Data Agency Theory: A Precise Theory of Justice for AI Applications

Leah Ajmani
University of Minnesota
Minneapolis, MN, USA

Logan Stapleton
University of Minnesota
Minneapolis, MN, USA

Mo Houtti
University of Minnesota
Minneapolis, MN, USA

Stevie Chancellor
University of Minnesota
Minneapolis, MN, USA

ABSTRACT

Data collection methods for AI applications have been heavily scrutinized by researchers, policymakers, and the general public. In this paper, we propose data agency theory (DAT), a precise theory of justice to evaluate and improve current consent procedures used in AI applications. We argue that data agency is systematically defined by consent policies. Therefore, data agency is a matter of justice. DAT claims data agency ought to be afforded in a way that minimizes the *oppression* of data contributors by data collectors. We then apply DAT to two salient consent procedures in AI applications: Reddit's Terms of Service agreement and the United States's IRB protocols. Through these cases, we demonstrate how our theory helps evaluate justice and generate ideas for improvement. Finally, we discuss the implications of using justice as an evaluation metric, comparing consent procedures, and adopting DAT in future research.

CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing theory, concepts and paradigms**.

KEYWORDS

artificial intelligence, user-generated data, empowerment, consent

ACM Reference Format:
Leah Ajmani, Logan Stapleton, Mo Houtti, and Stevie Chancellor. 2024. Data Agency Theory: A Precise Theory of Justice for AI Applications. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3630106.3658930

1 INTRODUCTION

In 2022, Politico reported that Crisis Text Line–a non-profit SMS suicide hotline in the United States—used one-on-one crisis conversations to train a customer service chatbot. This chatbot was intended as a for-profit spin-off to support the hotline [44]. Users and popular media criticized Crisis Text Line (CTL) for violating the spirit of the hotline and breaching data sharing rules in their Terms of Service [43]. CTL maintained that they did not violate reasonable expectations—hotline users "consented" to a lengthy Terms of Service, which specified that CTL could use data for business purposes. The CTL controversy is just one of many controversies highlighting how data-driven organizations can abuse peoples' expectations of data collection and use [35, 56]. Current Artificial Intelligence (AI) and Machine Learning (ML) technologies often scrape data from public venues, such as social media sites, with no attribution to the original sources. Data also comes from technology users who passively produce detailed artifacts such as behavioral logs, ratings, and personal information [6], which inform popular models like GPT and Bard.

Theories of data labor [6, 45] argue that data contribution is necessary work to fuel predictive technologies. However, those who contribute this data, referenced as *data contributors*, are rarely given their due right to shape data decisions on models [85]. For example, social media users are often the inadvertent contributors to AI/ML models; emerging generative AI models, such as GPT-4 and Dall-E, are primarily powered by the public contributions of internet users, writers, and artists [45]. However, publicly available data is often authored and disseminated in a specific context. People post content with an *imagined audience* of their peers or other users, not data collectors [51]. Previous work has explored how taking this data out of its intended context can increase the risk of dehumanization [15], violate one's understanding of privacy [58], and contribute to entrenched social injustices such as racial disparities [18].

Traditional models of consent have been argued to mitigate these failures; however, we think they fail to capture the systemic nature of consent in data-centric environments. AI applications often rely on massive amounts of data, such that gaining the ideal of individual informed consent may not be feasible at scale [76]. Social media data sites, which often make their data publicly available, rely on Terms of Service (TOS) agreements that users consent to before they can use the platform. Arguably, these agreements do not create consentful situations. TOS agreements have been scrutinized for being opaque [16, 26, 79, 93] and giving platforms unfettered access to data once consent is given [4, 48].

While AI scholars and activists have called for justice-centric perspectives [14, 42, 82], AI practitioners have not coalesced on a moral and pragmatic framework to accomplish these goals. Currently, justice in predictive systems (AI, ML, algorithms, etc.) is

FAccT '24, June 03–06, 2024, Rio de Janeiro, Brazil

ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3658930

used as a valuable umbrella term for describing a category of interrelated desiderata, such as equity [21], fairness [40], and harm mitigation [42]. Umbrella terms are helpful insofar as they promote accessible and intuitive understanding [71]. However, in AI justice literature, these broad strokes require researchers to abstract away from specific situations. AI ethicists must create a theory of justice specifically for data-centric policies and procedures in an AI pipeline. From the Writer's Guild of America strike surrounding AI threats [89] to organized AI advocacy [66], current events indicate a lack of AI justice frameworks that sufficiently capture the collective struggles humans have with predictive technologies. Precise theories create more ground for productive implementation and action [76], giving designers and developers the scaffolding to create more just systems. How can we align AI to justice when we have not established what justice is in this context?

We present *data agency theory*: consent policies constrain data contributors as a group by negotiating their agency around a dataset they created. Data agency theory (DAT) is built on emerging work on data labor [6, 45] and feminist work on redefining the role of consent as a systemic procedure [10, 30, 47] rather than an individual agreement. We adapt feminist philosopher Iris Young [90]'s theory that justice is the ridding and redressing of group-based oppression through institutional change. Theory adaptation is an applied ethics method that translates broad societal norms into technical practice [76]. Specifically, we adapt Young's theory of justice to AI applications by:

(1) Describing the systemic relationship between consent and agency in data settings
(2) Establishing data agency as a necessary dimension of justice in predictive systems

We conclude that justice in a predictive system demands considering how institutional routines (i.e., consent procedures and terms of services) transform agency at a group level. In other words, data agency is a *contributor* to justice and a *product* of consent policies in a predictive system. We highlight the power of having a problem-specific theory of justice by applying data agency theory to two salient data consent procedures: (1) social media data sharing and (2) human subjects research consent procedures. We demonstrate that a precise theory of justice carries unique evaluative and generative power. Finally, we discuss key opportunities for using data agency theory to compare consent procedures, generate new designs, and inform future research. DAT serves an urgent purpose in FAccT and AI development communities by formalizing our ethical intuitions about consent and agency into an operationalizable theory.

2 RELATED WORK

In this section, we outline the relevant work on consent and justice, particularly in data settings, such as social media, ML, and AI applications. DAT is heavily inspired by (1) work on data labor [6, 45, 85] and (2) feminist philosophy on consent, agency, and justice [22, 38, 90]. Therefore, we scope broad terms like "consent" to specific definitions from previous work. Table 1 provides a full overview of DAT-related terms.

2.1 Consent in Sociotechnical Settings

Our contemporary understanding of consent originated in feminist and sex-education movements. Consent, as a sociological construct, is mainly centered around those who are at a power deficit due to deeply-rooted institutions such as patriarchy, heterosexism, and toxic othering [1, 24, 47, 74]. Furthermore, it is widely accepted that consent is an impactful mechanism for navigating power imbalances in high-stakes situations, such as claims over one's body or sexuality [9, 37]. Social notions of consent, particularly in sexual contexts, have been iteratively progressing for decades, implying that consent is a dynamic concept [20]. Consent must be made free of coercive forces, being given enthusiastically and willingly [9]. However, consent is not a one-time nor binding agreement but rather an ongoing negotiation dependent on numerous contextual notes [8, 10, 53, 57].

