"How I Know For Sure": People's Perspectives on Solely Automated Decision-Making (SADM)

Smirity Kaushik[1], Yaxing Yao[2], Pierre Dewitte[3], Yang Wang[1]
[1]University of Illinois at Urbana-Champaign [2]University of Maryland, Baltimore County
[3]Katholieke Universiteit Leuven Centre for IT & IP
{smirity2, yvw}@illinois.edu, {yaxingya}@umbc.edu, {pierre.dewitte}@kuleuven.be

Abstract

Algorithms are used to make automated decisions that can affect individuals in numerous domains. The General Data Protection Regulation (GDPR) of the European Union (EU) has granted citizens some rights regarding solely automated decision-making (SADM), including obtaining an explanation of such processing. It is unclear, however, how organizations should support people in effectively exercising such rights. We conducted an online survey to understand people's perspectives on SADM. We found that our respondents had several misunderstandings about the SADM right, such as opt-out of SADM ahead of time. We also identified various attributes of SADM that our respondents desired to understand, including new attributes (e.g., actionable information about what they can practically do to improve future decision outcomes) not covered by implementation guidelines of the GDPR. Our respondents also anticipated many challenges with SADM, including not knowing when SADM is applied to them. We discuss design implications of our results on how to support people in coping with SADM, for instance, the design of icons to represent SADM processing and explanation templates that cover a common set of attributes and can be personalized to explain a specific SADM decision.

1 Introduction

From job applications to insurance premiums to targeted ads, algorithms have increasingly been used to make automated decisions that can affect individuals [16,66]. While using algorithms to automatically make decisions about individuals may allow companies to increase efficiency and save resources, they also pose significant threats to individuals and society such as privacy violations, social segregation, discrimination, and unjustified denials of service [4, 25, 28, 60, 62]. Furthermore, these automated systems often remain opaque to public scrutiny and understanding, thus leading to a lack of decision acceptance and trustworthiness in these algorithmic practices.

The SADM right. One key aspect of making automated decision-making fair, accountable, and transparent is to provide sensible explanations of these algorithmic decisions to individuals who might be affected. For instance, the European Union (EU) General Data Protection Regulation (GDPR), which entered into force on May 25 2018, defines such an algorithmic decision as "decision based solely on automated processing, including profiling, which produces legal effects concerning the data subject or similarly significantly affects him or her" [70]. Among other citizen rights, the law allows citizens (1) to obtain an explanation of the logic involved as well as the significance and the envisaged consequence of such processing, and (2) to request human intervention, express a point of view, and contest the decision [70]. Since these rules are closely related and to simplify the reference to these rules, we coined an umbrella term "Right against solely automated decision-making" (or the *SADM right*).

GDPR background. The GDPR applies to (1) companies that have an establishment in the EU and (2) companies not based in the EU but offer goods and services to people living in the EU or monitor their behavior if that behavior occurs in the EU [70]. For instance, if an American organization (e.g., company or university) has employees, regardless of their citizenship, who live in the EU (e.g., university staff who work in the study abroad program in France), then the US organization needs to comply with GDPR for those employees. The GDPR and its underlying principles have also substantially impacted other privacy legislation around the world. For instance, Brazil and India are adopting GDPR-like legislation. The UK enacted EU GDPR's requirements into a UK law (Data Protection Act 2018 [17]) and later amended it in 2019 to form a new UK-specific data protec-

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021.
August 8–10, 2021, Virtual Conference.

tion regulation (UK-GDPR [18]) for post-Brexit transition. The UK has also provided guidelines for explaining AI-based decisions or AI-assisted decisions where humans and AI are both involved [69]. In part inspired by the GDPR, the US has also taken significant steps to improve transparency and consumer privacy with state legislations such as California Consumer Privacy Act (2018) [56], California Privacy Rights Act (2020) [52] and proposals of an omnibus federal privacy law [45].

Research motivations. Despite the potential benefits of GDPR and the SADM right more specifically, it is unclear how people perceive SADM or how organizations should support people in exercising this right. This is in part because GDPR was not very specific about this right, which is still open to interpretation regarding its implementation. It is also why the EU A29WP Working Party (now replaced by the European Data Protection Board) has been creating more actionable guidelines on SADM [26]. Moreover, it is increasingly recognized that understanding citizens' perspectives on SADM is needed. For instance, both legal and HCI scholars [6, 29, 38, 55, 57, 63] have advocated for seeking citizens' inputs as crucial and timely in informing the refinement of these actionable guidelines and the ways in which companies can effectively support citizens in exercising the SADM right. Despite the need and importance of understanding citizens' perspectives on the SADM right, there is a lack of empirical research on this topic. Our research aims to help fill this gap.

Research questions. Whether organizations can provide people an effective explanation of their SADM practices is a crucial element of their ethical use of algorithms and cultivating consumer trust in SADM systems. However, what people consider as constituents of an effective explanation of SADM is an open question. Given the broad impact of GDPR on algorithms, people, and ethics, our study contributes to the understanding of people's expectations of the SADM right and the design of socio-technical mechanisms that empower people to exercise this right. Our long-term goal is to design such mechanisms. To inform such design, this paper focuses on three research questions:

- **RQ1:** What are people's understandings of SADM right?
- **RQ2:** What aspects of SADM do people want to know?
- **RQ3:** What challenges do people anticipate in exercising the SADM right?

RQ1 can elucidate how people perceive and expect to exercise the SADM right. RQ2 can provide a baseline list of aspects that companies can consider including in the explanations of their SADM practices. RQ3 can identify peoples' anticipated challenges of exercising the SADM right. Together, answering these questions can inform future design to support people's needs regarding SADM.

Study and findings. To answer our research questions, we conducted an online survey with 392 respondents from the UK and the US. Our research approach gives voice to ordinary citizens' perspectives. It is akin to governments conducting "citizen juries" that directly inquire and incorporate people's inputs on important public policy issues (e.g., [68]). We deliberately chose one country (UK) where the SADM right is directly supported while another (US) that currently does not. It could help us to explore whether they would have drastically different views on SADM. This was found not the case in our study, implying that our respondents shared common expectations of SADM systems and their explanations.

For RQ1, our respondents had many misunderstandings about the SADM right, e.g., incorrectly assuming that the right allows them to opt out of SADM ahead of time or that they could deny the use of personal information for SADM processing. For RQ2, our respondents desired to know more about SADM than what policy-makers have suggested organizations provide (e.g., type of information, source of information, logic used). For instance, our respondents expected to receive personalized explanations including factors considered. It indicated a stronger need of the respondents to seek justification for the decisions made, especially the negative ones. For RQ3, our respondents anticipated a wide variety of challenges. Some of these were unique to SADM, such as the difficulty for people to know when they are subject to SADM.

Research contributions. This research makes two primary contributions. First, it has many novel empirical results on people's perspectives on the SADM and the related right. Specifically, it uncovers people's misunderstandings of the right, which will hinder their effective exercise of the right to protect themselves. The study also identifies attributes of SADM that people want to understand beyond what has been recommended by policymakers. Our research also uncovers some unique challenges anticipated by our participants to exercise the SADM right. Second, we propose several design implications, based on our study results, for organizations to support citizens exercise this right. For instance, designing and using (standardized) icons to represent SADM processing, personalized explanation templates to explain SADM outcomes, and SADM sandboxes that allow users to explore the SADM systems to improve future decisions about them.

2 Related Work

In this section, we present the relevant literature on people's perceptions of algorithmic accountability, fairness, methods of creating explanations of AI systems, and people's challenges with legal concepts (with a focus on privacy policies).

