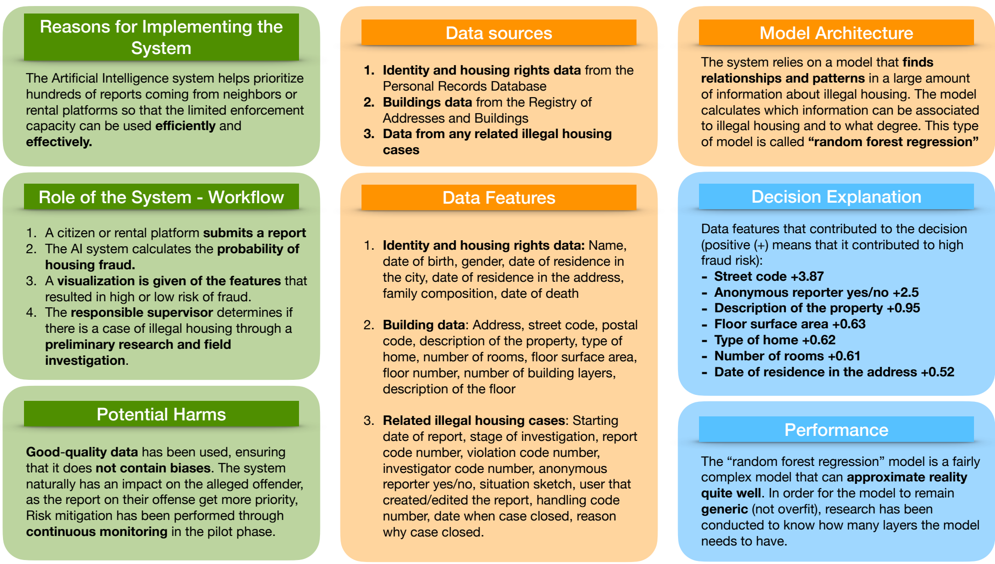

Fig. 3. Information Sheet provided to our participants. It includes the information relevant to the algorithmic decision-making process summarized from its entry in the algorithm register. It is color-coded. Green refers to information related to the scope of the system and it includes reasons for implementing the system, the role of the system, and potential harms. Orange refers to information about the decision rules of the system and it includes data sources, data features, and model architecture. Blue refers to information related to the outputs and it includes the decision explanation and performance information.

role in constructing knowledge and these findings cannot, thus, be considered truly objective [109]. We analyzed our data using *reflexive thematic analysis* with a combination of inductive and deductive orientation to data [17, 25]. Reflexive thematic analysis is a flexible method that allows an in-depth engagement with the data. This approach is adequate for answering our research question where we aim to identify patterns *in data* and interpret them [18]. We conducted the data analysis on *Atlas.ti*.[20]

Analysis Procedure. Data analysis was led by the first author. After transcribing –and translating when applicable– the recordings, the first and second authors cleaned the transcriptions. The first, second, and third authors read the transcriptions and got familiar with the material. The first author open-coded the transcripts and clustered the codes in code groups. The second and third authors partially coded the data and reviewed the code groups. The first author then crafted the themes. All authors reviewed and mapped the themes. In total, three main themes and six sub-themes were developed. The first author finally refined the codes based on the final themes. The final codebook can be found in our repository. Having different researchers analyze the data helped us reflect on

[20] Atlas.ti URL: https://atlasti.com.

different perspectives on the same data to develop richer insights into that data. Reflexivity helped researchers identify their own situatedness within the research and take responsibility for it [25].

Statement of Positionality. Reflexivity acknowledges that knowledge production is contingent on the researcher producing it [14]. As researchers living and working at a Western European university, we recognize that our perspectives shape the research and knowledge we generate. Our disciplinary backgrounds include engineering, cognitive science, computer science, HCI, and design. We have previously argued for making algorithmic decision-making processes contestable.

4 Results

The cooperative nature of contestability was a salient characteristic of contestation processes and was present throughout the interviews. We structure our results to highlight the cooperative work involved in contestability at three different points in time: (1) during the sense-making process that enables decision subjects to understand the provided information (post-hoc intervention[21]; Section 4.1), (2) during the contestation act (post-hoc intervention; Section 4.2), and (3) during the development and deployment of the AI system (ex-ante intervention; Section 4.3). We provide an overview of the themes and sub-themes in Table 3. We name themes as T*i.* and sub-themes as T*i.j.*, where *i* and *j* are the indexes of a particular theme and sub-theme. To improve readability, we avoid naming participants (P_k) for *each* statement that compose our themes and sub-themes. We, instead, give a sense of the prominence of each statement by using terms such as *a few*, *many*, *mostly*, *generally*, *unanimously*. A detailed mapping of the participants whose responses led to the statements in our results section is included in our repository. Additionally, we release our codebook, where we include the specific quotes that compose each statement. The codebook can also be found in our repository.

Table 3. Overview of themes and sub-themes.

Information and Procedural Needs
T1. Cooperation in Sense-Making – post-hoc intervention
T1.1. Strategizing Information Requests
T1.2. Facilitating Dialogue with Controllers
T2. Social Support in Contestation Acts – post-hoc intervention
T2.1. Seeking For Organizational Support
T2.2. Seeking For Peer Support
T3. Distributed Responsibility – ex-ante intervention
T3.1. Ensuring Algorithmic Accountability
T3.2. Fostering Social Transparency

4.1 T1. Cooperation in Sense-Making

The first theme highlights the need for cooperation in the sense-making process that precedes the contestation act. **This cooperative effort involves decision subjects, legal and AI experts that decision subjects could contact, and street-level bureaucrats acting as controllers.** Controllers are street-level bureaucrats that are involved in the first instance of the decision-making and that interact with the decision subject to inform them about their situation before starting a

[21]We use the term *post-hoc* intervention to refer to an intervention that happens once the algorithmic decision is made. For interventions that happen before the algorithmic decision is made, we use the term *ex-ante* intervention.

contestation act. We observed an effect of AI literacy on decision subjects' information needs for sense-making (e.g., the type of information that participants with different levels of AI literacy perceived as actionable).

4.1.1 *T1.1. Strategizing Information Requests.*

Participants in our study developed strategies for deciding which information to request. These strategies depended on participants' ability to make sense of the provided information –or their capacity to look for an independent expert advisor who could help them make sense of the information– and the risks and benefits of issuing an appeal using that information. Participants generally hypothesized two reasons for receiving the letter: (1) they had violated the regulation, or (2) they represented a false positive. In case (1), participants would accept the decision based on the justification and evidence they are shown. In case (2), there were two main situations that participants contemplated: (2a) they had rented their property out but they had a license for it, or (2b) they had not rented their property but the system indicated that they did.

In view of the above, participants, regardless of their AI literacy, unanimously prioritised knowing *why* they got detected by the algorithmic system. The reason behind this was that knowing *why* they got flagged was the first step towards knowing which of the situations they were in and resolving the issue. Participants pointed to the difference between a feature-based explanation, and a decision justification that clearly signals the reasons why a penalty is issued. The provided decision explanation (see Figure 3) did not satisfy their information needs because it did not provide a clear actionable path that could help decision subjects remedy the situation. P16, for instance, complained about the uselessness of receiving a feature-based explanation that points how the system had identified their property as an illegal holiday rental:

> *"Are you telling me that I have illegally rented my house or are you telling me there is a probability of me illegally renting my house? That probability could be based on a thousand things. Tell me the things I have actually missed. There has to be a concrete reasoning behind it, just give me that reasoning. Don't give me these numbers."* (P16)

Many participants additionally wanted to understand the decision basis (i.e., the policy behind the decision) to better discern whether their actions conform to the law or to double-check that the algorithmic decision basis was backed up by relevant policy. Some admitted that they might not have been aware of the regulation and would accept the first warning if this was duly motivated. The option of asking for legal advice to make sense of the decision and the policy was mentioned several times. The willingness to ask for legal advice depended on the required resources and the amount of the penalty.