While consent is a well-defined term in some contexts, such as law, it is still a social construct—it is dependent on ambiguous and contextual factors. In social media contexts, communicating consent is multi-faceted. For example, users may build a mental model of an *imagined audience* that determines the terms of their consent [51]. Previous work has noted that consent is a sociotechnical gap in online spaces [2, 38, 76]. For example, the power asymmetry between a data collection platform and a user in a technical setting is not comparable to humans in a social setting, such as a sexual encounter. Moreover, because there is no one-on-one interaction, as with sexual encounters, the implications of consent in technology systems are unclear. In technological systems, it can often be seen as a "moral magic" [4] that gives data collectors unfettered access and control over large swathes of data.

Our work builds off of consent design by suggesting a meta-theoretical framework to evaluate and compare prevalent consent procedures, such as heavily scrutinized Terms of Service agreements [26, 49, 64, 79]. We specifically take a sociological definition of consent as an agreement between data contributors and data collectors. Moreover, we frame consent in terms of justice, allowing us to reason about consent agreements as an institutional routine rather than an individual agreement. We use data agency theory to evaluate these different mechanisms for consent and generate future improvements. Specifically, we evaluate these mechanisms along the dimension of justice.

2.2 Justice in AI/ML

Originating from Justinian Law in the Byzantine Era, justice is the idea of giving every individual their due [39]. Justice is the opposite of arbitrariness; it is systemic fairness that has been deliberately designed through laws and policy. In AI ethics, there has been a long-standing interest in algorithmic justice and fairness as fundamental values. Justice has been invoked both as a property of a model–"algorithms should be just" [23, 40, 52]–and as a pre-existing value in systems where AI is utilized–such as the criminal justice system [80, 91]. Moreover, data collection, maintenance, and use procedures are often central to discussions on AI/ML justice and broader ethics [55, 82, 86]. This makes sense as the data used for AI applications are generated by humans [15], meaning that intricate networks of stakeholders and sociopolitical dynamics are at play. For instance, Vincent et al. [85] found that social justice

Term	Definition	Citations
Data Contributors	Users who either passively or actively provide data to a collective "dataset," such as a platform's API or a research dataset.	Arrieta-Ibarra et al. [6], Li et al. [45]
Data Collectors	Entities that pool data contributions together into a usable dataset. These could be large entities, such as tech companies, or single researchers.	Arrieta-Ibarra et al. [6], Li et al. [45]
Institutional Routine	Sweeping, standardized policies that are the norm for someone engaging with an institution, such as a user engaging with a social media platform	Young [90]
Consent	An institutional routine of agreement between data contributors and data collectors.	Im et al. [38]
Data Agency	An individual's capacity to shape the action around the data they created.	Davies [22], Vincent et al. [85]
Justice	The rectification of the oppression of social groups caused by institutional routine.	Young [90]

Table 1: Glossary of central terms to data agency theory. DAT is heavily inspired by (1) work on data labor in predictive systems and (2) feminist work on consent, agency, and justice.

mechanisms, such as worker strikes, have been adapted to give data contributors "levers" over data collectors and platforms.

We build on this work by proposing a theory central to designing and evaluating data management mechanisms across the AI pipeline. Furthermore, we adopt the Rawlsian view that justice and fairness are symbiotic in AI systems: increasing the justness of a model will, in turn, increase fairness [68]. However, we argue that Rawlsian distributive justice paradigms, often used in ML fairness and AI justice research [52], do not sufficiently mirror the structural inequities entrenched in data pipelines. Instead, we adopt Iris Young's perspective, a seminal critic of distributive justice paradigms, that seeking justice is rectifying the oppression of social groups caused by structures and systems [90]. We ground our theory in the parallels between Young's conceptualization of justice and current thought paradigms surrounding data-centric applications, such as power asymmetries, stakeholder groups, and sweeping consent policies.

3 DATA AGENCY THEORY (DAT)

3.1 Theory Overview

Justice is often framed as a distributive goal in predictive systems [42]. AI justice typically considers whether favorable outcomes are distributed equitably amongst a key population. Critiques of distributive justice note that it does not sufficiently consider systemic factors and structural inequities [90].

In this section, we establish **data agency theory (DAT)**—consent procedures define data-centric social groups and dictate each group's capacity to shape data action. Here, we use theory adaptation methods to translate Young's theory of societal justice into an operationalizable theory of data justice. DAT is a framing of data agency as a *product of* consent policies and a *contributor to* justice within a predictive system (Figure 1). DAT allows us to evaluate the justness of a predictive system by dissecting the data consent policies used. We outline the two fundamental premises of DAT. Each premise implies that AI/ML fairness research, a subset of responsible AI, must adopt two paradigm shifts to meaningfully progress toward the desired goal of measurable proxies for justice and fairness in an AI system.

3.1.1 Defining Data Agency. We define **data agency** as an individual's capacity to shape action around the data they create. More data agency gives contributors more control over how their data is used [45, 85].

It is important to distinguish agency from power. While agency focuses on an individual's capacity to act, power is something one has over another being, group, or circumstance [22]. In other words, agency is action while power is domination [50] and inescapably relational. We take the feminist view [22, 50, 90] that agency begets power; eventually, one can wield agency in a way that exerts power. Conversely, power can be distributed in a way that affects the agency of numerous people. Iris Young [90] focuses on what happens when power distributions affect people's agency in specific social or identity groups. Young argues that when a lack of agency is *systemic*, it affects a whole community of people—and, therefore, is a form of *oppression*. In turn, seeking justice involves redressing harm to specific social groups caused by institutional routines. We argue that this conceptualization of justice accurately describes data-centric technologies, such as predictive systems. Therefore, we explore what affects agency in data pipelines and how that ultimately translates to justice.

3.2 Premise 1. Consent outlines agency in a systematic way

First, we establish the relationship between consent and agency. In social settings, consent has historically tried to create more equitable power dynamics in complicated contexts, such as sexual encounters [88]. To this end, consent has been posited as a dynamic and mutual agreement free of coercive forces [8]. Consentful situations give individuals the ability to express agreement or disagreement through their consent. This sociological definition of consent describes online data-sharing agreements. For example,

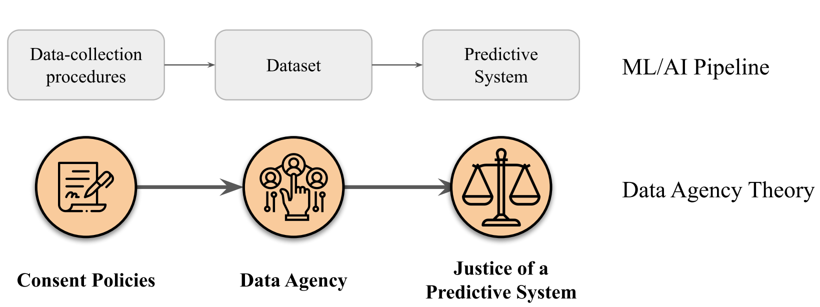

Figure 1: Overview of the central constructs in data agency theory (DAT). We argue for a Youngian conceptualization of justice where data agency is a *contributor* to justice and a *product* of consent policies.

individual privacy settings on Google allow users to opt into (i.e., consent to) data sharing with third-party advertisers. By taking that option, users make a dynamic and contextual agreement with Google to provide data.