2.1 Perceptions of Algorithmic Accountability

Algorithmic accountability has recently received a lot of traction with the increased use of algorithms in high-impact domains and the changes in the regulatory landscape, such as the

implementation of the EU's GDPR. Bovens describe accountability as 'a relationship between an actor and a forum, in which an actor must explain and to justify his or her conduct, the forum can pose questions and pass judgment, and the actor may face consequences' [7]. Recent scholarship has conceptualized algorithmic accountability, building on the traditional accountability literature. For instance, Wieringa [72] adapted Bovens's [7] widely accepted definition of accountability in the context of algorithmic accountability. The actors (i.e., algorithm developers, decision-makers) are accountable to explain or justify the algorithmic decisions (account) in a forum that can question these actors, often with consequences [72]. Additionally, it requires a perspective, i.e., identifying what needs to be accounted for in the algorithmic system [72]. For instance, Coglianese and Lehr [14] distinguished actors by their roles to determine the appropriate actor for a particular situation. They noted that it is important to identify 'who within an agency actually wields algorithm-specifying power' [14]. According to Kolkman [30], a forum provides a platform for users to understand the decision-making process and to engage with it. It can further impose consequences on the actors. The GDPR [1] through its SADM right provides that forum to an individual citizen in the context of algorithmic accountability and further imposes legal accountability (consequences) for actors. From a technical perspective, Kroll et al. [41] linked algorithmic accountability to a system's life cycle, identifying two possible approaches to algorithmic accountability: ex-ante (before the decision is made) and ex-post (after the decision is made). Neyland [51] argued that accountability should be considered throughout different stages of algorithmic system, i.e., design, implementation, and evaluation, as accountability is a shared responsibility between designers, developers, and users throughout the system's life cycle. Finally, several researchers [10, 13, 14] also argued that algorithmic accountability has a direct relationship to the measure of human involvement. For instance, in case of a human-out-of-the-loop system (or a SADM system), the degree of accountability increases manifolds as there is no human oversight. Our work was partly motivated by the notion of algorithmic accountability, which underpins the SADM right, making designers of SADM systems responsible for explaining and justifying their systems and practices.

2.2 Perceptions of Algorithmic Fairness

The concept of fairness is vast and ambiguous and is used differently across disciplines [48]. Fairness broadly refers to an equitable outcome that can be justified reasonably for a purpose within a context or domain. It further includes the dimension of unfairness which elaborates on what and who is considered capable of violating fairness. Fairness also seeks to clarify who is to be protected and where such protection can be operationalized [48]. Recent work has explored algorithmic fairness in different ways such as building fair decision-making algorithms, determining peoples' perceptions of fairness in these systems. For this paper, we explored prior literature focusing on people's perceptions of algorithmic fairness in general and in specific application domains. For instance, Grgic-Hlaca et al. examined why people perceive the use of certain features as unfair in making decisions about individuals in general [31]. They proposed that people's unfairness concerns are multi-dimensional, based on various aspects such as the relevance, volitionality, and reliability of decision and moral judgment. Woodruff et al. explored the impact of algorithmic bias on marginalized groups based on demographic features such as race [73]. Other scholars have analyzed people's perceptions of algorithmic fairness in specific domains, such as real estate and finance. For instance, Lee and Baykal investigated people's perceptions of fair division algorithms (e.g., those designed to divide rent among tenants) compared to discussion-based group decision-making methods [42]. They found that participants perceived the algorithmic decisions to be less fair than group-based decisions because the former did not account for people's social behavior. Saxena et al. investigated ordinary people's attitude toward three notions of individual fairness in the context of loan decisions [61]. They found that people tend to prefer calibrated fairness which selects individuals in proportion to their merit.

Prior research has also examined algorithmic fairness from experts' perspectives. For instance, Veale et al. interviewed public sector machine learning practitioners regarding the challenges of incorporating public values into their work [71]. They found a disconnect between organizational realities and current research into algorithmic fairness. They proposed incorporating domain knowledge by designing usable privacy tools aimed at private sector managers and public sector bureaucrats. Similarly, Holstein et al. conducted a systematic investigation of commercial product teams' challenges in developing fairer machine learning systems [34]. It highlights the disconnection between the challenges faced by teams in practice and the proposed solutions in the literature review. Research suggests that people care about the fairness of algorithms as well as potential discrimination and biases when companies make decisions about them. Our study helps to understand whether people would consider fairness issues such as biases in the context of SADM.

2.3 Transparent and Explainable AI

The algorithmic black box makes it difficult for users to know how an algorithm works, mainly because the information is either of a certain level of secrecy or intellectual property or too complicated for users to understand [64]. The principle of transparency is related to such a black box. It refers to provide people with the details of knowledge/information that a system gains from its users implicitly. Such details may include how the service works, the potential consequences and

other types of data management (e.g., sensible data) [2]. For example, in the 2016 US Presidential Election, algorithmic transparency on Facebook became a key issue to "end the profiling" [11]. Research has been arguing that practicing the principle of transparency may have a significant impact on people's knowledge and behaviors. Lee and Boynton stated that people are more like to use a system properly and form a sense of trust toward the system designer and developers if they understand how the systems work [43]. The notion of transparency is very relevant to our present work because explanation of SADM is a form of transparency.

Improving transparency of AI systems has been an active area of research in the AI and machine learning community. There is a growing body of work on creating explanations of AI systems or decisions/predictions to improve their transparency. Hu et al. presents a survey of existing methods [35], which mainly differ by two dimensions: scope of the explanation (global vs. local) and how the explanation is generated (intrinsic vs. post-hoc). Global explanations focus on how the whole model/system works, whereas local explanations describe specific decisions made by the model. For instance, 'model-agnostic' approaches such as LIME (Local Interpretable Model-Agnostic Explanations) explain a specific prediction of an algorithm by learning a local model around that predicted value [59]. However, these approaches are primarily helpful in supporting experts (e.g., data analysts) [9]. Intrinsic explanations are built-in part of the model. For instance, decision tree rules [44] are a built-in aspect of the decision tree, which lends itself to easy human interpretation. In comparison, deep learning models (artificial neural networks) are often difficult to explain the logic behind these models. Thus, researchers have created (post-hoc) explanations after these models have been built.

From a policy perspective, Doshi-Velez and Kortz argue that an explanation of an algorithm for end-users requires similar level of accountability that is ascribed to the human decision-makers. This may be achieved by applying certain technical considerations such as using local explanations to reach the outcome without divulging company trade-secrets [21]. Selbst and Barcos [63] explore the use of existing laws to fix the interpretability challenge of machine learning algorithms by focusing on not only their logic but also their fairness. From an HCI perspective, Binns et al. argue that end-users expect system explanations to be similar to those from human decision makers [5]. This is because the end-users apply similar perceptions of justice to the context of automated decision-making that is applicable in human decision-making [5]. Eslami et al. conducted a qualitative study to explore end-users' perceptions of personal information used to make targeted ads [24]. They found that increased visibility of inferences made about users can lead to "algorithm disillusionment," the idea that users realize the limitations of those algorithms which they thought to be perfect before [24]. In another study, Kizilcec [40] found that over-explanation i.e., supplementing the procedural explanation of the grade-adjusting algorithm with the outcome-specific information to increase the transparency of the algorithm further led to students' distrust in the system. There should be a balance between lack of explanation and over-explanation of the algorithm process to cultivate user trust.

2.4 People's Challenges with Privacy Policies

Since the SADM right has legal meanings, we also looked at prior research that shows how ordinary people may struggle with legal concepts and documents. In the domain of privacy, these issues have been around for a long time [12]. For instance,people struggle with privacy policies, a form of explanation that describes an organization's data/privacy practices and are required in the US. Grossklags and Goods point out that privacy policies are often unstructured, jargon-filled, and thus difficult to read and comprehend [32]. Prior work has also shown less than 1% of the general population read these documents [3].

2.5 Summary

These lines of work suggest that people may have multi-dimensional perspectives on automated decision-making, and they may encounter challenges with legal concepts. The prior literature also calls for research on human-centered perspectives on SADM to provide accountable, fair, and transparent, yet balanced explanation to automated decision-making. Our study helps fill the gap by investigating people's perspectives of SADM, such as their understandings and expectations of the SADM right, the kinds of attributes of automated decision-making that they wish to understand, and what challenges they anticipate in exercising the SADM right.

3 Method

To answer our research questions, we conducted a large-scale online survey with respondents from the UK and US to understand common understandings and perceptions of automated decision-making. We chose surveys over interviews because the former allowed us to study a much larger sample and identify common patterns in their perceptions. Our IRB approved this research.

We decided to focus on the UK and US in this study for many reasons. First, while the UK needs to be compliant with UK-GDPR [18], the US currently does not. However, the underlying transparency, accountability, and fairness principles (i.e., providing transparency about companies' data practices) are much more broadly supported. For example, the Fair Information Practices principles that undergird privacy-related legislation in the US include an openness principle (organizations should be open about "developments, practices and

policies with respect to personal data") and an individual participation principle (individuals can access data about them or confirm whether an organization has data about them) [20]. SADM is a concrete type of data practice. We deliberately chose one country where the SADM right is directly supported while another that has not to see whether they would have drastically different perspectives on the idea of SADM. Second, the US is the world's leading Internet economy and the home for most Internet giants that provide services to global users. The UK is the largest internet economy in the G20 [17] surpassing other European countries such as Germany and France. Third, the two countries are from two continents and might have different cultures. Lastly, English is the official language of both countries allowing us to use the same survey.