Among the hypothesized scenarios, it was only in case (2b) that some participants started questioning the algorithmic system beyond the output itself. This was motivated by the difficulties in showing proof of innocence in this particular scenario as compared to the other scenarios. Our participants' AI literacy and experience dealing with fairness in AI affected their interest in knowing *how* the decision-making took place (including *how* the algorithmic system worked) and the perceived actionability of such information. This can be explained by the effect that participants' AI literacy had on the sense-making process that precede a contestation. Participants with *low AI literacy* were mostly uninterested in knowing how the algorithmic system worked because they were not certain about how they would use this information as part of their contestation. *Medium AI literacy* participants were interested in receiving more information about the data used by the system since this would allow them to ask questions related to privacy and bias in data. They were also generally curious to know more about the system due to their technical background but

AI literacy	**Information needs - perceived as actionable**	**Contestation object - *what* to contest**
Low AI literacy	• *Why* of decision output	• Correctness of decision • Legal validity of AI usage • Appropriateness of decision basis • Lack of legitimate proof
Medium AI literacy *High AI literacy* (no experience in AI fairness)	+ Data-related information	+ Bias + Discrimination + Privacy of data
High AI literacy (experience in AI fairness)	+ Decision explanation + Model-related information + Development of the system	+ Explanation weights + Prioritization of data features + Lack of model robustness + Faulty development process

Table 4. Overview of the relationship between AI literacy, information needs and contestation objects. + symbols indicate that the presented items are cumulative (e.g., medium AI literacy participants wanted to know the *why* behind the decision output as well as information related to data). This arrangement shows tendencies we observed in our data and does not necessarily represent one-to-one relations: not all participants from a given subgroup requested all of the information and a few participants requested more information than it is indicated for their subgroup. For a one-to-one relation for each item, see the mapping in the supplementary material.

expressed doubts on how to use this information. For example, when asked about their willingness to know more about the AI system itself, P10 responded:

> *"Myself, because I'm quite a freak, I would [like to know more on a system level]. In general, I don't think people would care. They would be very focused on [fixing] their own problem."* (P10)

The option to contact experts in AI that would help them make sense and act upon information regarding the workings of the AI system was mentioned as an option by some participants with medium AI literacy. Among *high AI literacy* participants, the perceived actionability of AI-related information was further influenced by participants' experience in topics related to AI fairness. This was due to the effect that experience with AI fairness had on participants' ability to identify the subjectiveness of many of the design choices in the development of AI systems and their capacity to use this information as part of their contestation. Participants with high AI literacy who had not previously dealt with fairness-related topics were aligned with medium AI literacy participants and had doubts about how they could use information about the AI workings as part of the contestation. Instead, participants with high AI literacy and experience with AI fairness were willing to know and question aspects related to e.g., AI development. See table 4 for a detailed account of the information that each subgroup of participants deemed important and actionable. As shown in the table, our results indicate that the sense-making process that precedes contestation acts depends on decision subjects' AI literacy and AI fairness experience.

4.1.2 *T1.2. Facilitating Dialogue with Controllers.*

Responses from our interviews indicate that the way in which decision subjects make sense of the provided information is also influenced by the means used for dialogue between decision subjects and human controllers. The communication between controllers and decision subjects turns information into meaningful explanations. This communication also helps clarify technical jargon, a key aspect of the sense-making process according to our participants. One aspect that impacts the communicative effort and joint sense-making between controllers and decision subjects is the communication channel. Participants' preference for communication channels varied based on the stakes and complexity of the decision, who was responsible for the situation, personal experience, and language. However, there was consensus that communication channels should be designed to minimize the friction of engaging in a dialogue.

When it comes to the format in which information is presented, almost all participants wanted the effort to understand the information provided by the municipality to be minimal. This required the provided information to be relevant to their case, concise, simple, and clear. The reason behind this was the need to make the information digestible to different decision subjects, especially those with lower levels of AI literacy. A potential means for satisfying this requirement would be the progressive discovery of information based on relevance. As claimed by P8, this could be achieved through information hierarchies.

> *"Give me a diagram of my case. Then if I want to go in detail on anything in particular, put it all at the end. I would love for everything to be well explained at the end."* (P8)

Many deemed visual explanations (i.e., graphics) or explanatory videos of the decision-making process as appropriate mediums of communication due to their interactivity.

4.2 T2. Social Support in Contestation Acts

The second theme focuses on the contestation act itself. **Contestation acts require the joint effort of decision subjects, street-level bureaucrats acting as reviewers, third parties assigned to decision subjects, and fellow sufferers.** Reviewers are street-level bureaucrats involved in the contestation process [69]. Contestation acts were defined to participants as processes implemented *within* the organization rather than before a tribunal [83]. For this reason, cooperation with legal representatives was not mentioned as a central element of the contestation process (unlike the sense-making). We observed an effect of AI literacy on decision subjects' procedural needs for engaging in acts of contestation (e.g., the need for third parties to compensate for knowledge differentials).

4.2.1 *T2.1. Seeking Organizational Support.*

Reviewers were seen as key actors in supporting decision subjects on a one-to-one basis and in facilitating the act of contestation. Our participants generally preferred a human reviewer over an algorithmic reviewer. The reasons for this were varied. AI was seen as unable to change the output. Humans, in contrast, were seen as more appropriate since they could provide answers beyond frequently asked questions, and they could deal with *grey* areas (i.e., ill-defined situations).

Our participants expressed a general wish for the human reviewer to be cooperative and empathetic during the discussions. Many highlighted the need for a proactive attitude where *"both parties need to be willing to find a solution and a conclusion to the problem"* (P12). The reviewer thus needs to be an active listener as opposed to a *"nameless bureaucrat who doesn't really deal with my issue."* (P7). The reason for this was the wish of decision subjects to feel understood and *"to be heard before being given a warning."* (P21). In contrast to reviewers in existing public decision-making processes, many participants claimed that reviewers for algorithmic decision-making should be experts in AI so that they can effectively accompany them throughout the contestation process.

While participants acknowledged the need for cooperation with human reviewers, a few defined the contestation process as a fight. One of the main reasons for this was the *power differentials* between the decision subject and the reviewer. Power differentials are accentuated with knowledge differentials (e.g. when decision subjects have low AI literacy or when non-native decision subjects do not know the functioning of public administration). Many participants requested a third party (e.g., a watchdog) to mediate the conflict. The third-party could ask questions on decision subjects' behalf and could have information about similar cases. P2, for example, motivated the need to have an independent party involved in the contestation process to deal with information differentials:

> *"I would want a third party. Someone who is equally informed but who did not build the system. Just to have an objective perspective."* (P2)

The third-party should have both legal and technical knowledge (i.e., experience in data science) and should help decision subjects to move forward. A few participants acknowledged that the level of support needed from the third party would depend on the decision subjects' AI literacy, the level of satisfaction with the dialogue they had with the controller during sense-making, and decision subjects' legal knowledge.