Data contributors have a baseline amount of agency; they can choose which sites to use, how to present their profiles, and what they post. However, agency is also defined by the consent procedures they encounter in computing systems. For example, in an informed consent form, a research participant is told they can withdraw at any time. Therefore, the research participant (i.e., the data contributor) is informed of an action they can take. However, consent can be used to diminish agency in nefarious ways. For example, TOS for a social media site may outline that the data collector (i.e., platform) is allowed to share user data with any third parties. This is anti-agentic as data collectors can asymmetrically and unilaterally take away actions from data contributors.

Agency is more than individual actions and relationships with data collectors; it is a product of systemic factors (e.g., consent procedures). In predictive systems research, consent procedures around data are often framed at and for individuals: *Are we giving every individual their due right to their data?* Prior work supports the position that data rights are individual rights [82]. Data is often regulated as property that an individual has rights over rather than a collective resource that users contribute to [73]. While individual rights are a fantastic foundation for reasoning, consent procedures around data are relational and, therefore, may not map to the societal consequences of technology.

We posit that consent defines an individual's capacity to shape action with a predictive system, but it does so according to social groups rather than individuals. For example, current models of consent procedures (e.g., blanket TOSes) are designed for data contributors as a group rather than a single individual. Individuals cannot negotiate TOSes; their choice is an all-or-nothing adoption of policies decided by data collectors. While the datasets necessary for AI/ML applications are often the work of millions of data contributors, these broad policies conflict with the idea of individual data or individualized consent. Instead, predictive systems often view "datasets" as an amalgamation of data contributions that go beyond a single individual. Therefore, consent as an agreement between two individuals fails to capture the structural and systemic nature of data-centric settings. Alternatives to individualized consent models for technology have been proposed [38]—and we deliberately call on these alternatives as a new model for mapping consent to more relational principles and societal justice. While an individual should clearly have access to individual actions (such as refusal to use a system [46, 85], these individual actions must be defined and implemented in *systemic and procedural ways* to push against the consequences of one-size-fits-all consent procedures. In sum, consent defines agency in a systematic way, meaning that there are opportunities to either negatively neglect or positively protect the agency of data contributors through consent policies.

3.3 Premise 2. Agency is a matter of justice.

Next, we establish the relationship between agency and justice. This claim is an adaptation of Iris Young's critique of distributive justice, which states that the distributive paradigm in justice theory does not sufficiently recognize the impact of systemic factors, such as inherent power asymmetries and undemocratic policies [90]. Young's critique is relevant to data-centric systems, which are ripe with power imbalances that we have outlined previously [45]. We believe that Young's formulation of justice better mirrors how ML systems view stakeholders.

To translate this critique into a precise theory of justice for AI applications, we adopt Young's core principle: *equity for individuals should not override the rectification of group-based oppression* [90, p. 34]. As we argued in Premise 1, ethics in predictive systems often rely on social groups, such as data contributors, tech companies, and developers, rather than individuals. Therefore, we establish that data contributors are a social group and consent procedures are institutional routines—sweeping, standardized policies that a user must engage with. These consent processes are passed down by platforms (data collectors) to data contributors. Justice is a matter of identifying and rectifying the systemic oppression that data contributors face.

Agency—specifically agency around data-centric actions—is crucial to justice in predictive systems. Recall that agency is the social power to define one's actions and behaviors [63]. For DAT, we conceptualize agency as a question: *what is a data contributor's capacity to give direction and shape to a data-centric action?* Inspired by activity theory [11, 25], we presume that agency is not a starting place. Rather, it is created and transformed through action [75]. Viewing agency in this way captures how that agency is transformed throughout the AI pipeline; we can understand how actions around

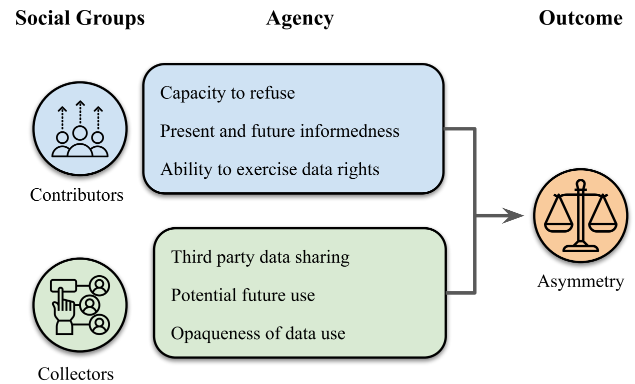

Figure 2: Operationalizing DAT requires outlining relevant social groups, mechanisms to shape data action, and resulting asymmetry.

data representation, use, and dissemination change agency. Furthermore, this conceptualization aligns with Young's definition of justice and the ability of a group to participate in decisions that impact their lives [90]. Therefore, mechanisms that affect data agency consequently affect the justness of a system.

This claim points to a new approach in AI justice and ML fairness to conceptualize *justice as ridding a system of oppression.* Researchers and developers must build a measurable and tractable conceptualization of justice in predictive scenarios. This shift towards social justice paradigms, such as Iris Young's, gives us grounds to consider the systemic nature of consent in data settings; consent policies outline a network of relationships with power (im)balances and domino effects across social groups.

4 OPERATIONALIZING DATA AGENCY THEORY

We see two applications of our theory concerning consent and justice: evaluative and generative.

4.1 Evaluative Power

A primary benefit of DAT is *evaluative* power, or the ability to assess the justness of consent processes. These assessments can happen against the procedure by itself or appraising it against other alternatives. When consent procedures focus too closely on individual action, minutiae can make direct comparisons of people or contexts difficult and can prevent generalized takeaways. Conversely, ML has prioritized abstract concepts such as fairness, equity, and justice, but these concepts can be difficult to operationalize in practice. Here, we translate the premises of data agency theory into actionable dimensions of consent policies. In sum, operationalizing DAT requires outlining relevant social groups, mechanisms to shape data action, and the asymmetry between the two (e.g., Figure 2).

Premise 1. Consent outlines agency in a systematic way. Recall that agency refers to one's capacity to shape present and future actions about the data they contributed. DAT allows us to assess agency in consent policies by highlighting the capacity for data action enabled or limited by policy. Furthermore, DAT focuses on what happens when unquestioned power distributions affect the agency of specific social or identity groups. At that point, the lack of agency is *systemic* and, therefore, is a form of *oppression* [90]. DAT gives a name to and a lens for describing the outcomes of consent policies on data contributors and collectors.

Premise 2. Agency is a matter of justice. In DAT, seeking justice is redressing harms to specific social groups caused by institutional routines. We argued in Section 3 that this conceptualization of justice better suits data-centric technologies, such as predictive systems. Consent policies are "institutional routines" as they are sweeping, standardized policies that are the norm for a user experience. We note how these policies contribute to structural injustice in predictive systems along dimensions of asymmetry. As Young [90] says, "*for every oppressed group there is a group that benefits from the other group's oppression.*" Therefore, injustice feeds off of structural asymmetry between social groups.

4.2 Generative Power

DAT can also serve as a *generative* framework, providing opportunities to consider how policies could be improved to be more consentful and, therefore, more just. We envision DAT's generative power to be broad. In some cases, our design suggestions reimagine consent throughout the design and implementation of the AI pipeline. There may also be opportunities where generative suggestions must move beyond technical or engineering solutions. As a theory of justice, DAT often points to policy, structural, or other changes that must happen to allow for a more just system.