3.1 A Video Tutorial of The SADM Right

Since our respondents might not know the SADM right, we created a short video with animations to educate them about the right[1]. We chose this format over others because a video with animations, an audio track, and text captions can be more engaging and accessible than text or audio alone. Below is a summary of how we designed this video tutorial.

First, we searched descriptions and examples of the SADM right from reliable resources such as the GDPR website [70], UK-ICO (UK Information Commissioner's Office) [53], CNIL (French Data Protection Authority) [50], and EU Working Party-29 (AWP29) Guidelines [26]. We specifically sought: (1) introduction of the right as part of the GDPR; (2) explanation of salient features of the right (e.g., under the SADM right, the users were introduced to the terms profiling and automated decision-making, along with an example of each term); and (3) a real-life example illustrating how the right can be exercised. We included a bank-related example from the EU's official website [15]. The example reads, *"You use an online bank for a loan. You are asked to insert your data and the bank's algorithm tells you whether the bank will grant you the loan or not and gives the suggested interest rate. You must be informed that you may express your opinion, contest the decision and demand that the decision made via the algorithm be reviewed by a person. Additionally, the company must also explain to you how the automated decision-making algorithm makes a decision about you and its envisaged legal or significant consequences for you."*

Second, the descriptions and examples were then integrated into a script (written in English) by one researcher, based on which another researcher created various graphical animations using Adobe Illustrator, Adobe Photoshop and iMovie. A third researcher, a legal scholar with expertise in the GDPR, reviewed the script and the animations to ensure their correctness and neutrality. When inaccuracies were identified, we updated the script and the animations accordingly. We performed this process iteratively until all researchers agreed on both the script and animations.

Third, we audio-recorded the final script in English, embedded the audio in the animations, and generated the video. The video was then incorporated into the survey. Before conducting the actual study in Fall 2018, we did a pilot on Amazon Mechanical Turk (US participants) and Prolific (UK participants) with the initial survey design. We received some feedback that the voice of the video was too monotonic. To address this issue, we on-boarded a broadcasting professional to re-record the audio track with more tones and variations to make the audio more engaging.

3.2 Survey Flow

Our survey had a total of 21 questions, including both open-ended and multiple-choice questions. We started by asking people's understanding of the SADM right. The first question asked whether respondents had heard of this right before, and the second asked about their understanding of the right. We then asked another two questions about their understanding of 'user profiling' and 'automated decision-making,' respectively. Next, respondents were asked to watch our video tutorial on the SADM right. Then, we asked a simple attention checking question ("In the video, did you spot a computer screen?") to ensure they paid attention to the video. We then asked them two multiple-choice questions (MCQ) to examine their understandings of the right, e.g., "which of the following statements best explains the Right against solely automated decision making?", "which of the following is an example of the Right against solely automated decision making?" For each MCQ question in the survey, they can select all/multiple responses that apply. However, out of the five answer options there was only one correct answer option to each of the two multiple-choice question, based on the legal definition of the right. For instance, the statement that best explained the SADM right was, *"It allows users to request companies to explain how the AI makes automated decisions about them."* Similarly, the correct example of SADM right was, *"An individual stumbled upon the social media settings page that shows an automatically generated profile that the company uses to provide a personalized news feed and other targeted ads. This person requests the company to explain the process of automatic profiling."* We then asked an open-ended question about decisions that may affect them significantly.

Next, to understand respondents' expectations about attributes to explain for the SADM right, we asked an open-ended question, "Regarding the example from the video about automatically making decisions on bank loans, what aspects of automated decision making would you like to be explained?" This was followed by a multiple-choice question (MCQ) asking respondents to choose attributes they would like to be included in the explanations about SADM. They can choose multiple responses from a randomized list of six

[1] https://youtu.be/BrQMqmPEWQs

pre-defined attributes, adapted from concrete examples in the 'Guidelines on Automated Decision-Making' created by the EU Working Party (A29WP [26]). These attributes include:

- *Type of information:* "Company must inform me what types of data (e.g., my name, age, address) are used in making automatic decisions about me."
- *Explaining logic involved:* "Company must inform me how data is used or what algorithm or method is used in making automatic decisions about me."
- *Fairness of algorithm:* "Company must inform me how it ensures that the algorithm or method used in making automatic decisions about me is fair and unbiased."
- *Source of information:* "Company must inform me where or how the company found/obtained my data that is used in making automatic decisions about me."
- *Company infrastructure:* "Company must provide me customer care services to contest automated decision making."
- *Data accuracy and updates:* "Company must update both my data and the algorithm for accurate decision making as well as inform me about these updates."

We then asked respondents to recall whether they have encountered an incident where they could have exercised this right. If they answered "Yes," we then asked them to provide details about that incident, how they could have exercised the right, what benefits and challenges they could have if they try to exercise the right, and any possible solution to address the challenges. If they answered "No," we directly asked them about the benefits, challenges, and solutions. Lastly, we asked their overall understanding of the right using a Likert scale question followed by an open-ended question.

3.3 Data Analysis

For the open-ended questions, we conducted a thematic analysis [8], a standard method for analyzing qualitative data. We used a software called Dedoose to code the open-ended survey responses qualitatively. Three researchers coded together and discussed a 10% subset of the data. Once the coders achieved a good understanding of the data, the three coders continued to code another 10% subset of the data independently, then discussed and reconciled their coding to develop the initial codebook. The inter-coder reliability is 0.89 (Cohen's Kappa), which is considered good [27]. Due to a large amount of qualitative data from our survey, we decided to involve another two research assistants to help with the coding process. The inter-coder reliability for the additional two coders was 0.88 (Cohen's Kappa). Using the agreed-upon code-book, one researcher and the two new coders repeated the procedure mentioned above to ensure that all coders shared a good understanding of the coding process. Then each of the five coders independently coded a subset of the rest of the data. We added new codes to the code-book when existing codes cannot capture the data. Upon completion, the final code-book contained over 300 lower-level codes. We then grouped all codes into higher-level themes, such as people's understanding or misunderstandings of the SADM right before and after watching the video (i.e., *understanding* questions), attributes of SADM that people desire to know (i.e., *attribute* questions), and different types of expected challenges in exercising the right (i.e., *challenge* questions).

Furthermore, we compared how UK and US respondents answered these key questions on understanding, attribute, and challenges regarding the percentage of respondents who selected each answer in the multiple-choice questions or mentioned a theme in the open-ended questions. Since each respondent can select/mention multiple answers in each question, we cannot perform Chi-Square tests comparing the two groups because the data is not independent (e.g., two answers were chosen by the same respondent). Instead, we treated each answer/theme as a separate, binary question. A respondent can only select/mention an answer or not. We then conducted tests of proportions between the UK and US respondents for each answer. Since these are essentially post-hoc comparisons, we applied Bonferroni correction to adjust the family-wise p value. Specifically, we divided the common 0.05 p value threshold by the number of answers/themes in each question. The adjusted p value threshold for statistical significance ranges from 0.007 to 0.01 depending on the specific questions

3.4 Participants

Similar to prior research, we used Amazon Mechanical Turk (MTurk, based in the US) to recruit US respondents (e.g., [39]) and Prolific (based in the UK) to recruit UK respondents (e.g., [54]). We could not recruit enough UK respondents from MTurk because MTurk is much less popular than Prolific in the UK. Similarly, we recruited US respondents on MTurk because it is much more popular than Prolific in the US. Therefore, we had to use two platforms, each for respondents from one country. We required the respondents to be residing in the US/UK and had more than 95% task acceptance on both platforms.

We manually removed incomplete or randomly filled responses based on their quality. For instance, some respondents pasted the same text (e.g., a long paragraph about a product, which is completely irrelevant to our topic) to each answer. Others entered unintelligible characters or words. We removed 15 responses from MTurk (used for US respondents) and two responses from Prolific (used for UK respondents) due to quality issues. After filtering, we had 392 valid responses in total, including 192 from the US and 200 from the UK. For our US sample, 61% were male and 39% were female. Respondents' ages ranged from 25 to 35 (SD = 1.15).

88% of the respondents had at least some college education, and 21% had a technology background. For our UK sample, 62% were female and 38% were male. Their ages ranged from 25 to 34 (SD = 1.15). 70% had a college education and 7% had a technology background. The average time for completing the survey was about 25 minutes. After quality check, each US participant was paid $3 and each UK participant was paid $3.5 (to meet the payment requirement of Prolific).