4.2.2 ***T2.2. Seeking Peer Support**.* Participants generally prioritized clarifying their own case at the individual level—for all scenarios (1), (2a), and (2b) in section 4.1.1. The possibility of contesting aspects of the algorithmic system—scenario (2b)—, however, was conceived to be more feasible if done collectively. When asked about the possibility of contesting the algorithmic system, P11 mentioned:

> *"If more of us get this letter, then maybe a consortium could be formed, and then through that consortium, we would discredit the AI system. If I was the only one of my social circle getting this letter, I wouldn't immediately go towards discrediting their AI. "* (P11)

Some suggested that using similar cases where the algorithmic system repeatedly made an error could be the basis of the collective contestation. This would be a means for others not to go through the same issues if the system incorrectly flags them. The collective was regarded as *"a place that is organized by citizens, by people that have gone through this"* (P18). Within those previously affected, high AI literacy individuals, or experts with some status (e.g., professors) could be the technical guides to help escalate the situation. According to P6 and P15, attracting the attention of the media and turning the issue into a political matter would be required.

> *"[If] there is a group that we all together try to say [that] this shouldn't work like that, and this becomes a thing, then it could form a very interesting, small nerdy rebellion against the AI system. I think that this is a collective issue, which needs space and people, and attention. (...) It can be like an Anonymous kind of thing, but for AI and for governmental AI systems."* (P15)

Some participants highlighted that a collective could help citizens affected by the system to remedy their situation. A collective would provide decision subjects insights into similar cases. They claimed that this could also enable spotting of error patterns across false positives. This was deemed especially important for people with *low AI literacy* and with no immediate social support structures for providing emotional and procedural help.

4.3 T3. Distributed Responsibility

The third theme highlights the **need for street-level bureaucrats acting as controllers, policy makers, and other members of the public administration to cooperate and to ensure**

appropriate responsibility attribution. We did not observe an effect of AI literacy on claims about responsibility attribution.

4.3.1 ***T3.1. Ensuring Algorithmic Accountability.***

In general terms, participants appreciated and wanted to exercise their right to contest the algorithmic decision but dealing with the consequences of errors made by the algorithmic system was perceived to be unfair. Many mentioned the burden of showing proof of innocence and the effort needed to make sense of the information that would enable them to do so. Overall, there was consensus on the fact that correcting AI's mistakes is not the decision subject's responsibility. If such a burden is put on the decision subject and this represents a false positive, a few participants requested *compensations* for the time wasted and the effort devoted to contesting.

There were several views on whose responsibility it was to contest the system. P15, for example, mentioned that, *"what I would like is the AI to be contested by the employees before they send you the letter."* This would require human controllers to be able to identify such false positives, for which P21 suggested an approach. The suggested workflow would entail: (1) the municipality contacts the individual that has been flagged by the system before any warning is issued, (2) the municipality provides the reasons why they contact the individual, (3) there is a discussion around the reasons why the citizen has been flagged to verify that it is not a false positive, (4) if it turns out to be a false positive, the human reviewer restrains the system from flagging that decision subject again.

When the system is not developed in-house and responsibility is distributed across actors (e.g., dataset creators, model developers, system consumers), P13 pointed to the complexity of attributing responsibility correctly.

> *"If the City Hall outsourced the implementation of the system, then the outsourcing company would be responsible for correcting the system. But the citizen is unaware of that aspect and vis-a-vis the citizen the ultimate responsible is the City Hall. Then, ultimately, the City Hall should take responsibility"* (P13)

Certifying the algorithmic system before deployment was suggested as a means of unburdening the decision subject and ensuring a fair responsibility attribution.

4.3.2 ***T3.2. Fostering Social Transparency.***

Throughout the interviews, we observed that the unique nature of the public administration (e.g., far-reaching impacts, goals of social good) shaped the way in which our participants reacted to the presented decision-making process. On the one hand, because of the nature of public administration, a few participants requested transparency of cooperative activities (e.g., how AI implementation projects take place) in the context of the public administration (i.e., there were requests for *social transparency* [33, 93]). Social transparency within public administration was seen as a pre-requisite for implementing ex-ante contestability mechanisms. This was translated, for instance, in requests for participatory development approaches. To avoid corrective measures, P1 highlighted the importance of *probationary periods.* Probationary periods should be conducted in a way that does not impact ongoing activities and should be used to issue first warnings. P11 suggested that the municipality should consult decision subjects around their preferences towards the system at the early stages of AI development.

> *"There is probably a research team that has time and resources to organize 30-min video calls with each case to have a discussion like this in the early stages. Where they show these slides, and they have the different model architectures and data sources, potential harms, performance. Then I would be more interested."* (P11)

On the other hand, the nature of public administration led some participants to believe that the choices made during the system development were the correct ones. For example, P21 claimed that

"I really assume that they are indeed taking care that the data is good quality." Similarly, because the public administration was the entity behind this system, some assumed that there would be more accountability and diligence when dealing with false positives. A few participants also made comparisons between the public and private spheres. Algorithmic decision-making processes in the public sphere were believed to be more contestable and were considered to have higher ethical standards.

5 Discussion

Our study aimed to generate in-depth *empirical* insights into *decision subjects'* information and procedural needs for meaningful contestability in a *high-risk* decision-making scenario in the public sector (i.e., an illegal holiday rental detection scenario). To this end, we conducted 21 interviews with participants with experience renting their properties out with varying levels of AI literacy. Instead of conceiving their right to autonomy as purely individual self-determination, our results suggest that participants' capacity for contestability was shaped and dependent on their interactions with other actors involved in decision-making. In this section, we summarize our results and position them in existing literature. We then discuss the implications for practice and research of our work.

5.1 Results in Relation to Previous Work

Information Needs for Contestability. Our results show that decision subjects have different strategies for deciding which information to request when contesting an algorithmic decision. These strategies depend on the perceived actionability of the provided information, and the risks and benefits of contesting the decision-making process. Regardless of AI literacy, there is a consensus in prioritizing the *why* (i.e., reasons, proof [44, 82]) behind the decision as a first step towards exercising their right to contest. The extent to which decision subjects want to know *how* the decision-making process took place depends on their AI literacy. It also depends on participants' experience with AI fairness. Especially among subjects with low AI literacy, knowing *how* the decision was made is not a priority. Among those who are interested in knowing *how* the decision was made, and unlike previous work on human-AI collaboration [53] and recommender systems [88], decision subjects do not want to "know everything". They are rather selective in choosing relevant information about the system that could help them contest the decision-making process [1]. This could be due to the differences in purpose (i.e., the aim of contesting *vs.* improving human-AI collaboration *vs.* getting better recommendations) and our participants' intrinsic need for practically helpful information because they are hypothesizing around a contestation scenario. The object and means of contestation (*what* participants in our study want to contest and *how* they want to proceed), in turn, depend on the perceived appropriateness of the information they receive and their ability to understand and use it as part of their contestation. Our findings further suggest that the sense-making process that preceeds a contestation is a cooperative process that participants engage in through expert advice or through dialogue with controllers. The means that enable such dialogue (i.e., communication channel, explanation medium), therefore, also affect the sense-making process. Even if theoretical claims have recognized the importance of justifications [44], or explanations [68] for contestability, there has been a comparatively small emphasis on empirically examining how decision subjects (individually or collectively) make sense of that information and how this empowers them to contest an algorithmic decision.

Procedural Needs for Contestability. Participants in our study request support from the decision-making organization and from peers to deal with the contestation process. This includes the presence of a third party to balance power and knowledge differentials. This suggests that participants perceive how algorithmic systems widen power gaps because of their complexity and opacity [68].