5 CASE STUDY

Next, we present two case studies of popular data collection policies: social media data sharing and human-subjects data collection. Using data agency theory (DAT), we evaluate the justness of these two policies and generatively demonstrate how they can be improved. Table 2 summarizes our two evaluations into their key components.

5.1 Social Media Data Sharing in Terms of Service

> *Access, search, or collect data from the Services by any means [automated or otherwise] except as permitted in these Terms or in a separate agreement with Reddit. We conditionally grant permission to crawl the Services in accordance with the parameters set forth in our robots.txt file, but scraping the Services without Reddit's prior consent is prohibited. [69]*

Terms of Service (TOS) agreements are policies a technology user agrees to by using an application or service. The user is bound to the policies in these documents, which codify data procedures—such as data use, sharing, storage, and collection. In the example above, we show Reddit's policy for data sharing within their Terms of Service (TOS) agreement, but similar agreements are used on Instagram,[1] and Twitter[2].

Data sharing policies are standard in social platform TOS agreements; we argue this is partly because TOS agreements facilitate the use of data across many applications relevant to ML and AI.

[1] https://help.instagram.com/
[2] https://twitter.com/en/privacy

Setting	Collectors	Contributors	Asymmetry	Impact
Social Media TOS	Reddit & Third Parties	Redditors	Consent to third parties given by Reddit, not Redditors	Redditors cannot properly contextualize their data during consent and use
Human Subjects Consent	Researcher	Research Subjects	Researchers have deeper technical knowledge of potential future data use	Subjects are most informed during the time of study consent, but this is never re-evaluated

Table 2: Summary of DAT evaluations.

Internally, data may be shared to facilitate targeted advertising [41], personal recommendation [92], and business relationships with outside parties. Platforms such as Twitter and Reddit have APIs that permit data sharing with third parties. This data access has been collectively recognized for its value to AI and ML practitioners [7, 62, 64, 93]. However, TOS agreements have been critiqued for being illegible, dense, and opaque [26, 49]. They are often presented in long forms that users agree to without reading or understanding, requiring high reading levels and comprehension. At their worst, social media platforms have used TOS agreements to justify sharing data in nefarious ways, exemplified by the Facebook-Cambridge Analytica [35] and Crisis Text Line [44] controversies.

5.1.1 Evaluation. DAT provides an effective lens to explain what parts of TOS data-sharing policies violate notions of consent and justice. We operationalize DAT by outlining the main social groups and their capacity to shape data action.

Data collectors. Focusing on Reddit's policy, this TOS authorizes third parties to scrape Reddit by getting prior consent from the platform (Reddit). Therefore, we have two data collectors: the platform and the third party. Here, scraping is a data action that third parties can do within certain bounds specified in an agreement that is unknown at the time of consent. The platform Reddit dictates those bounds and, thus, situates itself as a data collector and policymaker.

Specifically, third parties can collect the user-generated content on the website if they ask Reddit first or abide by the `robots.txt` file. Reddit can choose to give consent or not. Therefore, we have two mechanisms to shape data action: (1) the platform's capacity to govern and gatekeep access to user-generated data and (2) a third party's capacity to scrape the user-generated data.

Data contributors. Next, we consider data contributors and their capacity to shape data action. Data contributors are users of Reddit who actively contribute data through posting, commenting, or upvoting content. While data contributors could act as a third party and scrape the website, they most likely do not have the technical skills or time to do so. Moreover, data contributors cannot shape data-sharing decisions once they have consented and produced data. Finally, contributors cannot change, alter, or adjust these policies.

Asymmetry. If we care about justice here, DAT suggests we must consider the systemic asymmetries this policy creates and propagates. Recall that Reddit can govern and gatekeep the dataset, a collective of user contributions hosted by the platform. However, there is asymmetry insofar as data contributors do not have a similar capacity to govern and gatekeep their own data from third parties. Through this TOS agreement, Reddit has taken matters of data control away from data contributors (users of the site) and now brokers data-sharing agreements with third parties. Implicitly, this means that the platform has control over *who* can scrape data from the platform, but data contributors do not have similar control over the data they produce.

Moreover, this policy propagates the existing structural asymmetry between data contributors and collectors. In prior work on data labor [6, 45], we know that data contributors are rarely given credit for their work and often left to their own means of gaining leverage [85]. Consent does not remediate this pre-existing asymmetry when considering agency and Reddit's policy. There are no mechanisms for data contributors to regain ownership or exert leverage. For example, individuals cannot erase their own data from these datasets or withdraw their consent. In fact, deletion bots have emerged as a workaround to this asymmetry. DAT suggests that this TOS policy infringes on justice by allowing a third party and platform to take a data-centric action while data contributors have no information or recourse.

5.1.2 Generation. How can consent be renegotiated to improve data agency in this situation? Recall that control is being given to a third party without mechanisms for data contributors to engage with Reddit, something prior work has likened to an unfettered stream of data from users [70].

To prevent an unjust ecosystem of data agency, Reddit and platforms like it could make their data-sharing permissions available to data contributors. This would involve designing consent procedures where users can tailor their data-sharing permissions according to their preferences. Building on Fiesler and Proferes [27], there are several opportunities where users have expressed differences in their preferences for sharing, including the content of their posts and intended use by third parties. DAT informs how Reddit could allow users to negotiate data agency through more consentful interfaces when contextual considerations are most relevant.

Pragmatically, what might this look like? One opt-in could happen at the profile level, where data contributors (who we normally call users) can make different profiles with different sharing permissions. Similar to the intentions behind throwaway accounts on Reddit or fake private Instagrams (Finstas) [36], users can have profiles where their data is only shared with specific institutions, such as researchers or non-profits. Taken a step further, Reddit could design mechanisms where users can consent to data-sharing post-by-post, which could be supported by future research and design propositions similar to Im et al. [38]'s design ideas in affirmative consent. To make this easier, users could set specific words where, if present in their post, they may not want to share or subreddits of content they wish to withhold.

5.2 Human-Subjects Research for AI: Informed Consent

> *The informed consent process involves three key features: (1) disclosing to potential research subjects information needed to make an informed decision; (2) facilitating the understanding of what has been disclosed; and (3) promoting the voluntariness of the decision about whether or not to participate in the research [59]*

Another data collection mechanism in AI is using human subjects to provide data dumps of their social media use [17]. Human subjects are the United States (US) concept of a research participant or subject formally involved in a research study. This method allows researchers to cross-reference an individual's social media data with medical records [12], mental health assessments [31], or in-person interviews to understand their behavior. While human subjects research is a less common technique to gather data than social media data scraping, ethicists have noted that human-centered concerns are relevant in AI contexts that use public data [15, 86]. This section evaluates the US principles for Institutional Review Board (IRB) consent as a broader indicator of human subjects procedures for AI.

In human subjects research overseen by an IRB, consent is a standardized and regulated practice. IRB standards are based on principles outlined in the Belmont Report (1979) [83]: respect for persons, beneficence, and justice. Under the "respect for persons" principle lies the specific guidelines on practicing *informed consent* [59]. Consent is about respecting an individual. This practice contrasts social media TOS agreements (as just discussed), which often frame consent as procedural.