4 Results

This paper focuses on (1) peoples' understanding and misunderstandings of the SADM right; (2) attributes of SADM that our respondents desired to understand; (3) challenges anticipated by the respondents in exercising this right. We found that the UK and US responses were broadly consistent and did not observe any notable differences in these results between the two groups. As detailed in Section 3.3, we conducted tests of proportions to compare the UK vs. US responses and found no statistically significant difference.

4.1 Peoples' Understanding of The Right

4.1.1 Before watching the video tutorial

We asked our respondents about their understanding of the SADM right, before watching the video tutorial of the right. As summarized in Table 1, we present peoples' misunderstandings and reasonable understandings of the right based on the GDPR definition of the SADM right, both before and after watching the video tutorial.

Misunderstandings. Many respondents incorrectly assumed that the SADM right inherently allows people to deny being subjected to SADM or let the companies use their personal information, including profiling for any SADM decision-making. For instance, P148 from UK thought that *"it's the right for you to not consent to automatised decisions"*. Some respondents even anticipated that the right could allow them to completely opt-out of automated decisions made by computers or algorithms. However, this is a misconception because companies can legally do user profiling and SADM if consumers agree to the service contract. In practice, this may allow any Internet-based services (with a valid service agreement and a privacy notice) to subject people to SADM processing. Furthermore, respondents thought that they could choose *ex-ante* (i.e., before a decision is made) between automated and human-involved decision making. For instance, P151 from UK believed that, *"you have the right to request that a human looks at your application for something before a decision is made."* However, the GDPR (under Art 22(3) [70]) allows for human intervention *ex-post* (i.e., only after) the user is subjected to SADM decisions. Majority of respondents reported a lack of understanding of the SADM right, primarily, because they had not heard this right before.

Before tutorial	US	UK
Deny subjecting to SADM	58	57
Human involvement	49	33
Don't know/unsure	45	70
User control and choice	25	25
Prohibit profiling	11	3
Inform about SADM	10	4
Obtain explanation	6	12

After tutorial	US	UK
Deny subjecting to SADM	66	60
Inform about SADM	43	40
Obtain explanation	34	44
Human involvement	27	29
Prohibit profiling	16	11
Don't know/unsure	16	21
Contest against SADM	6	7

Table 1: The top table shows our respondents' answers to the open-ended question, understanding of the SADM right before watching the tutorial. The bottom table shows our respondents' answers to the open-ended question, understanding of the SADM right after watching the tutorial.

Reasonable understandings. Some respondents correctly assumed that they could request for human review of automated decisions ex-post, i.e., after the automated decision is made. Art 22(3) under GDPR provides the right to the consumers to obtain human intervention, express their concern, or even contest the solely-automated decision. We also found people mostly preferred human review in the case of negative outcome (e.g., loan rejection) of a SADM decision. For instance, according to P108 (UK), *"if I have applied for something and have been rejected automatically by [a] computer I can appeal and have it looked at by human."*

Some respondents also preferred to be informed when subjected to SADM and obtain an explanation for SADM decisions made. Art 13(2)(f) GDPR highlights these requirements, stating that companies must inform users about the existence of SADM, including profiling, and provide meaningful information about the logic involved to make the decision.

"I don't know." Many respondents reported having no prior knowledge of the right. While some respondents guessed a basic meaning of the right. Others were unable to even guess. For instance, P36 from US reported that, *'I have not got even the smallest of clues'* about what SADM right means.

4.1.2 After watching the video tutorial.

After watching the tutorial, over 97% of US and 90% of UK respondents correctly answered the attention-checking question, suggesting that the vast majority of them watched the video carefully. They also felt the video was informative in

helping them understand SADM. For example, P85 from US found that *"Some of the questions, such as the "automated decision making" were a little confusing but [..] the video made it a lot easier to understand."* Our results also suggest that most respondents gained a basic understanding of the right after watching the video. Based on the responses from the two multiple-choice questions examining their understanding of the right after watching the video, 80% of both US and UK respondents' expressed understanding was consistent with the legal definition. The enhanced understanding of the SADM right also helped respondents to answer subsequent questions such as those related to attributes of SADM process to explain and challenges to exercise SADM right.

At the end of the survey, over 80% of US respondents and 64% of UK respondents reported either extremely clear or somewhat clear understanding of the SADM right, based on the responses to the Likert scale question. Our respondents also provided their self-reported understanding of the right through an open-ended question towards the end of the survey. We found that their understanding was consistent with the legal definition and covered significant aspects of the right. These aspects included participants' assumptions that the right allows people to be informed about being subjected to SADM, to obtain explanations of the SADM process, to request human involvement, and to contest against solely automated decisions. However, some respondents still had misunderstandings after watching the video. For instance, some of them still believed that the right prohibits user profiling for targeted ads. A large proportion of respondents were still keen to deny being subjected to SADM if it affects them significantly. For instance, P57 from UK reported that his opinion about this right remained same as before watching the video, i.e., *"This is a person's right to get fair treatment free from pre-programmed decision making algorithms."* They prefer not to be subjected to SADM process because they lack trust in the SADM systems to make fair decisions, especially in case of high-stake decisions impacting them legally or significantly.

4.2 SADM Attributes People Want to Know

To capture people's expectations about the attributes of the 'solely' automated decision-making that they desire to understand, we asked an open-ended question, followed by a multiple-choice question. Table 2 summarizes the answers to the two questions, respectively. These answers are attributes frequently mentioned by our respondents.

4.2.1 Attributes from survey open-ended responses

First, we present the results from the open-ended question. We further classify them as part of the local or global explanation (as defined in section 2.3), wherever applicable. Our respondents' answers (summarized in top of Table 2) covered major themes such as the type of information used, the process of

Open-ended	US	UK
Type of information	77	88
Personalized explanation	56	52
Unique factors	44	34
Source of information	35	56
No explanation	12	20
Human involvement & appeal	19	24
Fairness of algorithm	6	10

Multiple-choice	US	UK
Type of information	40	36
Explaining logic involved	34	36
Fairness of algorithm	34	35
Source of information	34	31
Company infrastructure	32	29
Data accuracy and updates & appeal	28	27

Table 2: The top table shows our respondents' answers to the open-ended question regarding what attributes they would like to understand and the corresponding number of US and UK respondents who mentioned each attribute. The bottom table shows our respondents' answers to the multiple-choice question where they can select multiple answers where each answer option (i.e., attribute to explain) was suggested by the policy-makers (in this case, the Working Party A29WP [26]). In the top table, the attributes in bold were only expressed by the respondents, i.e., not mentioned by A29WP and thus not shown in the right table.

SADM, and the source of information. A large percentage of respondents (US 40%, UK 44%) reported that they were interested in knowing about the 'type of information' used. They also desired to know the 'source of information' (US 18%, UK 28%). For instance, P136 from UK wanted to *"[..] know what information is used to make the decision and where it was obtained from."*

Furthermore, many respondents desired to know the 'unique factors considered' (US 23%, UK 17%), i.e., the criteria used in the decision-making and the *weights* of those factors that were unique to their profiles. For instance, P46 from the US emphasized that *"I would like for them to explain how much weight they put on what specific data information that have about me"* and *"how they determine what data is important and what data is not."* Similarly, P14 from UK requested to know *"what factors are taken into consideration when applying for a loan ...[and what] could hinder your chances of getting a loan."* 'Factors considered' differ from the 'type of information' used because the factors could include a subset of the user information collected and possibly other non-user information (e.g., market interest rate) as criteria for decision-making. For instance, a company can collect

different types of user information such as age and gender, but age is the most important factor considered for SADM (e.g., age greater than a certain number). These attributes can be classified as part of local explanation since they help to explain how the model makes a specific decision.

The need to know the factors considered relates to the expectations to receive *personalized* explanations. People reported that they wanted to have personalized explanations of automated decisions (US 29%, UK 26%) as a way to seek justification for the decision made for them. This explanation should include reasons for a negative decision, their impact, limitations, and ways to improve/alter the decision. For instance, P136 from UK wanted to know *"How my circumstances were assessed - what factors led to me being accepted or turned down, if it was a points based system, how close was I/ what would I need to do to achieve the necessary score."* She preferred a personalized explanation of a particular decision rather than a general explanation of the SADM process. She also desired actionable information about what she can do to improve the decision outcome. Similarly, P170 from US stated that *"It would be the same questions I would ask of an employee"*. Here P170 expected to receive a machine-provided explanation similar to the experience of receiving an explanation from a human decision-maker. One possibility is to create chatbots that can converse with the user to explain the decision made automatically. In another instance, both P51 from US and P195 from UK pointed out the risks of user profiling in SADM and therefore expected greater transparency in explanations. P51 suspected that the automated decisions could use *"the data that was biased in any way in an unfair manner"* and therefore demanded to know *"complete description of how each aspect causes a change in their decisions [..] as well as the changes that could be made"*. Similarly, P195 questioned *"how certain features such as age and gender can"* influence a decision to make *"people be potentially deprived a loan"*.