Our results further show that, for contesting aspects of the decision-making process that involve the algorithmic system itself, participants deem *collective action* as more effective than individual appeals. Even if the possibility of collective action was tangentially mentioned in theoretical frameworks [3], the insights from our participants provide detailed descriptions on what the collective could look like (e.g., led by AI experts) and what would define collective success (e.g., media attention, turning the issues into a political matter). When dealing with algorithmic failures (e.g., false positives), our results suggest that individual decision subjects do not want to bear the burden of identifying and contesting such failures. Participants in our study suggest that algorithmic failures should be corrected by *human controllers* (i.e., street-level bureaucrats involved in the first-order decision-making [3]). The very act of having to go through the process of contesting a false positive is considered to be unfair. Our participants also request transparency of the cooperative work that happens among *actors at previous stages of the AI development and deployment pipeline* as well as due responsibility attribution. The need to ensure transparency and due responsibility in a chain of distributed actors is aligned with theoretical claims for contestability by design [3]. It highlights the need to ensure awareness of risks and responsibilities across decision chains.

5.2 Implications for Practice

This section highlights the implications that our work has for public agencies integrating AI systems in decision-making processes.

Building Capacity for Supporting Contestability. The needs we identified for enabling decision subjects to meaningfully engage in acts of contestation are in tension with the challenges for contestability found by Alfrink et al. [2]. Those challenges include limited capacities of civil servants, organizational limits or resource constraints. These tensions indicate that there might be a mismatch between the capacity required to *ideally* address decision subjects' needs during contestation processes and the *reality* of what public administration can offer them in practice based on the available resources. For decision subjects to feel heard and understood, a balance between decision subjects' needs and the allocation of limited resources needs to be found. While participants in our study were generally not against using algorithmic systems for first-order decision-making (this could lead to public savings), they did insist on having a *human* reviewer during the contestation process. However, the organizational challenges of redistributing resources (e.g., economic, human, infrastructural) from the first-order decision-making to the contestation loop cannot be ignored. This is especially true when the algorithmic system suffers from functionality failures [77] in a resource-deficient context [84]. First, it is important to consider *who* is involved in the first-order decision-making and *who* in the contestation loop. How actors involved in different phases of the process have access to each other's information [26], the extent to which there is effective communication between them [77] or the scrutability of the system that mediates the process [68] are all aspects that make organizational change challenging. Furthermore, the relationship between the resources allocated for the current first-order human-led decision-making process and the resources needed for future contestation processes might not be a one-to-one relation. If such algorithms malfunction [77] and human oversight is motivated by legal compliance rather than quality control [2, 40], the harms generated when deployed at scale might multiply. It is, therefore, important to first ensure effective human oversight through e.g., explanations, cognitive forcing functions, or reinforcement learning paradigms [20, 21, 105]. Once appropriate human oversight is ensured, one way to build capacity for contestability would be to partly augment human reviewers' capacities (e.g., through chatbots [66] or methods to detect insincere contestations [8]) while ensuring decision subjects *feel heard*. If human oversight mechanisms are not effective or the option to augment human reviewers' capacities does not allow decision subjects to be heard and to exercise their right to contest automated decisions meaningfully (Article 22(3) of the GDPR [107]), the usage of AI systems might need to be interrupted.

Enabling Collective Action. Our results suggest that participants sought organizational and peer support to engage in acts of contestation. The conception of contestability might, therefore, need to account for the social nature of contestability. One way to do so is through collective contestations. Designing for collective contestability can involve indirect forms of control [2] through *representative bodies* of decision subjects [27, 101]. Examples of collective contestations include the *Contestation Café* suggested by Collins and Redström [27]. The *Contestation Café* [27] is a speculative concept for community contestation, where decision subjects could learn to identify and contest unfair decisions. In a similar vein, *end-user driven audits* [29, 30, 59, 89] use the lived experiences of everyday users of algorithmic systems to uncover harmful algorithmic behaviors, which has, in turn, led to collective contestations (see Shen et al. [89] for a list of examples). An alternative line of work rather explores collective contestations as ex-ante mechanisms by e.g., involving decision subjects in the early stages of the AI design pipeline. This allows decision subjects to get actively involved in crafting the desired algorithmic behavior and in avoiding harmful consequences downstream [3]. If public agencies decide to explore this option, participatory frameworks such as *WeBuildAI* [60] could represent an interesting starting point. In *WeBuildAI* [60], stakeholders–including decision subjects–, can represent their views through computational models that contribute to algorithmic policy creation. For collective action like *Contestation Cafés* [27], end-user-driven audits [89] or WeBuildAI [60] to be of any use, participation is required. Participation, in turn, requires incentives [29] (e.g., available time, interest). An option to promote collectives could be for public administration to (financially) sustain them while ensuring collective action remains independent from the decision-making entity.

Defining Normative Boundaries for Contestability. A number of policy decisions should precede these contestation acts. These include determining what can be contested (both ex-ante and post-hoc), who can contest algorithmic decision-making processes, who is accountable for them, and what type of reviews or scrutiny mechanisms should be put in place [68]. Our work urges policymakers to further define normative boundaries for contestability.

5.3 Implications for Research

This section elaborates on the implications of our work for the CSCW research community.

Characterizing Individual and Collective Sense-Making of Personalized Explanations. According to our results, for a piece of information to be *actionable*, this information needs to be relevant and translatable into an "effective goal-oriented action" [102] (i.e., contesting). The relevance and potential of an information item to be translated into action, in turn, depends on decision subjects' ability to make sense and critically reflect on it to evaluate its appropriateness [85]. Our results, therefore, suggest that personalized, actionable explanations might be needed to address decision subjects' varying information needs for contestability. In contrast to *actionability* in *recourse* (i.e., set of factors that can be changed to obtain the desired outcome [52, 63, 92]), when dealing with contestability that goes beyond the decision outcome (i.e., it concerns the whole lifecycle of the system [3]), there is not one single definition for *actionable information*. There is, therefore, not a single response as to what information empowers decision subjects to meaningfully contest an algorithmic decision-making process [82, 91]. For explanations to be actionable for different decision subjects, they should, therefore, afford varying levels of sufficiency (i.e., content depth) and configuration. This could be operationalized by, for example, implementing explanations with hierarchies of information and varying levels of detail [28] or making explanations interactive [58].

Furthermore, the sense-making process of those explanations is not necessarily an individual process. It is additionally influenced by the actors that decision subjects could contact for help (e.g., legal representatives, human controllers). Further research is needed to know how different decision subjects make sense –individually and collectively– of personalized actionable explanations that are aimed explicitly at enacting contestability and that present varying levels of (1) availability, (2) content, (3) detail, (4) modality (i.e., audio vs. visual), and (5) paradigm (i.e., textual vs. graphical vs. interactive) [111]. Previous work on personalized explanations for recommender systems [96] could represent a good starting point for exploring personalized, actionable explanations *for contestability*. Personalized explanations for contestability will have to navigate the tension between opening algorithmic systems to scrutiny, and the need to align with privacy and confidentiality requirements [115].