While the IRB has defined key features of informed consent [59], operationalizing informed consent is still a vigorously debated topic within the research community. In health research settings, there are conflicting standards of readability versus comprehensiveness, so it is unclear whether subjects are fully informed [32, 34]. These problems persist in sociotechnical research settings, where researchers have called for expansions of informed consent. For example, Im et al. [38] distinguish informedness as just one of five major principles for consentful situations in sociotechnical settings. Moreover, the notion of "informedness" is dependent on how we view knowledge, which is often viewed through a Eurocentric lens [81]. While informed consent provides a foundation for ethical research practices, it is unclear whether it truly captures agentic consent practices.

5.2.1 Evaluation. DAT allows us to unpack the nuances of informed consent as it relates to AI and ground them in justice. Focusing on the definition of informed consent from the IRB (quoted above), this policy states that informed consent is garnered so long as the researcher promotes disclosure, understanding, and voluntariness at the time of consent. Once a subject gives consent, researchers can collect the subject's data. Once a subject has consented to research participation, they establish a data-sharing relationship, with the subject becoming a **data contributor** and the researcher as a **data collector**.

Asymmetry. By its namesake, informed consent is intended to minimize the information asymmetry between researchers and subjects, the two parties most prominently featured in scientific relationships. There is an inherent asymmetry because a researcher has designed the study. This information disparity can be exacerbated by a researcher's technical knowledge of the situation, ability to shape future research use, and need to meet publication standards.

Under DAT, evaluating the consent form is necessary to evaluate the justness of any resulting AI systems. Positively, the consent form attempts to increase a participant's agency by providing (1) information about what they consent to and (2) a reminder of their ability to refuse, withdraw, or alter participation. The design of the consent process, whether through written or verbal consent, is intended to increase a participant's agency when consent is requested. It is important to note here that informedness is not generated by simply providing information [34]. Rather, the information should be presented to promote understanding of the human subject.

However, DAT also asks us to consider the process of shaping data action rather than a one-time mechanism. Thus, under DAT we would build on this policy to give participants data action in the long term, similar to the data actions that researchers have. Take, for example, a research study with an IRB for building a dataset for researchers to use in the long-term. If participants consent once, researchers can continuously update the dataset and use it for future projects. Although the researchers at the time may have had the best of intentions in developing the dataset, the control of the dataset rests in the hands of the researchers. This means there is asymmetry created when data contributors will not be asked to reconsent to transform their data into a new context. Therefore, consent with data contributors would need to be iteratively adressed, rather than mitigated once. For example, subjects would need to be informed about follow-up work that uses their data. DAT would say, therefore, we are not sufficiently remediating asymmetry in this situation.

In this situation, *communicating* the intended use of a dataset through a consent form relieves asymmetry between contributors and collectors in scientific research, and is a positive foundation for consent. However, DAT points to an asymmetry when future use of datasets are not explained or made clear.

5.2.2 Generative. Under DAT, informed consent creates a space where asymmetries are partially minimized. However, we note two main areas for improvement: (1) ensuring that the relevant information is interpretable and (2) expanding informed consent past the initial data collection.

As noted in our evaluation, *how* a subject is informed is critical to reducing the inherent information asymmetry between researcher and subject. This sentiment is already echoed in informed consent procedures; institutions and governments often provide approved examples of interpretable consent forms. However, recent work has noted that standards of "readability" often conflict with other standards of comprehensiveness and length [32]. Informed consent procedures could pacify this tension by using standardized readability tests, such as the SMOG formula [49], or providing multiple versions of consent forms, such as a summary with key points and a more comprehensive version. This would allow researchers to capture the principles and standards of "informedness," without relying on solely on institutional discretion or subfield norms [3].

Furthermore, discussions about study information and voluntariness are only happening at the time of initial consent, leaving few standards for consent withdrawal and sustained information. In research studies, particularly in AI and mental health, data collected from human subjects helps generate publicized benchmark datasets that are used in future research, such as the eRisk dataset [60]. These datasets can contain sensitive information, such as mental health diagnostic assessments. Because they are publically accessible to researchers, a subject's data could be used in a different context than the initial research study where their consent was requested. We suggest that the human-subjects consent procedure be improved by replicating initial informed consent throughout the research pipeline. For example, researchers could discreetly inform participants about how their dataset will be used in future publications. Ideally, this check-in also includes a mechanism for participants to withdraw their consent, either by removing their data from the public dataset or declining to be included in specific research projects.

6 DISCUSSION

Recent work in AI has focused heavily on expanding our evaluation metrics beyond quantitative ones, such as accuracy, and towards human-centered concepts, such as bias, fairness, and equality [14, 87]. These metrics affect how we normatively evaluate technological systems [42]. For example, understanding the fairness of an ML model can help us evaluate whether it is appropriate to use in high-stakes settings, such as the judicial system. Justice allows us to critique how our algorithms uphold and underpin institutional inequities, particularly in high-stakes spaces such as mental health support [61], criminal justice [91], and child-welfare [78]. Therefore, incorporating justice into our evaluation criteria is essential for minimizing sociotechnical gaps around values. In other words, the absence of justice considerations will lead to technologies that unintentionally cause harm.

Justice is a *fundamental* societal value that permeates how we conceptualize welfare [67, 84], morality [28, 77], and relationships with fellow society members [33]. Despite being an intuitive concept of "right" and "wrong," justice is difficult to operationalize as an evaluation criterion [42]. In this paper, we propose data agency theory (DAT) as a guide for evaluating the justness of consent procedures as they relate to predictive systems. DAT allows us to evaluate the justness of systems in situations where real-world policies impact users. For example, we show that Terms of Service agreements make Reddit a black-box data broker, which impacts justice across the AI pipeline. In this section, we describe key opportunities for data agency theory.

6.1 Data Agency Theory as a Comparative Framework

One reason why consent is so nuanced is that it relies heavily on context. As consent theories from feminism and sexual liberation note, the context in any given interaction is constantly evolving and is an ongoing negotiation [8, 10, 53, 57]. We adopt this view of consent and apply it to two examples of AI-relevant settings—social media data use and human subjects research— to reason about the justness of these consent procedures. We note that human-subjects research consent (informed consent) actively tries to minimize the information and power asymmetry between researcher and subject. Meanwhile, Reddit's Terms of Service agreement maximizes this power asymmetry. In this way, data agency theory can serve as a framework to compare and contrast various consent procedures.

In AI settings, relating various consent procedures can help us identify a shared set of explicit norms. If informed consent is the ideal, how can we replicate it in other data collection methods? If individual informed consent is not feasible at scale, what are alternatives that still protect data subjects?

A comparative framework on consent in AI settings can particularly help with the gap between our values (what we should be doing) and our methods (what we can do). Chancellor [14] notes this as a prevalent gap in human-centered machine learning, and Sisk et al. [76] note it as a specific problem with consent practices. In this paper, we frame justice as a desired value and scope Young's justice theory down to apply to consent procedures. Moreover, we leverage data agency theory to evaluate how various consent procedures impact justice.

6.2 Data Agency Theory for Designers

There is enormous potential to support data agency through design work and HCI methods. We define agency in terms of one's capacity to shape action. In system design, a user's agency is often influenced by which features are designed and how they are presented within the system. Could designers increase a user's data agency by providing them with certain data-related features?