It is worth noting that some respondents expressed that they did not want any explanation about automated decision-making (US: 12, UK: 20). They did not report specific reasons for this preference when answering this question. However, later when discussing their anticipated challenges in exercising the right, some respondents said that knowing SADM can be boring, or not worth their time and effort. For instance, P145 from the UK said, *"effort required to exercise this right is greater than the 'reward' I would expect to gain."* Here, he seemed to derive his preference based on a cost-benefit analysis of exercising the right. As we will discuss later, one common challenge people anticipated was that exercising the right can be too time-consuming. Another possible reason could be that people feel resigned about SADM. Privacy scholars [67] have suggested that most Americans give up data for relevant ads, not because of convenience, but resignation. Rather than participating in a rational exchange, consumers are giving up their personal information with 'a feeling of futility' [67]. Future research can further investigate why some people do not want an explanation of SADM.

4.2.2 Attributes in policy-makers' recommendations

Next, we compare themes from the open-ended question with the six pre-defined attributes from the MCQ-based question. As described in section 3.2, these pre-defined attributes were adapted from the GDPR Working Party-29 Guidelines [26] and included: a) type of information, b) explaining logic involved, c) fairness of algorithm, d) source of information, e) company infrastructure, and f) data accuracy and updates.

Table 2 (right) shows US and UK respondents' choices of these pre-defined attributes of SADM where each respondent can choose multiple attributes. A majority of respondents from both countries wanted to know the 'type of information' used, followed by 'explaining logic involved', 'fairness of algorithm', and 'source of information.' 'Data accuracy and updates' and 'company infrastructure' were least selected.

It is interesting to note that three major themes from the open-ended question corroborate with similar pre-defined attributes from the multiple-choice question in terms of how commonly they were expected by our respondents. These included 'type of information', 'explaining logic involved' and the 'source of information' used.

However, we also observed some differences between the responses to the two questions. For instance, while about one-sixth of respondents selected 'Fairness of algorithm' in the multiple-choice question, a much smaller percentage of respondents expressed it in the open-ended question. Since we asked the open-ended question first, this might suggest respondents did not immediately think about the fairness aspect. Or in other words, they might be paying more attention to a negative decision about themselves and a need for a personalized explanation of the decision rather than whether the SADM process is fair.

Last but not least, we found that some themes were unique to the open-ended question, such as 'seeking justification' for the decision, requesting a personalized explanation including the weight of the 'factors considered', or not willing to be presented with any explanation at all. This is an important list of attributes because they were sought by the respondents but were not covered in policy makers' recommendations. This gap might lead to ineffective SADM explanations that do not satisfy people's expectations or needs.

4.3 Anticipated Challenges

We heard from our respondents about what attributes they prefer in the explanations for the SADM decision. However, will they be able to exercise this right easily if they want to? We next asked respondents about their perceived challenges of using this right. Participants reported several types of challenges that may arise at different stages of exercising the

Major challenges anticipated	US	UK
Hard to safeguard against SADM processing	40	54
Common Challenges	63	75
Hard to identify and fight bias in algorithm	62	69
Hard to contest SADM decisions	32	29
Review process not user friendly	26	18

Table 3: Main perceived challenges of exercising SADM mentioned by our respondents.

SADM right. Some of these challenges (e.g., hard to rectify incorrect information about individuals) have been reported in other contexts. In contrast, other challenges seem more salient in the context of the SADM right. We will briefly summarize the former and then focus on the latter. Table 3 summarizes the number of US and UK respondents who mentioned each type of these challenges.

4.3.1 Common Challenges

Our respondents reported several perceived challenges of the SADM right. These challenges are also common in other contexts for protecting people's privacy, such as: lack of user awareness as well as control of personal data collection and sharing, lack of user trust on companies, lack of transparency of company privacy practices, time-consuming to communicate with companies, and difficult to rectify incorrect information that companies have about individuals.

4.3.2 Challenges More Salient in SADM

From the participants' responses, we identified several challenges that are either more salient or have different implications in the context of SADM right. We present them below.

Hard to identify and fight biases in algorithms. Respondents were concerned that the algorithms used in the SADM process could be wrought with biases. Such a bias could be hard to detect, verify, prove to cause discrimination. P184 from the US pointed out that as an initial step, it will be difficult to check *"how the algorithm is updated and how fair and unbiased it is."* He further suspected that the companies could be reluctant to provide such information, making it all the more difficult to assess. Furthermore, if the algorithm is found to be biased, it may be challenging to provide proof of bias to the company and request them to make it unbiased for future decision-making.

Hard to safeguard against SADM processing. Many participants reported that they might not be aware of being subjected to SADM processing while using an internet-based service, e.g., social media, banking. For instance, P 136 from the US expressed that *"It might be difficult to know in the first place whether I was affected by automated decision making or not."* It could be because companies might not actively notify people about SADM processing or, as P21 from the US noted, *"It's hidden in legal terms that most people don't understand so you don't ever even notice it."* Additionally, respondents felt that once the SADM system decides for them, it would be challenging to request a re-evaluation of the outcome. Users might find it challenging to explain their circumstances for companies to re-evaluate the decisions about them. For instance, P36 from the UK explained that while applying for a job or a loan, one of the biggest challenges is that SADM *"decisions [..] don't take into consideration personal circumstances or personality"* of an individual and that the person is *"not being able to explain [themselves].."* in case of negative outcome. As a result, people expected to have an ex-ante opt-out option to not be subjected to the SADM processing. However, they also worried that even if the option existed, they may not know it either due to their own blind-spot or the lack of transparency from companies.

Hard to contest SADM decisions. Respondents from both the UK and the US pointed out that it will be difficult to contest the SADM decisions. Companies can show resistance or be *"hesitant about having human review the decision of the automated system"* as P105 (US) put it. He anticipated that one way of doing so is to make it difficult to contact the company for human involvement in SADM. Based on his real life experiences, he cited that *"unfortunately, most customer service is anonymously automated. Seldom is there an actual human being in which I could contact to deal with this issue."* While reaching customer services might be a need and challenge in commerce, getting a real person to check the machine-made decisions is at the core of the SADM right.

Review process not user-friendly. Even if the users get access to the customer representative, such customer staff may lack relevant training to review the SADM decisions. For instance, P81 from US questioned the quality of human involvement to review a decision. He expressed, *" I don't anticipate any meaningful human interventions that would contradict the results of an automated system."* It could also be possible that companies may not invest in a dedicated team of experts to review automated decision, and *"presumably, the "human intervention" could just be an intern who clicks okay to a computer prompt"*. Additionally, the companies may even lack human resources to respond to large amounts of complaints, leading to slow redress of decision.

People were also concerned that even with required human involvement, the review process *"won't be impartial"* (P80, UK). P88 from the US suspected companies would be *"unwilling to change decision"* because *"[it] likely makes them too much money for them to actually be willing to do anything to change it."* This highlights people's lack of trust in companies to carefully review SADM decisions and change them if needed.

5 Discussion

This study explored research questions regarding people's understanding of the SADM right, aspects of SADM that they want to be explained, and their perceived challenges in exercising the SADM right. Our respondents incorrectly assumed that the right allows them to deny being subjected to SADM or to opt-out altogether. Respondents expected a few novel attributes of SADM to be explained, which are not covered by government guidelines on SADM [26]. These novel attributes ranged from receiving personalized explanations to seeking justification for adverse outcomes. Lastly, our respondents reported four broad perceived challenges of exercising the SADM right: i) *Hard to safeguard against SADM processing*: respondents anticipated that they might not be aware of whether or when companies subject them to SADM processing; ii) *Hard to identify and fight algorithmic biases*: respondents were concerned that algorithms might harbor biases that would be hard to detect, verify, and prove to cause discrimination; iii) *Hard to contest the outcome of SADM*: respondents anticipated that companies could show resistance to review adverse outcomes/decisions by complicating ways to contact them or denying human reviewers to re-evaluate the SADM decision; and iv) *Unfriendly review process*: respondents anticipated that companies may lack the infrastructure to respond to SADM decision review requests or that human reviewers may lack the relevant training to conduct the review. Table 4 summarizes the main findings. Next, we will discuss design and policy implications that can mitigate some of the misunderstandings and challenges and support people's informational needs of SADM.