Opening Up Sites for Contestation in AI Development and Deployment Pipelines. For contestability to be exercised by stakeholders (other than decision subjects) at earlier stages of the AI pipeline, tools that enable "real-time questioning, curiosity, and scrutiny" [57] of algorithmic systems by human controllers are needed. While some tools are already available that enable the scrutiny of algorithmic systems to surface information about decisions and models (e.g., What-if Tool [108]), further research is necessary to identify the needs of professional human controllers to interactively shape algorithmic behaviour and prevent false positives from repeatedly happening [57]. For due responsibility attribution across *algorithmic supply chains,* [26] tracking and documenting data flows represents the first step towards contestability—documentation which is required by the EU AI Act [36]. Exercising contestability throughout the algorithmic supply chain could, in turn, represent a step towards a deeper engagement with the system [99]. It would help actors distributed across the supply chain not only gain visibility over the supply chain itself, but it would also allow those actors to be attributed due responsibility when required. Enabling contestability throughout algorithmic supply chains [26] faces two main challenges that would benefit from further research. First, documenting discretionary choices made throughout the development and deployment pipeline of algorithmic systems is not straightforward [113]. There is a need to raise awareness around the value-laden (and therefore contestable) nature of "undisclosed yet impactful" [22] choices made throughout the pipeline. There is also a need to provide resources for practitioners to identify and effectively document such choices [9, 43, 70]. We echo prior work [45, 113] and encourage the CSCW community to look into strategies for scaffolding collaborative reflexive practices throughout the AI development and deployment pipelines. Second, even if those choices are acknowledged and documented, different actors across the supply chain might suffer from accountability horizon (i.e., limited capacity for system designers to understand the deployment context and for system consumers to influence its design) [26]. Therefore, legal and institutional mechanisms would be required to ensure visibility and influence over those design choices [26].

6 Caveats and Limitations

In this section, we discuss relevant caveats and report the limitations of our study.

Participant Recruitment. To answer our research question, we sought to generate an in-depth understanding of decision subjects' needs for meaningful contestability and, therefore, decided to conduct qualitative interviews. In line with the Big Q qualitative research paradigm [24], we used purposive sampling to recruit participants that could help us generate nuanced insights into those needs. We, therefore, ensured that we had a diverse pool of participants in terms of AI literacy and ensured that the number of participants with low, medium, and high AI literacy was equally distributed. Among our participants, there was a more prominent representation of two countries (i.e., Netherlands and Spain). Similarly, our interviewees were all highly educated individuals (i.e., all had at least a bachelor's degree) and were used to interacting with digital platforms. Even if these choices are an intrinsic trade-off of Big Q qualitative research [24] in favor of generating in-depth insights, we acknowledge that our study might be subject to *representativeness limitations* [61].

Material Used for the Interviews and Transferability of Results. The letter we used as a prompt in the interview was designed and informed (e.g., penalty, contestation timeframe) by the guidelines that the Dutch public administration follows. Such a choice was made due to the origin of the suggested AI system and its specifications (i.e., the municipality of Amsterdam). The interviews, however, did not necessarily include citizens dealing with the Dutch public administration. Only a few participants mentioned the discrepancies between their experience with public administration communications and the material we presented. They considered this to be an irrelevant detail (e.g., *"They give me a timeframe. I don't care if it's 30 days [contestation timeframe in their residence country] or 6 weeks or whatever."* (P10)). However, we acknowledge that this mismatch might have affected how some other participants engaged in the interview. Similarly, the materials used for the interviews were based on a single case: a risk-scoring scenario for fraud detection within the public sector. We expect our findings to be transferable to other contexts where AI systems are used as part of policy enforcement efforts in the public sector. The transferability of our results to contexts other than policy enforcement support in the public sector will need further verification and should not be fully assumed.

Reflections on External Validity. For exploring the usage of algorithmic systems that have not yet been deployed, previous work has shown that scenario- or vignette-based qualitative methods can be useful instruments [19]. Several studies have also shown that how people react to studies in a "lab-based" environment is a good approximation to how they would react in the real world [110]. Furthermore, our recruitment strategy (i.e., participants who have experience renting their homes out) ensured that our participants had a hypothetical personal stake in the topic, as suggested by Clarke and Braun [24] when using interviews for capturing people's perceptions and understandings about a specific topic. However, a few of our participants indicated that they would not take the time to look at the information sheet (see Section 3.3 for information about the materials we used) if they had not been required to do so as part of the interview. In some cases, it was when participants engaged with the information sheet that they were able to raise concerns about the algorithmic system. This affected the object of contestation (i.e., *what* they wanted to contest). Results might have varied if participants were to contest a real-world algorithmic decision and had not inspected the information sheet.

7 Conclusion

This paper provided in-depth empirical insights into how to operationalize contestability in algorithmic decision-making processes based on decision subjects' information and procedural needs. To this end, we selected an algorithmic system used for identifying illegal holiday rentals as our case and conducted 21 semi-structured interviews with participants with experience renting their homes out and varying levels of AI literacy. Our participants highlighted the need for cooperation during the sense-making process that enables contestability. Strategies that participants used for making sense of the provided information varied based on participants' AI literacy (e.g., unlike the rest of participants, *low AI literacy* participants did not want to know *how* the decision-making process took place) and experience with AI fairness (e.g., only *high AI literacy* participants *with experience in AI fairness* considered information about the development of the system actionable). Our participants additionally asked for support mechanisms both from the decision-making organization and from fellow decision subjects to effectively engage in acts of contestation. Lastly, our participants requested ex-ante interventions to ensure accountability in algorithmic decision-making. Our work suggests that making algorithmic decision-making processes contestable by design is far from a trivial transition from currently available appeal mechanisms for human-led decision-making. There are, instead, several urgent future research directions that deal with the cooperative work behind contestability. We believe that the CSCW community is uniquely positioned to make valuable contributions on this front.

Acknowledgments

We thank the participants in our interviews. We also thank Sofie-Amalie Dideriksen and Uğur Genç for helping us pilot test the interview protocol.

This work was partially supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955990.

References

[1] Kars Alfrink, Ianus Keller, Neelke Doorn, and Gerd Kortuem. 2022. Tensions in transparent urban AI: designing a smart electric vehicle charge point. *AI & SOCIETY* (3 2022). doi:10.1007/s00146-022-01436-9

[2] Kars Alfrink, Ianus Keller, Neelke Doorn, and Gerd Kortuem. 2023. Contestable Camera Cars: A Speculative Design Exploration of Public AI That Is Open and Responsive to Dispute. (2 2023). doi:10.1145/3544548.3580984

[3] Kars Alfrink, Ianus Keller, Gerd Kortuem, and Neelke Doorn. 2022. Contestable AI by Design: Towards a Framework. *Minds and Machines* (8 2022). doi:10.1007/s11023-022-09611-z

[4] Saja Aljuneidi, Wilko Heuten, Markus Tepe, and Susanne Boll. 2023. Did that AI just Charge me a Fine? Citizens' Perceptions of AI-based Discretion in Public Administration. In *Proceedings of the 2023 ACM Conference on Information Technology for Social Good*. ACM, New York, NY, USA, 57–67. doi:10.1145/3582515.3609518

[5] Ali Alkhatib and Michael Bernstein. 2019. Street-Level Algorithms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. doi:10.1145/3290605.3300760

[6] Marco Almada. 2019. Human intervention in automated decision-making. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. ACM, New York, NY, USA, 2–11. doi:10.1145/3322640.3326699

[7] Ariful Islam Anik and Andrea Bunt. 2021. Data-Centric Explanations: Explaining Training Data of Machine Learning Systems to Promote Transparency. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA. doi:10.1145/3411764.3445736

[8] Shubham Atreja, Jane Im, Paul Resnick, and Libby Hemphill. 2023. AppealMod: Shifting Effort from Moderators to Users Making Appeals. (1 2023).