For example, in our case study on Reddit's Terms of Service agreement, we suggest designing profile-level features where users can tailor their data-sharing permissions. These design recommendations echo current non-traditional social media use, such as the creation of Finstas [36] or throwaway accounts on Reddit. We suggest that intentionally designing for these uses could increase a user's agency over their own data by giving them new sociotechnical mechanisms to opt into (i.e., consent to) data sharing. Moreover, HCI researchers could build off of previous work on user's perceptions of data sharing. For example, Fiesler & Profores [27] found that Twitter users had a series of contextual notes that impacted their feelings about contributing data to research. Are there new designs, such as the participatory governance of consent policies or up-to-date mechanisms for data withdrawl [38], that make these contextual notes transparent? Our theory could ground future work that explores how these contextual notes interact with a user's sense of agency.

6.3 Data Agency Theory for Research Ethics

At its core, data agency theory ties together notions of justice, consent, and action. Within the scope of AI applications, consent procedures relate to work in HCI, ML, and other areas. However, with ambiguous concepts, such as justice, it can be difficult to find common ground across these disciplines and engage in meaningful discussion. Justice is often thought of as a wicked problem [19]: a problem that cannot be clearly divided into a problem statement and problem solution. Therefore, having well-defined, agreed-upon definitions is a necessary step in iterating on solutions and achieving progressive improvement [72]. In this paper, we offer a conceptualization of consent as a mediator of agency and means toward more just practices. At a broad level, we hope this contributes to shared vocabulary technologists can rely on when discussing justice across various disciplines.

In AI and ML settings, such as NeurIPS, researchers could use data agency theory to inform discussions of their own consent procedures. Recently, ML venues have required broader impacts or ethics considerations statements. Prunkl et al [65] note that these statements can help institutionalize ethics and inform a clear set of standards for the field. However, Nanayakkara et al [54] found that broader impact statements were scattered across topics and did not meet the self-stated goals of the venue and larger research community. DAT could scaffold discussions of justice in the ML community, thereby institutionalizing it as a normative value.

Finally, we see opportunities for data agency theory to ground critical theories around emerging corporate AI technologies. In addition to academic inquiry into the injustices AI can promote [5, 18], the internet is filled with stories of growing distrust, experienced harms, and personal risks that individuals have confronted. In 2018, the Financial Times declared "Techlash" as the word of the year: the growing public animosity towards large platform technology companies [29]. Through critical technoculture discourse analysis (CTDA), Andre Brock [13] underpins the need to study how these technology practices intersect with identity, class, and power. In data agency theory, we bring relational power to the forefront of the conversation around AI and ML practices. We argue that the inherent power asymmetry between data contributors and data collectors should be ameliorated through consent procedures. We hope that future critical theories expand DAT to explore other institutionalized power asymmetries that are worsened by poor data collection procedures, such as those rooted in identity disparities or classism.

7 LIMITATIONS AND FUTURE WORK

A major limitation is that our designs are speculative reimaginations of current consent procedures. Future work could operationalize our design recommendations to evaluate their feasibility and implementations. Moreover, this paper offers a theoretical contribution and demonstrates its evaluative and generative power. Future work could explore empirical substantiations of our theory, such as simulating data actions in large datasets or interviewing data contributors about their perceptions of agency. Finally, throughout this paper, we use a sociological definition of consent. We realize that this definition can be messy; it is ambiguous and highly contextual. Future work could explore how unambiguous definitions of consent, such as legal definitions, may help users regain agency in courtrooms and policies.

8 CONCLUSION

This work offers a precise theory of justice for predictive systems that use human-generated data. Data agency theory states that consent procedures systemically affect an individual's agency and, therefore, affect the justness of a predictive system. We apply data agency theory to two cases: (1) Reddit's terms of service and (2) IRB informed consent procedures. In both cases, we unpack the social roles outlined by the consent policy and the asymmetry between those social roles. We discuss the potential for data agency theory to ground meta-theoretical conversations around consent and justice. We hope that, despite the contextual nature of consent, DAT gives researchers the tools to unpack the constructs and relationships that continue to propagate injustice throughout our data pipelines.

9 STATEMENTS

Ethical Considerations Statement. As a theoretical paper, this work did not involve any human or data subjects. Therefore, there were minimal ethical implications for our methods, and we did not seek IRB approval.

Positionality Statement. All of the authors of this work have experience using data from public sites, such as social media, for their research. Therefore, it is important to note that the researchers of this work have reaped benefits from broad consent procedures.

Adverse Impact Statement. Misinterpretations of our work could lead to a dilution of justice in AI systems. We are not suggesting that reimagined consent procedures will solve all structural problems with AI systems. Therefore, adversarial actors could use our paper to shroud non-consent-based injustices.

REFERENCES

[1] Antonia Abbey, Pam McAuslan, and Lisa Thomson Ross. 1998. Sexual Assault Perpetration by College Men: The Role of Alcohol, Misperception of Sexual Intent, and Sexual Beliefs and Experiences. *J. Soc. Clin. Psychol.* 17, 2 (June 1998), 167–195.

[2] Mark S Ackerman. 2000. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *Human–Computer Interaction* 15, 2-3 (Sept. 2000), 179–203.

[3] Leah Hope Ajmani, Stevie Chancellor, Bijal Mehta, Casey Fiesler, Michael Zimmer, and Munmun De Choudhury. 2023. A Systematic Review of Ethics Disclosures in Predictive Mental Health Research. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1311–1323.

[4] Larry Alexander. 1996. The Moral Magic of Consent (II). *Legal Theory* 2, 3 (Sept. 1996), 165–174.

[5] Julia Angwin, Jeff Larson, Lauren Kirchner, and Surya Mattu. 2016. Machine Bias. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Accessed: 2023-3-14.

[6] Imanol Arrieta-Ibarra, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier, and E Glen Weyl. 2018. Should we treat data as labor? Moving beyond "free". *AEA Pap. Proc.* 108 (May 2018), 38–42.

[7] Jenae Barnes. 2023. Twitter Ends Its Free API: Here's Who Will Be Affected. *Forbes Magazine* (Feb. 2023).

[8] Melanie Beres. 2010. Sexual miscommunication? Untangling assumptions about sexual communication between casual sex partners. *Cult. Health Sex.* 12, 1 (Jan. 2010), 1–14.

[9] Melanie A Beres. 2007. 'spontaneous' sexual consent: An analysis of sexual consent literature. *Fem. Psychol.* 17, 1 (Feb. 2007), 93–108.

[10] Melanie Ann Beres. 2014. Rethinking the concept of consent for anti-sexual violence activism and education. *Fem. Psychol.* 24, 3 (Aug. 2014), 373–389.

[11] Olav W Bertelsen and Susanne Bødker. 2003. Activity theory. *HCI models, theories, and frameworks: Toward a multidisciplinary science* (2003), 291–324.

[12] M L Birnbaum, S K Ernala, A F Rizvi, E Arenare, A R Van Meter, M De Choudhury, and J M Kane. 2019. Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *NPJ Schizophr* 5, 1 (Oct. 2019), 17.

[13] André Brock. 2018. Critical technocultural discourse analysis. *New Media & Society* 20, 3 (March 2018), 1012–1030.

[14] Stevie Chancellor. 2023. Toward Practices for Human-Centered Machine Learning. *Commun. ACM* 66, 3 (Feb. 2023), 78–85.

[15] Stevie Chancellor, Eric P S Baumer, and Munmun De Choudhury. 2019. Who is the "human" in human-centered Machine Learning. *Proc. ACM Hum. Comput. Interact.* 3, CSCW (Nov. 2019), 1–32.