People's perspective on SADM right	Main findings
Understandings of the SADM right	**Misunderstandings:** - Deny subjecting to SADM - User Choice and Control (e.g. opt out) - Don't know **Reasonable understandings:** - Obtain explanation - To be informed about decision - Prohibit profiling (e.g. for targeted ads) - Request for human intervention - Contest against SADM right
Attributes of SADM that people desired to be explained	**Attributes (open-ended):** - Type of information - Personalized explanation - Unique factors - Source of information - Human involvement and appeal - Fairness of algorithm
Challenges anticipated in exercising the SADM right	- Hard to safeguard against SADM processing - Hard to identify and fight bias in algorithm - Hard to contest SADM decisions - Review process not user friendly - Other common challenges

Table 4: A summary of the main findings from the study.

5.1 Design Implications

5.1.1 Help People Understand SADM

Personalized explanation templates. Our findings suggest that people have different expectations or informational needs for explanations of SADM. Many respondents anticipated personalized explanations of the SADM outcomes. They expect companies to provide explanations that can justify the decision regardless of the outcome and provide actionable suggestions on what people may do to improve the decision outcome, especially in negative decisions. While designing explanations, caution must be taken to balance over-simplified vs. over-complicated explanations [25, 40, 49]. To strike a good balance, companies can present hierarchical explanations with multiple levels of details and personalization, which is similar to the privacy nutrition labels for IoT devices [23]. For instance, it would be useful to design explanation templates that would allow people to zoom into details about a specific decision, personalized based on individuals' demographics (e.g., age) and computing knowledge. It could also include suggestions to improve the outcome next time. Suppose a bank rejects a person for a credit card; it could provide an explanation template with multiple levels of personalized explanations. At the basic level, it could include the outcome and a high-level reason (e.g., low credit score). At the intermediate level, the user could zoom into details such as attributes considered for the outcome and which attributes the user could improve for a better outcome. A more detailed level could include additional information such as what other personal data the company uses to make SADM decisions and its source, how the company ensures the fairness and accuracy of the algorithm used, and other aspects (see Table 2).

Sandboxes to play with SADM systems. Another key finding of our study is that our respondents expected explanations to include how a decision is relevant to them. It includes 1) describing the (significant) effect that the SADM decision has on the individual and 2) providing individually tailored and practically actionable recommendations to improve future outcomes for the individual. A concrete design idea is that companies can implement interactive interfaces that describe how different factors affect a decision, allow users to interact with various factors to see their impact and provide personalized recommendations to improve acceptance for future outcomes. For instance, a car insurance company can provide customers with an interactive sandbox to explain the extra insurance premium for certain events (e.g., speeding). It can also allow customers to play with the algorithm (or sandbox) by testing different factors such as levels and times of speeding, driver age, and past accident records to see the dependence of outcome on these factors. However, as a challenge, the sandbox could possibly allow for reverse-

engineering the model, which would not be ideal in anomaly detection cases (e.g., fraud prevention).

5.1.2 Mitigate Misunderstandings and Challenges

Icons for SADM processing. Since our respondents had many misconceptions about the SADM right, it would be beneficial for companies to consider these misconceptions when designing their platforms to better support users exercise the SADM right. Additionally, our respondents anticipated that it would be challenging to detect whether/when they are subject to SADM. One way to communicate whether a user is subject to SADM is to show indicators of SADM (visual icons or other modalities). Various privacy icons have been proposed to convey complicated privacy concepts. Some of the icons are about the privacy notices in various domains (e.g., online tracking [19,47], social media [36], web cams [22,58], web links [37]), while others are to convey privacy choices [33,65].

It is worth noting that even though there are icons representing targeted ads, targeted ads are only one example of a much broader set of SADM practices. We are not aware of any existing icons for SADM processing. For instance, imagine that a bank website shows a SADM icon next to its credit card application and loan application. An e-commerce site shows the icon next to its recommended products to a user, or a social media site displays the icon next to the recommended friends to a user. These icons can represent whether the corresponding decisions about the prevailing user are made by an automated system using algorithms alone (SADM) or by involving humans in the process (human-AI hybrid).

In addition, these icons could increase transparency by highlighting the uncertainty of the decision made by using solely automated systems. This was another aspect that some of our respondents reported as a challenge of the right. The icons would be even more useful if they are standardized (e.g., by self-regulating trade organizations or the Internet standard organization, W3C). While these icons are likely to have their challenges (e.g., people may ignore or misunderstand them), when they are designed and evaluated appropriately (e.g., see good examples of nutrition labels and the recent CCPA opt-out icons), they can help communicate SADM.

Social support for contesting algorithmic decisions. Our respondents felt it would be challenging to contest the outcome of SADM due to unfriendly review processes. They also anticipated the difficulty for them to explain their personal circumstances to appeal a negative decision. One design direction is to create tools or platforms that allow people to share and learn from each other about their strategies and experiences in working with specific companies for their redress/contest requests. This is similar to crowd-sourcing help for tech support, such as [46]. End-user tools could be designed to guide users to create and share contest requests (e.g., answering a set of questions and attaching supporting documents). These shared user experiences can also further motivate companies to improve how they handle people's redress requests. How SADM systems can be designed to support contestability is another exciting future direction.

5.2 Policy Implications

Our results also have some policy implications. For instance, policymakers could consider adding novel attributes to policy guidelines for companies to explain decisions made by SADM. As suggested by participants, these attributes include 'seeking justification' for the decision, requesting 'personalized explanation' including the weight of the 'factors considered,' or an option to opt-out of receiving any explanation at all. Policymakers could also consider mandating notifications to alert users when subjected to SADM processing, similar to website cookie notifications. Lastly, policymakers and industry organizations could initiate regulatory standards that could be applied internationally (e.g., a standard explanation template) similar to ISO standards.

6 Limitations

Our study has several limitations. First, we recruited adult participants, but our participants ended up with a small age range (about 25-35 years old), which was not intentional. Second, our study data does not explain the lack of differences between the US and UK responses. Third, we analyzed but did not observe any correlations between the "no explanation" responses and other participant data (e.g., demographics). Fourth, our results may not be generalizable since we cannot claim that our sample represents the populations in those two countries. Lastly, our study focused on people's desire for explanation rather than how SADM explanations are currently implemented. Future work could attempt to address these limitations and explore these directions further.

7 Conclusion

Automated decision-making, including profiling based solely on computer algorithms, continues to grow in importance as more companies adopt this practice to improve efficiency. The challenge, therefore, is to find the balance between opportunities for the companies and the impact on end-users. We studied people's understandings and expectations of the right against solely automated decisions because they are major stakeholders and would be directly impacted by these automated decisions. We presented design implications for companies to support citizens in exercising this right. Future research can explore concrete designs and modalities to explain various attributes of automated decision-making that people desire to understand.

8 Acknowledgement

We thank our participants for sharing their insights. We also thank the anonymous reviewers for their thoughtful feedback.

References

[1] 2018. The EU General Data Protection Regulation | Human Rights Watch. (June 2018). (Accessed on 09/12/2018) https://www.hrw.org/news/2018/06/06/eu-general-data-protection-regulation.

[2] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society* 20, 3 (2018), 973–989.

[3] Yannis-NYU Bakos, David R Trossen, and others. 2009. Does Anyone Read the Fine Print? Testing a Law and Economics Approach toStandard Form Contracts. (2009).

[4] Solon Barocas. 2014. Data mining and the discourse on discrimination. In *Data Ethics Workshop, Conference on Knowledge Discovery and Data Mining*. 1–4.

[5] 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. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 377.

[6] Rainer Böhme and Stefan Köpsell. 2010. Trained to accept?: a field experiment on consent dialogs. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2403–2406.

[7] Mark Bovens. 2007. Analysing and assessing accountability: A conceptual framework 1. *European law journal* 13, 4 (2007), 447–468.

[8] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development*. sage.

[9] Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. 2015. The role of explanations on trust and reliance in clinical decision support systems. In *Healthcare Informatics (ICHI), 2015 International Conference on*. IEEE, 160–169.

[10] Corinne Cath, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. 2018. Artificial intelligence and the 'good society': the US, EU, and UK approach. *Science and engineering ethics* 24, 2 (2018), 505–528.