[9] Agathe Balayn, Mireia Yurrita, Jie Yang, and Ujwal Gadiraju. 2023. "Fairness Toolkits, A Checkbox Culture?" On the Factors that Fragment Developer Practices in Handling Algorithmic Harms. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York, NY, USA, 482–495. doi:10.1145/3600211.3604674

[10] Jeffrey Bardzell, Shaowen Bardzell, and Erik Stolterman. 2014. Reading critical designs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1951–1960. doi:10.1145/2556288.2557137

[11] Christine Barter and Emma Renold. 2000. 'I wanna tell you a story': Exploring the application of vignettes in qualitative research with children and young people. *International Journal of Social Research Methodology* 3, 4 (1 2000), 307–323. doi:10.1080/13645570050178594

[12] Gianluca Bei and Filippo Celata. 2023. Challenges and effects of short-term rentals regulation. *Annals of Tourism Research* 101 (7 2023), 103605. doi:10.1016/j.annals.2023.103605

[13] Uladzislau Belavusau and Kristin Henrard. 2019. A Bird's Eye View on EU Anti-Discrimination Law: The Impact of the 2000 Equality Directives. *German Law Journal* 20, 05 (7 2019), 614–636. doi:10.1017/glj.2019.53

[14] Roni Berger. 2015. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qualitative Research* 15, 2 (4 2015), 219–234. doi:10.1177/1468794112468475

[15] R.J. Bies and J. F. Moag. 1986. Interactional Justice: Communication Criteria of Fairness. . *Research on Negotiations in Organizations* 1 (1986), 43–55.

[16] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. 'It's Reducing a Human Being to a Percentage'; Perceptions of Justice in Algorithmic Decisions. (1 2018). doi:10.1145/3173574.3173951

[17] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (1 2006), 77–101. doi:10.1191/1478088706qp063oa

[18] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (3 2021), 37–47. doi:10.1002/capr.12360

[19] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward Algorithmic Accountability in Public Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. doi:10.1145/3290605.3300271

[20] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (4 2021), 1–21. doi:10.1145/3449287

[21] Zana Buçinca, Siddharth Swaroop, Amanda E Paluch, Susan A Murphy, and Krzysztof Z Gajos. 2024. Towards Optimizing Human-Centric Objectives in AI-Assisted Decision-Making With Offline Reinforcement Learning. *arXiv preprint arXiv:2403.05911* (2024).

[22] Scott Allen Cambo and Darren Gergle. 2022. Model Positionality and Computational Reflexivity: Promoting Reflexivity in Data Science. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–19. doi:10.1145/3491102.3501998

[23] C W Churchman. 1967. Free for All. *Management Science* 14, 4 (12 1967), B–141–B–146. doi:10.1287/mnsc.14.4.B141

[24] Victoria Clarke and Virginia Braun. 2013. *Successful qualitative research: A practical guide for beginners*. Sage publications ltd. 1–400 pages.

[25] Victoria Clarke and Virginia Braun. 2021. *Thematic analysis: a practical guide*. SAGE Publications Ltd.

[26] Jennifer Cobbe, Michael Veale, and Jatinder Singh. 2023. Understanding accountability in algorithmic supply chains. In *2023 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 1186–1197. doi:10.1145/3593013.3594073

[27] Robert Collins and Johan Redström. 2022. The Contestation Café. In *Nordic Human-Computer Interaction Conference*. ACM, New York, NY, USA, 1–1. doi:10.1145/3546155.3547290

[28] Anamaria Crisan, Margaret Drouhard, Jesse Vig, and Nazneen Rajani. 2022. Interactive Model Cards: A Human-Centered Approach to Model Documentation. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 427–439. doi:10.1145/3531146.3533108

[29] Wesley Hanwen Deng, Boyuan Guo, Alicia Devrio, Hong Shen, Motahhare Eslami, and Kenneth Holstein. 2023. Understanding Practices, Challenges, and Opportunities for User-Engaged Algorithm Auditing in Industry Practice. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–18. doi:10.1145/3544548.3581026

[30] Alicia DeVos, Aditi Dhabalia, Hong Shen, Kenneth Holstein, and Motahhare Eslami. 2022. Toward User-Driven Algorithm Auditing: Investigating users' strategies for uncovering harmful algorithmic behavior. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–19. doi:10.1145/3491102.3517441

[31] Jonathan Dodge, Q. Vera Liao, Yunfeng Zhang, Rachel K. E. Bellamy, and Casey Dugan. 2019. Explaining Models: An Empirical Study of How Explanations Impact Fairness Judgment. (1 2019). doi:10.1145/3301275.3302310

[32] Tim Draws, Alisa Rieger, Oana Inel, Ujwal Gadiraju, and Nava Tintarev. 2021. A Checklist to Combat Cognitive Biases in Crowdsourcing. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 9, 1 (10 2021), 48–59. https://ojs.aaai.org/index.php/HCOMP/article/view/18939

[33] Upol Ehsan, Q. Vera Liao, Michael Muller, Mark O. Riedl, and Justin D. Weisz. 2021. Expanding Explainability: Towards Social Transparency in AI systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–19. doi:10.1145/3411764.3445188

[34] Upol Ehsan, Samir Passi, Q. Vera Liao, Larry Chan, I-Hsiang Lee, Michael Muller, and Mark O. Riedl. 2021. The Who in Explainable AI: How AI Background Shapes Perceptions of AI Explanations. (7 2021).

[35] Carsten Eickhoff. 2018. Cognitive Biases in Crowdsourcing. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, USA, 162–170. doi:10.1145/3159652.3159654

[36] European Commission. 2021. *Proposal for regulation of the European parliament and of the council - Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain Union legislative acts.*

[37] European Commission. 2021. Regulation of the European Parliament and of the council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

[38] Luciano Floridi. 2020. Artificial Intelligence as a Public Service: Learning from Amsterdam and Helsinki. *Philosophy & Technology* 33, 4 (12 2020), 541–546. doi:10.1007/s13347-020-00434-3

[39] Christopher Frauenberger. 2016. Critical Realist HCI. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, 341–351. doi:10.1145/2851581.2892569

[40] Ben Green. 2022. The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review* 45 (7 2022), 105681. doi:10.1016/j.clsr.2022.105681

[41] J. Greenberg. 1993. The social side of fairness: Interpersonal and informational classes of organizational justice. *Justice in the workplace: Approaching fairness in human resource management.* (1993), 79–103.

[42] Tamarinde L. Haven and Dr. Leonie Van Grootel. 2019. Preregistering qualitative research. *Accountability in Research* 26 (4 2019), 229–244. Issue 3. doi:10.1080/08989621.2019.1580147

[43] Amy Heger, Elizabeth B. Marquis, Mihaela Vorvoreanu, Hanna Wallach, and Jennifer Wortman Vaughan. 2022. Understanding Machine Learning Practitioners' Data Documentation Perceptions, Needs, Challenges, and Desiderata. (6 2022).

[44] Clément Henin and Daniel Le Métayer. 2021. Beyond explainability: justifiability and contestability of algorithmic decision systems. *AI & SOCIETY* (7 2021). doi:10.1007/s00146-021-01251-8

[45] Simon David Hirsbrunner, Michael Tebbe, and Claudia Müller-Birn. 2022. From critical technical practice to reflexive data science. *Convergence: The International Journal of Research into New Media Technologies* (11 2022), 135485652211322. doi:10.1177/13548565221132243

[46] Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E. Imel, and David C. Atkins. 2017. Designing Contestability. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM, New York, NY, USA. doi:10.1145/3064663.3064703

[47] Rhidian Hughes. 1998. Considering the Vignette Technique and its Application to a Study of Drug Injecting and HIV Risk and Safer Behaviour. *Sociology of Health & Illness* 20, 3 (5 1998), 381–400. doi:10.1111/1467-9566.00107

[48] Shalmali Joshi, Oluwasanmi Koyejo, Warut Vijitbenjaronk, Been Kim, and Joydeep Ghosh. 2019. Towards Realistic Individual Recourse and Actionable Explanations in Black-Box Decision Making Systems. (7 2019).