[16] Stevie Chancellor, Michael L Birnbaum, Eric D Caine, Vincent M B Silenzio, and Munmun De Choudhury. 2019. A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 79–88.

[17] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ Digit Med* 3 (March 2020), 43.

[18] H F Cheng, L Stapleton, A Kawakami, and others. 2022. How child welfare workers reduce racial disparities in algorithmic decisions. *Proceedings of the* (2022).

[19] Austin Choi-Fitzpatrick, Douglas Irvin-Erickson, and Ernesto Verdeja. 2022. *Wicked Problems: The Ethics of Action for Peace, Rights, and Justice.* Oxford University Press.

[20] Maddy Coy, Liz Kelly, Fiona Vera-Gray, Maria Garner, and Ava Kanyeredzi. 2016. From 'no means no' to 'an enthusiastic yes': Changing the Discourse on Sexual Consent Through Sex and Relationships Education. In *Global Perspectives and Key Debates in Sex and Relationships Education: Addressing Issues of Gender, Sexuality, Plurality and Power.* Palgrave Macmillan UK, London, 84–99.

[21] Jay Cunningham, Gabrielle Benabdallah, Daniela Rosner, and Alex Taylor. 2023. On the Grounds of Solutionism: Ontologies of Blackness and HCI. *ACM Trans. Comput.-Hum. Interact.* 30, 2 (April 2023), 1–17.

[22] Bronwyn Davies. 1991. THE CONCEPT OF AGENCY: A Feminist Poststructuralist Analysis. *Social Analysis: The International Journal of Social and Cultural Practice* 30 (1991), 42–53.

[23] Mateusz Dolata, Stefan Feuerriegel, and Gerhard Schwabe. 2022. A sociotechnical view of algorithmic fairness. *Information Systems Journal* 32, 4 (July 2022), 754–818.

[24] Maeve Duggan. 2017. Online Harassment 2017. https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/. Accessed: 2023-2-23.

[25] Yrjö Engeström and Others. 1999. Activity theory and individual and social transformation. *Perspectives on activity theory* 19, 38 (1999), 19–30.

[26] Casey Fiesler, Cliff Lampe, and Amy S Bruckman. 2016. Reality and Perception of Copyright Terms of Service for Online Content Creation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1450–1461.

[27] Casey Fiesler and Nicholas Proferes. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4, 1 (Jan. 2018), 2056305118763366.

[28] Robert Folger, Russell Cropanzano, and Barry Goldman. 2013. What is the relationship between justice and morality? In *Handbook of organizational justice.* Psychology Press, 215–245.

[29] Rana Foroohar. 2018. Year in a Word: Techlash. *Financial Times* (Dec. 2018).

[30] Jaclyn Friedman and Jessica Valenti. 2019. *Yes Means Yes!: Visions of Female Sexual Power and a World without Rape.* Basic Books.

[31] Manas Gaur, Vamsi Aribandi, Amanuel Alambo, Ugur Kursuncu, Krishnaprasad Thirunarayan, Jonathan Beich, Jyotishman Pathak, and Amit Sheth. 2021. Characterization of time-variant and time-invariant assessment of suicidality on Reddit using C-SSRS. *PLoS One* 16, 5 (May 2021), e0250448.

[32] Stefan C Grant. 2021. Informed Consent-We Can and Should Do Better. *JAMA Netw Open* 4, 4 (April 2021), e2110848.

[33] Terry D Hargrave, Glen Jennings, and William Anderson. 1991. The development of a relational ethics scale. *J. Marital Fam. Ther.* 17, 2 (April 1991), 145–158.

[34] Gail E Henderson. 2011. Is informed consent broken? *Am. J. Med. Sci.* 342, 4 (Oct. 2011), 267–272.

[35] Joanne Hinds, Emma J Williams, and Adam N Joinson. 2020. "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *Int. J. Hum. Comput. Stud.* 143 (Nov. 2020), 102498.

[36] Xiaoyun Huang and Jessica Vitak. 2022. "Finsta gets all my bad pictures": Instagram Users' Self-Presentation Across Finsta and Rinsta Accounts. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1 (April 2022), 1–25.

[37] Terry Humphreys. 2007. Perceptions of sexual consent: the impact of relationship history and gender. *J. Sex Res.* 44, 4 (Nov. 2007), 307–315.

[38] Jane Im, Jill Dimond, Melody Berton, Una Lee, Katherine Mustelier, Mark S Ackerman, and Eric Gilbert. 2021. Yes: Affirmative Consent as a Theoretical Framework for Understanding and Imagining Social Platforms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 403)*. Association for Computing Machinery, New York, NY, USA, 1–18.

[39] Justinian I (Emperor of Byzantium). 1875. *Corpus juris civilis.* Baumgaertner.

[40] Atoosa Kasirzadeh. 2022. Algorithmic Fairness and Structural Injustice: Insights from Feminist Political Philosophy. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) *(AIES '22)*. Association for Computing Machinery, New York, NY, USA, 349–356.

[41] Jan Kietzmann, Jeannette Paschen, and Emily Treen. 2018. Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. *J. Advert. Res.* 58, 3 (2018), 263–267.

[42] Matthew Le Bui and Safiya Umoja Noble. 2020. We're Missing a Moral Framework of Justice in Artificial Intelligence: On the Limits, Failings, and Ethics of Fairness. (July 2020).

[43] Alexandra S Levine. 2022. Suicide Hotline Left Ethics Board Out Of The Loop About Data-Sharing With For-Profit Spinoff. *Forbes Magazine* (Feb. 2022).

[44] Alexandra S Levine. 2022. Suicide hotline shares data with for-profit spinoff, raising ethical questions. https://www.politico.com/news/2022/01/28/suicide-hotline-silicon-valley-privacy-debates-00002617. Accessed: 2023-3-13.

[45] Hanlin Li, Nicholas Vincent, Stevie Chancellor, and Brent Hecht. 2023. The Dimensions of Data Labor: A Road Map for Researchers, Activists, and Policymakers to Empower Data Producers. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1151–1161.

[46] Hanlin Li, Nicholas Vincent, Janice Tsai, Jofish Kaye, and Brent Hecht. 2019. How Do People Change Their Technology Use in Protest? Understanding. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 1–22.

[47] K Livingston. 2015. The queer art & rhetoric of consent: Theories, practices, pedagogies. *https://www.semanticscholar.org › paper › The-queer-art-...https://www.semanticscholar.org › paper › The-queer-art-...* (2015).

[48] Juniper L Lovato, Antoine Allard, Randall Harp, Jeremiah Onaolapo, and Laurent Hébert-Dufresne. 2022. Limits of Individual Consent and Models of Distributed Consent in Online Social Networks. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 2251–2262.

[49] Ewa Luger, Stuart Moran, and Tom Rodden. 2013. Consent for all: revealing the hidden complexity of terms and conditions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2687–2696.

[50] Catharine A MacKinnon. 1989. *Toward a Feminist Theory of the State.* Harvard University Press.

[51] Alice E Marwick and Danah Boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13, 1 (Feb. 2011), 114–133.

[52] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.* 54, 6 (July 2021), 1–35.

[53] Charlene L Muehlenhard, Terry P Humphreys, Kristen N Jozkowski, and Zoë D Peterson. 2016. The Complexities of Sexual Consent Among College Students: A Conceptual and Empirical Review. *J. Sex Res.* 53, 4-5 (April 2016), 457–487.