[11] Electronic Privacy Information Center. 2016. EPIC - AI and Human Rights. (2016). https://www.epic.org/ai/index.html (Accessed on 09/17/2018).

[12] Robert P Charrow and Veda R Charrow. 1979. Making legal language understandable: A psycholinguistic study of jury instructions. *Columbia law review* 79, 7 (1979), 1306–1374.

[13] Danielle Keats Citron and Frank Pasquale. 2014. The scored society: Due process for automated predictions. *Washington Law Review* 89 (2014), 1.

[14] Cary Coglianese and David Lehr. 2016. Regulating by robot: Administrative decision making in the machine-learning era. *Georgetown Law Journal* 105 (2016), 1147.

[15] European Commission. 2018. Can I be subject to automated individual decision-making, including profiling? https://ec.europa.eu/info/law/law-topic/data-protection/reform/rights-citizens/my-rights. (August 2018).

[16] Thomas H Davenport and Jeanne G Harris. 2005. Automated decision making comes of age. *MIT Sloan Management Review* 46, 4 (2005), 83.

[17] Culture Media Sports Department for Digital. 2017. A New Data Protection Bill: Our Planned Reforms. (August 2017). (Accessed on 09/16/2018) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/635900/2017-08-07_DP_Bill_-_Statement_of_Intent.pdf.

[18] Culture Media Sports Department for Digital. 2019. The Data Protection, Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019. (February 2019). (Accessed on 02/23/2021) https://www.legislation.gov.uk/uksi/2019/419/made.

[19] Disconnect, Inc. 2014. Disconnect Privacy Icons. (2014). https://github.com/disconnectme/privacy-icons.

[20] Pam Dixon. 2008. A Brief Introduction to Fair Information Practices. (2008). https://tinyurl.com/7mmym54v (Accessed on 09/17/2018).

[21] Finale Doshi-Velez, Mason Kortz, Ryan Budish, Chris Bavitz, Sam Gershman, David O'Brien, Stuart Schieber, James Waldo, David Weinberger, and Alexandra Wood. 2017. Accountability of AI under the law: The role of explanation. *arXiv preprint arXiv:1711.01134* (2017).

[22] Serge Egelman, Raghudeep Kannavara, and Richard Chow. 2015. Is This Thing On? Crowdsourcing Privacy Indicators for Ubiquitous Sensing Platforms. In *Conference on Human Factors in Computing Systems (CHI)*. ACM, 1669–1678. DOI:http://dx.doi.org/10.1145/2702123.2702251

[23] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the Experts: What Should Be on an IoT Privacy and Security Label?. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 447–464.

[24] Motahhare Eslami, Sneha R Krishna Kumaran, Christian Sandvig, and Karrie Karahalios. 2018. Communicating Algorithmic Process in Online Behavioral Advertising. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 432.

[25] Motahhare Eslami, Kristen Vaccaro, Karrie Karahalios, and Kevin Hamilton. 2017. " Be Careful; Things Can Be Worse than They Appear": Understanding Biased Algorithms and Users' Behavior Around Them in Rating Platforms.. In *ICWSM*. 62–71.

[26] Working Party EU. 2018. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. (February 2018). (Accessed on 09/17/2018) https://iapp.org/media/pdf/resource_center/W29-auto-decision_profiling_02-2018.pdf.

[27] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. 2013. *Statistical methods for rates and proportions*. John Wiley & Sons.

[28] Tarleton Gillespie, Pablo J Boczkowski, and Kirsten A Foot. 2014. *Media technologies: Essays on communication, materiality, and society*. MIT Press.

[29] Bryce Goodman and Seth Flaxman. 2017. European Union regulations on algorithmic decision-making and a" right to explanation". *AI magazine* 38, 3 (2017), 50–57.

[30] Robert Gorwa. 2019. What is platform governance? *Information, Communication & Society* 22, 6 (2019), 854–871.

[31] Nina Grgic-Hlaca, Elissa M. Redmiles, Krishna P. Gummadi, and Adrian Weller. 2018. Human Perceptions of Fairness in Algorithmic Decision Making: A Case Study of Criminal Risk Prediction. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. International World Wide Web Conferences Steering Committee, Lyon, France, 903–912. DOI:http://dx.doi.org/10.1145/3178876.3186138

[32] Jens Grossklags and Nathan Good. 2007. Empirical studies on software notices to inform policy makers and usability designers. In *International Conference on Financial Cryptography and Data Security*. Springer, 341–355.

[33] Hana Habib, Yixin Zou, Yaxing Yao, Alessandro Acquisti, Lorrie Cranor, Joel Reidenberg, Norman Sadeh, and Florian Schaub. 2021. Toggles, dollar signs, and triangles: How to (in) effectively convey privacy choices with icons and link texts. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–25.

[34] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé, Miro Dudik, and Hanna Wallach. 2019. Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, Glasgow, Scotland Uk, 1–16. DOI:http://dx.doi.org/10.1145/3290605.3300830

[35] Xia Mengnan Du Hu, Ninghao Liu. 2020. Techniques for Interpretable Machine Learning. *Commun. ACM* 63, 1 (2020), 68–77. https://cacm.acm.org/magazines/2020/1/241703-techniques-for-interpretable-\machine-learning/fulltext

[36] Renato Iannella and Adam Finden. 2010. Privacy Awareness: Icons and Expression for Social Networks. In *International Workshop for Technical, Economic and Legal Aspects of Business Models for Virtual Goods*. Article 1, 15 pages. http://virtualgoods.org/2010/VirtualGoodsBook2010_13.pdf.

[37] Saraschandra Karanam, Janhavi Viswanathan, Anand Theertha, Bipin Indurkhya, and Herre Van Oostendorp. 2010. Impact of Placing Icons Next to Hyperlinks on Information-Retrieval Tasks on the Web. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 32. eScholarship, 2834–2839. https://escholarship.org/content/qt27w0n9kc/qt27w0n9kc.pdf.

[38] Farzaneh Karegar. 2018. *Towards Improving Transparency, Intervenability, and Consent in HCI*. Ph.D. Dissertation. Karlstad University Press.

[39] Aniket Kittur, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing User Studies with Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 453–456. DOI:http://dx.doi.org/10.1145/1357054.1357127

[40] René F Kizilcec. 2016. How much information?: Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2390–2395.

[41] Joshua A Kroll, Solon Barocas, Edward W Felten, Joel R Reidenberg, David G Robinson, and Harlan Yu. 2016. Accountable algorithms. *U. Pa. L. Rev.* 165 (2016), 633.

[42] Min Kyung Lee and Su Baykal. 2017. Algorithmic Mediation in Group Decisions: Fairness Perceptions of Algorithmically Mediated vs. Discussion-Based Social Division. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, Portland, Oregon, USA, 1035–1048. DOI: http://dx.doi.org/10.1145/2998181.2998230

[43] Tae Ho Lee and Lois A Boynton. 2017. Conceptualizing transparency: Propositions for the integration of situational factors and stakeholders' perspectives. *Public Relations Inquiry* 6, 3 (2017), 233–251.

[44] Benjamin Letham, Cynthia Rudin, Tyler H McCormick, David Madigan, and others. 2015. Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics* 9, 3 (2015), 1350–1371.

[45] David Meyer. 2019. Who should enforce a US federal privacy law? https://iapp.org/news/a/who-should-enforce-a-federal-privacy-law/. (2019). (Accessed on 02/23/2021).

[46] Microsoft. 2019. Microsoft Community. https://answers.microsoft.com/en-us/page/gettingstarted. (2019). (Accessed on 02/23/2021).

[47] Mozilla. 2020. Privacy Icons. (February 2020). https://wiki.mozilla.org/Privacy_Icons.

[48] Deirdre K Mulligan, Joshua A Kroll, Nitin Kohli, and Richmond Y Wong. 2019. This Thing Called Fairness: Disciplinary Confusion Realizing a Value in Technology. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–36.

[49] Menaka Narayanan, Emily Chen, Jeffrey He, Been Kim, Sam Gershman, and Finale Doshi-Velez. 2018. How do humans understand explanations from machine learning systems? an evaluation of the human-interpretability of explanation. *arXiv preprint arXiv:1802.00682* (2018).

[50] Commission nationale de l'informatique et des libertes. 2013. Homepage | CNIL. (2013). (Accessed on 09/17/2018) https://www.cnil.fr/en/home.

[51] Daniel Neyland. 2016. Bearing account-able witness to the ethical algorithmic system. *Science, Technology, & Human Values* 41, 1 (2016), 50–76.