[49] Ekaterina Jussupow, Izak Benbasat, and Armin Heinzl. 2020. Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. In *Proceedings of the 28th European Conference on Information Systems (ECIS)*.

[50] Margot E Kaminski and Jennifer M Urban. 2021. THE RIGHT TO CONTEST AI. *Columbia Law Review* 121, 7 (2021), 1957–2048. https://www-jstor-org.tudelft.idm.oclc.org/stable/27083420

[51] Amir-Hossein Karimi, Gilles Barthe, Bernhard Schölkopf, and Isabel Valera. 2022. A survey of algorithmic recourse:contrastive explanations and consequential recommendations. *Comput. Surveys* (4 2022). doi:10.1145/3527848

[52] Amir-Hossein Karimi, Bernhard Schölkopf, and Isabel Valera. 2021. Algorithmic Recourse. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 353–362. doi:10.1145/3442188.3445899

[53] Sunnie S. Y. Kim, Elizabeth Anne Watkins, Olga Russakovsky, Ruth Fong, and Andrés Monroy-Hernández. 2023. "Help Me Help the AI": Understanding How Explainability Can Support Human-AI Interaction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–17. doi:10.1145/3544548.3581001

[54] Sunnie S. Y. Kim, Elizabeth Anne Watkins, Olga Russakovsky, Ruth Fong, and Andrés Monroy-Hernández. 2023. Humans, AI, and Context: Understanding End-Users' Trust in a Real-World Computer Vision Application. In *2023 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 77–88. doi:10.1145/3593013.3593978

[55] Daniel Kluttz, Nitin Kohli, and Deirdre K. Mulligan. 2018. Contestability and Professionals: From Explanations to Engagement with Algorithmic Systems. *SSRN Electronic Journal* (2018). doi:10.2139/ssrn.3311894

[56] Daniel Kluttz and Deirdre K. Mulligan. 2019. Automated decision support technologies and the legal profession. *SSRN Electronic Journal* (2019). doi:10.2139/ssrn.3443063

[57] Daniel N Kluttz, Nitin Kohli, and Deirdre K Mulligan. 2022. Shaping our tools: Contestability as a means to promote responsible algorithmic decision making in the professions. In *Ethics of Data and Analytics*. Auerbach Publications, 420–428.

[58] Todd Kulesza, Margaret Burnett, Weng-Keen Wong, and Simone Stumpf. 2015. Principles of Explanatory Debugging to Personalize Interactive Machine Learning. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 126–137. doi:10.1145/2678025.2701399

[59] Michelle S. Lam, Mitchell L. Gordon, Danaë Metaxa, Jeffrey T. Hancock, James A. Landay, and Michael S. Bernstein. 2022. End-User Audits: A System Empowering Communities to Lead Large-Scale Investigations of Harmful Algorithmic Behavior. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022), 1–34. doi:10.1145/3555625

[60] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D Procaccia. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (11 2019). doi:10.1145/3359283

[61] Min Kyung Lee and Katherine Rich. 2021. Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. doi:10.1145/3411764.3445570

[62] Q. Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. (1 2020). doi:10.1145/3313831.3376590

[63] Q. Vera Liao, Yunfeng Zhang, Ronny Luss, Finale Doshi-Velez, and Amit Dhurandhar. 2022. Connecting Algorithmic Research and Usage Contexts: A Perspective of Contextualized Evaluation for Explainable AI. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 10, 1 (10 2022), 147–159. doi:10.1609/hcomp.v10i1.21995

[64] Charles Lindblom. 2018. The science of "muddling through". In *Classic readings in urban planning*. Routledge, 31–40.

- [65] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30 (2017).
- [66] Henrietta Lyons, Tim Miller, and Eduardo Velloso. 2023. Algorithmic Decisions, Desire for Control, and the Preference for Human Review over Algorithmic Review. In *2023 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 764–774. doi:10.1145/3593013.3594041
- [67] Henrietta Lyons, Eduardo Velloso, and Tim Miller. 2019. Designing for Contestation: Insights from Administrative Law. In *2019 Workshop on Contestability in Algorithmic Systems at CSCW Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '19)*.
- [68] Henrietta Lyons, Eduardo Velloso, and Tim Miller. 2021. Conceptualising Contestability: Perspectives on Contesting Algorithmic Decisions. (2 2021). doi:10.1145/3449180
- [69] Henrietta Lyons, Senuri Wijenayake, Tim Miller, and Eduardo Velloso. 2022. What's the Appeal? Perceptions of Review Processes for Algorithmic Decisions. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. doi:10.1145/3491102.3517606
- [70] Michael Madaio, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 2022. Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (3 2022), 1–26. doi:10.1145/3512899
- [71] Michael A. Madaio, Luke Stark, Jennifer Wortman Vaughan, and Hanna Wallach. 2020. Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. doi:10.1145/3313831.3376445
- [72] Anna Madill, Abbie Jordan, and Caroline Shirley. 2000. Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology* 91, 1 (2 2000), 1–20. doi:10.1348/000712600161646
- [73] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (2 2019), 1–38. doi:10.1016/j.artint.2018.07.007
- [74] Shirley Nieuwland and Rianne van Melik. 2020. Regulating Airbnb: how cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism* 23, 7 (4 2020), 811–825. doi:10.1080/13683500.2018.1504899
- [75] Christina A. Pan, Sahil Yakhmi, Tara P. Iyer, Evan Strasnick, Amy X. Zhang, and Michael S. Bernstein. 2022. Comparing the Perceived Legitimacy of Content Moderation Processes: Contractors, Algorithms, Expert Panels, and Digital Juries. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (3 2022), 1–31. doi:10.1145/3512929
- [76] Virgil Pavlu, Shahzad Rajput, Peter B. Golbus, and Javed A. Aslam. 2012. IR system evaluation using nugget-based test collections. In *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, New York, NY, USA, 393–402. doi:10.1145/2124295.2124343
- [77] Inioluwa Deborah Raji, I. Elizabeth Kumar, Aaron Horowitz, and Andrew Selbst. 2022. The Fallacy of AI Functionality. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 959–972. doi:10.1145/3531146.3533158
- [78] Peyman Rasouli and Ingrid Chieh Yu. 2022. CARE: coherent actionable recourse based on sound counterfactual explanations. *International Journal of Data Science and Analytics* (9 2022). doi:10.1007/s41060-022-00365-6
- [79] Horst W. J. Rittel and Melvin M. Webber. 1973. Dilemmas in a general theory of planning. *Policy Sciences* 4, 2 (6 1973), 155–169. doi:10.1007/BF01405730
- [80] Alan Rubel, Clinton Castro, and Adam Pham. 2021. *Algorithms and Autonomy*. Cambridge University Press. doi:10.1017/9781108895057
- [81] Helen Sampson and Idar Alfred Johannessen. 2020. Turning on the tap: the benefits of using 'real-life' vignettes in qualitative research interviews. *Qualitative Research* 20, 1 (2 2020), 56–72. doi:10.1177/1468794118816618
- [82] Claudio Sarra. 2020. Put Dialectics into the Machine: Protection against Automatic-decision-making through a Deeper Understanding of Contestability by Design. *Global Jurist* 20, 3 (10 2020). doi:10.1515/gj-2020-0003
- [83] Claudio Sarra et al. 2020. Defenceless? An analytical inquiry into the right to contest fully automated decisions in the GDPR. *An Anthology of Law* (2020), 235–252.
- [84] Devansh Saxena, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2021. A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021), 1–41. doi:10.1145/3476089
- [85] Timothée Schmude, Laura Koesten, Torsten Möller, and Sebastian Tschiatschek. 2023. On the Impact of Explanations on Understanding of Algorithmic Decision-Making. In *2023 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 959–970. doi:10.1145/3593013.3594054
- [86] Jakob Schoeffer, Niklas Kuehl, and Yvette Machowski. 2022. "There Is Not Enough Information": On the Effects of Explanations on Perceptions of Informational Fairness and Trustworthiness in Automated Decision-Making. (5 2022). doi:10.1145/3531146.3533218