[54] Priyanka Nanayakkara, Jessica Hullman, and Nicholas Diakopoulos. 2021. Unpacking the Expressed Consequences of AI Research in Broader Impact Statements. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 795–806.

[55] Terrence Neumann, Maria De-Arteaga, and Sina Fazelpour. 2022. Justice in Misinformation Detection Systems: An Analysis of Algorithms, Stakeholders, and Potential Harms. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1504–1515.

[56] The AP News. 2023. Google settles $5 billion lawsuit over tracking people using 'incognito mode'. https://apnews.com/article/google-incognito-mode-tracking-lawsuit-settlement-8b30c9397f678bc4c546ab84191f7a9d. Accessed: 2024-1-17.

[57] Josef Nguyen and Bonnie Ruberg. 2020. Challenges of Designing Consent: Consent Mechanics in Video Games as Models for Interactive User Agency. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13.

[58] H Nissenbaum. 2004. Privacy as contextual integrity. *Wash Law Rev.* (2004).

[59] Office for Human Research Protections (OHRP). 2010. Informed Consent Checklist (1998). https://www.hhs.gov/ohrp/regulations-and-policy/guidance/checklists/index.html. Accessed: 2023-2-22.

[60] Javier Parapar, Patricia Martín-Rodilla, David E Losada, and Fabio Crestani. 2016. CLEF eRisk: Early Risk Prediction on the Internet. https://erisk.irlab.org/. Accessed: 2023-1-19.

[61] Sachin R Pendse, Amit Sharma, Aditya Vashistha, Munmun De Choudhury, and Neha Kumar. 2021. "Can I Not Be Suicidal on a Sunday?": Understanding Technology-Mediated Pathways to Mental Health Support. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 545)*. Association for Computing Machinery, New York, NY, USA, 1–16.

[62] Jay Peters and Jon Porter. 2023. More than 7,000 subreddits have gone dark to protest Reddit's API changes. https://www.theverge.com/2023/6/12/23755974/reddit-subreddits-going-dark-private-protest-api-changes. Accessed: 2024-1-22.

[63] Stuart Price. 1997. *The complete AZ media & communication handbook.* Hodder & Stoughton.

[64] Nicholas Proferes, Naiyan Jones, Sarah Gilbert, Casey Fiesler, and Michael Zimmer. 2021. Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media + Society* 7, 2 (April 2021), 20563051211019004.

[65] Carina E A Prunkl, Carolyn Ashurst, Markus Anderljung, Helena Webb, Jan Leike, and Allan Dafoe. 2021. Institutionalizing ethics in AI through broader impact requirements. *Nature Machine Intelligence* 3, 2 (Feb. 2021), 104–110.

[66] Organizers Of Queerinai and others. 2023. Queer In AI: A Case Study in Community-Led Participatory AI. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1882–1895.

[67] Rawls. 2004. A theory of justice. *Ethics* (2004).

[68] John Rawls. 2001. *Justice as Fairness: A Restatement*. Harvard University Press.

[69] Reddit. 2021. Reddit User Agreement. https://www.redditinc.com/policies/user-agreement-september-12-2021. Accessed: 2023-3-14.

[70] Martin Rehm, Frank Cornelissen, Alan J Daly, and Miguel Del Fresno García. 2020. Drinking from the Firehose: The Structural and Cognitive Dimensions of Sharing Information on Twitter. *Am. J. Educ.* 127, 1 (Nov. 2020), 77–105.

[71] Arie Rip and Jan-Peter Voβ. 2019. Umbrella terms as mediators in the governance of emerging science and technology. *Nanotechnology and Its Governance* (2019), 10–33.

[72] Horst W J Rittel and Melvin M Webber. 1973. Dilemmas in a general theory of planning. *Policy Sci.* 4, 2 (June 1973), 155–169.

[73] Jeffrey Ritter and Anna Mayer. 2017. Regulating data as property: a new construct for moving forward. *Duke Law Technol. Rev.* 16 (2017), 220.

[74] Emily F Rothman, Deinera Exner, and Allyson L Baughman. 2011. The prevalence of sexual assault against people who identify as gay, lesbian, or bisexual in the United States: a systematic review. *Trauma Violence Abuse* 12, 2 (April 2011), 55–66.

[75] Annalisa Sannino. 2015. The principle of double stimulation: A path to volitional action. *Learning, Culture and Social Interaction* 6 (Sept. 2015), 1–15.

[76] Bryan A Sisk, Jessica Mozersky, Alison L Antes, and James M DuBois. 2020. The "Ought-Is" Problem: An Implementation Science Framework for Translating Ethical Norms Into Practice. *Am. J. Bioeth.* 20, 4 (May 2020), 62–70.

[77] Linda J Skitka, Christopher W Bauman, and Elizabeth Mullen. 2008. Morality and justice: An expanded theoretical perspective and empirical review. In *Justice*. Vol. 25. Emerald Group Publishing Limited, 1–27.

[78] L Stapleton, M H Lee, D Qing, M Wright, and others. 2022. Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. *2022 ACM Conference* (2022).

[79] K Sun, A H Mhaidli, S Watel, C A Brooks, and others. 2019. It's my data! Tensions among stakeholders of a learning analytics dashboard. *Proceedings of the 2019* (2019).

[80] Harry Surden. 2019. Artificial Intelligence and Law: An Overview. (June 2019).

[81] Juan M Tauri. 2018. Research ethics, informed consent and the disempowerment of First Nation peoples. *Research Ethics* 14, 3 (July 2018), 1–14.

[82] Linnet Taylor. 2017. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4, 2 (Dec. 2017), 2053951717736335.

[83] United States. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Department of Health, Education, and Welfare, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.

[84] Hal R Varian. 1975. Distributive Justice, Welfare Economics, and the Theory of Fairness. *Philos. Public Aff.* 4, 3 (1975), 223–247.

[85] Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. 2021. Data Leverage: A Framework for Empowering the Public in its Relationship with Technology Companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 215–227.

[86] Jessica Vitak, Katie Shilton, and Zahra Ashktorab. 2016. Beyond the Belmont Principles: Ethical Challenges, Practices, and Beliefs in the Online Data Research Community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 941–953.

[87] Angelina Wang, Vikram V Ramaswamy, and Olga Russakovsky. 2022. Towards Intersectionality in Machine Learning: Including More Identities, Handling Underrepresentation, and Performing Evaluation. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 336–349.

[88] Alan Wertheimer. 1996. Consent and Sexual Relations. *Legal Theory* 2, 2 (June 1996), 89–112.

[89] Alissa Wilkinson. 2023. The looming threat of AI to Hollywood, and why it should matter to you. Accessed: 2023-7-26.

[90] Iris Marion Young. 1990. *Justice and the Politics of Difference*. Princeton University Press.

[91] Aleš Završnik. 2021. Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology* 18, 5 (Sept. 2021), 623–642.

[92] Qian Zhang, Jie Lu, and Yaochu Jin. 2021. Artificial intelligence in recommender systems. *Complex & Intelligent Systems* 7, 1 (Feb. 2021), 439–457.

[93] Michael Zimmer and Proferes Nicholas John. 2014. A topology of Twitter research: disciplines, methods, and ethics. *Aslib Journal of Information Management* 66, 3 (Jan. 2014), 250–261.