[52] International Association of Privacy Professionals (IAPP). 2020. The California Privacy Rights Act of 2020. https://iapp.org/resources/article/the-california-privacy-rights-act-of-2020/. (2020). (Accessed on 02/23/2021).

[53] Information Commissioner's Office. 2018. Home | ICO. (2018). (Accessed on 09/17/2018) https://ico.org.uk/.

[54] Stefan Palan and Christian Schitter. 2018. Prolific.ac - A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27.

[55] Andrew S Patrick and Steve Kenny. 2003. From privacy legislation to interface design: Implementing information privacy in human-computer interactions. In *International Workshop on Privacy Enhancing Technologies*. Springer, 107–124.

[56] Sam Pfeifle. 2018. California passes landmark privacy legislation. https://tinyurl.com/b5es9y9x. (2018). (Accessed on 02/23/2021).

[57] Irene Pollach. 2007. What's wrong with online privacy policies? *Commun. ACM* 50, 9 (2007), 103–108.

[58] Rebecca S Portnoff, Linda N Lee, Serge Egelman, Pratyush Mishra, Derek Leung, and David Wagner. 2015. Somebody's Watching me? Assessing the Effectiveness of Webcam Indicator Lights. In *Conference on Human Factors in Computing Systems (CHI)*. ACM, 1649–1658. DOI:http://dx.doi.org/10.1145/2702123.2702164

[59] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. Why should i trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 1135–1144.

[60] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry* (2014), 1–23.

[61] Nripsuta Saxena, Karen Huang, Evan DeFilippis, Goran Radanovic, David Parkes, and Yang Liu. 2019. How Do Fairness Definitions Fare? Examining Public Attitudes Towards Algorithmic Definitions of Fairness. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. http://arxiv.org/abs/1811.03654 arXiv: 1811.03654.

[62] Nick Seaver. 2013. Knowing algorithms. *Media in Transition* 8 (2013), 1–12.

[63] Andrew Selbst and Solon Barocas. 2017. Regulating inscrutable systems. *WeRobot 2017* (2017).

[64] Donghee Shin and Yong Jin Park. 2019. Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior* 98 (2019), 277–284.

[65] Stanford Legal Design Lab. 2020. Icons for legal help. (2020). https://betterinternet.law.stanford.edu/design-guide/icons-for-legal-help/.

[66] Alan B Tickle, Robert Andrews, Mostefa Golea, and Joachim Diederich. 1998. The truth will come to light: Directions and challenges in extracting the knowledge embedded within trained artificial neural networks. *IEEE Transactions on Neural Networks* 9, 6 (1998), 1057–1068.

[67] Joseph Turow, Michael Hennessy, and Nora Draper. 2015. The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation. *Available at SSRN 2820060* (2015).

[68] UK Information Commissioner's Office (ICO). 2019. *ICO and The Turing consultation on Explaining AI decisions guidance*. Technical Report. https://tinyurl.com/2dk7rvrd Publisher: ICO, (Accessed on 02/23/2021).

[69] UK Information Commissioner's Office (ICO). 2020. *Explaining decisions made with AI*. Technical Report. https://tinyurl.com/ico-explaining-decisions Publisher: ICO, (Accessed on 02/23/2021).

[70] European Union. 2016. Regulation 2016/679 of the European parliament and the Council of the European Union. *Official Journal of the European Communities* 2014, March 2014 (2016), 1–88. DOI:http://dx.doi.org/pri/en/oj/dat/2003/l_285/l_28520031101en00330037.pdf

[71] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, Montreal QC, Canada, 1–14. DOI:http://dx.doi.org/10.1145/3173574.3174014

[72] Maranke Wieringa. 2020. What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 1–18.

[73] Allison Woodruff, Sarah E. Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of CHI2018 (CHI '18)*. Association for Computing Machinery, Montreal QC, Canada, 1–14. DOI:http://dx.doi.org/10.1145/3173574.3174230

Survey Questions:

Q1 Have you heard about the “Right against solely automated decision making" before?

- ○ Yes (1)
- ○ No (2)

Q2 What do you think the “Right against solely automated decision making" means? Pleasebriefly describe your understanding below.

_____

_____

_____

Q3 Please: 1) Briefly describe your understanding of "User Profiling"?
2) Give an example to illustrate your understanding.

_____

_____

_____

Q4 Please: 1) Briefly describe your understanding of "Automated decision making"?
2) Give an example to illustrate your understanding.

_____

_____

_____

Q5 Now we will explain the meaning of the ***“Right against solely automated decision making***". Please either read the description or watch the video carefully. The next few questions will be based on your understanding from the resources below.

Video: youtu.be/BrQMqmPEWQs

Q3.6 In the video, did you spot a computer screen?

- ○ Yes (1)
- ○ No (2)

Q7 Do you have any suggestions to improve this video? Please write them down below.

_____

_____

_____

Q8 Based on the information you have just learned from the video/text description, which of the following statements best explains the “Right against solely automated decision making"? (Choose all that apply)

- ☐ It allows users to request companies to delete their personal data (1)
- ☐ It allows users to request companies to explain how the AI makes automated decisions about them (2)
- ☐ It allows users to object to the processing of their personal data (3)
- ☐ It allows users to request companies to correct their personal data (4)
- ☐ Other (Please Specify) (5) _____

Q9 Based on the information you have just learned from the video/text description, which of the following is an example of the “Right against solely automated decision making"? (choose allthat apply)

- ☐ A college level football player stumbled upon a photo of them on a webpage of a sports magazine. This person wonders how the magazine editor found the photo. (1)
- ☐ An individual stumbled upon the social media settings page that shows an automatically generated profile that the company uses to provide a personalized news feed and other targeted ads. This person requests the company to explain the process of automatic profiling. (2)
- ☐ An individual request the bank to transmit details of this person’s bank transactions to a new Budget planning app. (3)
- ☐ An individual recently switched jobs but still appears on the social events webpage of the previous organization. This person requests the organization to remove such information. (4)

☐ Other (Please specify) (5) _____

Q10
If the companies start to make decisions about you automatically, what kind of decisions do you think might significantly affect you? Please explain briefly.

_____

_____

_____

Q11 Regarding the example from the video about automatically making decisions on bankloans, what aspects of automated decision making would you like to be explained?

_____

_____

_____

Q12 Which of the following aspects of automated decision making would you like to be explained/provided with? (select all that apply)

☐ Company must inform me where or how the company found/obtained my data that is used in making automatic decisions about me (1)

☐ Company must inform me what types of data (e.g., my name, age, address) are used in making automatic decisions about me (2)

☐ Company must inform me how data is used or what algorithm or method is used in making automatic decisions about me (3)

☐ Company must inform me how it ensures that the algorithm or method used in making automatic decisions about me is fair and unbiased (4)

☐ Company must update both my data and the algorithm for accurate decision making as well as inform me about these updates (5)

☐ Company must provide me with customer care services to contest automated decision making (6)

☐ Other, please specify (7) _____

Q13 In your past internet usage experience, have you ever encountered an incident where you could have exercised your “Right against solely automated decision making"?

○ Yes (1)

○ No (2)

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = Yes

Q14 Could you briefly explain this past incident? How could you have exercised this right?

_____

_____

_____

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = Yes

Q15 What are some challenges you think you could have encountered while exercising your “Right against solely automated decision making" for the incident you mentioned in the previous question? Please explain briefly.

_____

_____

_____

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = Yes

Q16 What could be some possible solutions you would suggest to companies in order to overcome the challenges you mentioned in the previous question? Please explain briefly.

_____

_____

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = No

Q17 Based on your understanding of the “Right against solely automated decision making", how do you expect to be benefited from exercising this right online? Please briefly describe below.

_____

_____

_____

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = No

Q18 Based on your understanding of the “Right against solely automated decision making", what are some of the challenges you anticipate while exercising this right? Please briefly describe below.

_____

_____

_____

Display This Question:

If In your past internet usage experience, have you ever encountered an incident where you could hav... = No

Q19 What are some possible solutions you would suggest to companies in order to overcome the challenges you mentioned in the previous question? Please briefly describe below.

_____

_____

_____

Q20 On the scale of 1 to 5, how well were you able to understand the "Right against solely automated decision making"?

	Extremely unclear (25)	Somewhat unclear (26)	Neither clear nor unclear (27)	Somewhat clear (28)	Extremely clear (29)
Understanding of "Right against solely automated decision making" (1)	o	o	o	o	o

Q21 Please briefly describe below your overall understanding of the “Right against solely automated decision making"?