[87] Andrew D Selbst and Julia Powles. 2017. Meaningful information and the right to explanation. *International Data Privacy Law* 7, 4 (11 2017), 233–242. doi:10.1093/idpl/ipx022
[88] Ruoxi Shang, K. J. Kevin Feng, and Chirag Shah. 2022. Why Am I Not Seeing It? Understanding Users' Needs for Counterfactual Explanations in Everyday Recommendations. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 1330–1340. doi:10.1145/3531146.3533189
[89] Hong Shen, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors. (5 2021). doi:10.1145/3479577
[90] Herbert A Simon. 1990. Bounded rationality. *Utility and probability* (1990), 15–18.
[91] Ronal Singh, Tim Miller, Henrietta Lyons, Liz Sonenberg, Eduardo Velloso, Frank Vetere, Piers Howe, and Paul Dourish. 2023. Directive Explanations for Actionable Explainability in Machine Learning Applications. *ACM Transactions on Interactive Intelligent Systems* (1 2023). doi:10.1145/3579363
[92] Kacper Sokol and Peter Flach. 2020. Explainability fact sheets. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA. doi:10.1145/3351095.3372870
[93] H. Colleen Stuart, Laura Dabbish, Sara Kiesler, Peter Kinnaird, and Ruogu Kang. 2012. Social transparency in networked information exchange. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, New York, NY, USA, 451–460. doi:10.1145/2145204.2145275
[94] Jacob Kraemer Tebes. 2005. Community science, philosophy of science, and the practice of research. *American journal of community psychology* 35, 3-4 (2005), 213–230.
[95] J. W. Thibaut and L. Walker. 1975. Procedural Justice: A Psychological Analysis. *L. Erlbaum Associates, Hillsdale.* (1975).
[96] Nava Tintarev and Judith Masthoff. 2007. Effective explanations of recommendations. In *Proceedings of the 2007 ACM conference on Recommender systems*. ACM, New York, NY, USA, 153–156. doi:10.1145/1297231.1297259
[97] Tom R. Tyler. 2003. Procedural justice, legitimacy, and the effective rule of law. *Crime and justice* 30 (2003), 283–357.
[98] Berk Ustun, Alexander Spangher, and Yang Liu. 2019. Actionable Recourse in Linear Classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 10–19. doi:10.1145/3287560.3287566
[99] Kristen Vaccaro, Karrie Karahalios, Deirdre K. Mulligan, Daniel Kluttz, and Tad Hirsch. 2019. Contestability in Algorithmic Systems. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 523–527. doi:10.1145/3311957.3359435
[100] Kristen Vaccaro, Christian Sandvig, and Karrie Karahalios. 2020. "At the End of the Day Facebook Does What ItWants". *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (10 2020), 1–22. doi:10.1145/3415238
[101] Kristen Vaccaro, Ziang Xiao, Kevin Hamilton, and Karrie Karahalios. 2021. Contestability For Content Moderation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021), 1–28. doi:10.1145/3476059
[102] Orlin Vakarelov and Kenneth Rogerson. 2020. The Transparency Game: Government Information, Access, and Actionability. *Philosophy & Technology* 33, 1 (3 2020), 71–92. doi:10.1007/s13347-019-0340-z
[103] Niels van Berkel, Jorge Goncalves, Daniel Russo, Simo Hosio, and Mikael B. Skov. 2021. Effect of Information Presentation on Fairness Perceptions of Machine Learning Predictors. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. doi:10.1145/3411764.3445365
[104] Arnaud Van Looveren and Janis Klaise. 2021. Interpretable Counterfactual Explanations Guided by Prototypes. 650–665. doi:10.1007/978-3-030-86520-7{_}40
[105] Helena Vasconcelos, Matthew Jörke, Madeleine Grunde-McLaughlin, Tobias Gerstenberg, Michael S. Bernstein, and Ranjay Krishna. 2023. Explanations Can Reduce Overreliance on AI Systems During Decision-Making. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (4 2023), 1–38. doi:10.1145/3579605
[106] Suresh Venkatasubramanian and Mark Alfano. 2020. The philosophical basis of algorithmic recourse. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 284–293. doi:10.1145/3351095.3372876
[107] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. (11 2017).
[108] James Wexler, Mahima Pushkarna, Tolga Bolukbasi, Martin Wattenberg, Fernanda Viegas, and Jimbo Wilson. 2019. The What-If Tool: Interactive Probing of Machine Learning Models. (7 2019). doi:10.1109/TVCG.2019.2934619
[109] Carla Willig. 2013. *EBOOK: introducing qualitative research in psychology*. McGraw-hill education (UK).
[110] S. Woods, M. Walters, Kheng Lee Koay, and K. Dautenhahn. 2006. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *9th IEEE International Workshop on Advanced Motion Control, 2006.* IEEE, 750–755. doi:10.1109/AMC.2006.1631754
[111] Xuhai Xu, Mengjie Yu, Tanya R. Jonker, Kashyap Todi, Feiyu Lu, Xun Qian, João Marcelo Evangelista Belo, Tianyi Wang, Michelle Li, Aran Mun, Te-Yen Wu, Junxiao Shen, Ting Zhang, Narine Kokhlikyan, Fulton Wang, Paul Sorenson, Sophie Kahyun Kim, and Hrvoje Benko. 2023. XAIR: A Framework of Explainable AI in Augmented Reality. (3 2023).

doi:10.1145/3544548.3581500
[112] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. doi:10.1145/3313831.3376301
[113] Mireia Yurrita, Agathe Balayn, and Ujwal Gadiraju. 2023. Generating Process-Centric Explanations to Enable Contestability in Algorithmic Decision-Making: Challenges and Opportunities. In *2023 Human-Centered XAI Workshop at CHI Conference on Human Factors in Computing Systems (CHI '23)*.
[114] Mireia Yurrita, Tim Draws, Agathe Balayn, Dave Murray-Rust, Nava Tintarev, and Alessandro Bozzon. 2023. Disentangling Fairness Perceptions in Algorithmic Decision-Making: the Effects of Explanations, Human Oversight, and Contestability. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–21. doi:10.1145/3544548.3581161
[115] Mireia Yurrita, Dave Murray-Rust, Agathe Balayn, and Alessandro Bozzon. 2022. Towards a multi-stakeholder value-based assessment framework for algorithmic systems. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 535–563. doi:10.1145/3531146.3533118
[116] Anneke Zuiderwijk, Yu-Che Chen, and Fadi Salem. 2021. Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly* 38, 3 (7 2021), 101577. doi:10.1016/j.giq.2021.101577

Received January 2024; revised October 2024; accepted February 2025