Analysis of Non-Discrimination Policies in the Sharing Economy

Author(s) Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
email address or ORCID

***Abstract*—Recent research has exposed a serious discrimination problem affecting applications of Sharing Economy (SE), such as Uber, Airbnb, and TaskRabbit. To control for this problem, several SE apps crafted a new form of usage policies, known as non-discrimination policies (NDPs). These policies are intended to raise awareness of the problem and describe the measures taken to mitigate any forms of discriminatory behavior affecting the app. However, there is still a major knowledge gap of how such policies should be drafted and structured. To bridge this gap, in this paper, we introduce a first-of-its-kind framework to support content analysis of NDPs in the SE market. Our analysis is conducted over a dataset of 108 SE apps, sampled from a broad range of application domains. Our results show that *a)* most SE apps do not provide an anti-discrimination policy *b)* the majority of policies are either extremely brief or provided as a part of other policies, and *c)* only a handful of apps provide a clear statement of how their anti-discrimination tactics are enforced through the functional features of the app. Our expectation is that this type of analysis can help SE app developers to draft more comprehensive and less ambiguous NDPs as well as help users make more informed decision in the SE market.**

***Index Terms*—Privacy, Sharing Economy, Requirements, NFR, Privacy Policies**

I. Introduction

Sharing (also known as *Shared* or *Gig*) Economy refers to a sustainable form of business exchange that is built around sharing, rather than owning, of resources [1]. Over the past decade, applications of the Sharing Economy, such as Uber, TaskRabbit, and Airbnb, have caused major disturbances in established classical markets, enabling people to exchange and monetize idle and underused assets at unprecedented scales [2]–[4]. As of today, there are thousands of SE platforms, operating in a market sector that is projected to grow to close to 335 billion U.S. dollars industry by 2025 [5].

The unique form of Peer-to-Peer (P2P) business exchange that SE platforms have enabled has been linked to significant levels of economic growth, especially in communities at the lower end of the economic ladder, helping unemployed and partially employed individuals to generate income, increase reciprocity, and access resources that are unattainable otherwise [1], [4], [6]–[8]. However, recent research has exposed a serious discrimination problem affecting these platforms [9]–[11]. Discrimination, as a general term, refers to cases where *"members of a minority group (women, Blacks, Muslims, immigrants, etc.) are treated differentially (less favorably) than members of a majority group with otherwise identical characteristics in similar circumstances"* [12]. In the context of SE, discrimination (also known as *digital* discrimination) refers to a phenomenon where **an online business transaction is influenced by race, gender, age, or any other non-business related characteristic of consumers or receivers**. This phenomenon is mainly facilitated by the P2P connection initiated between SE users, encouraging different forms of established bias (e.g., racism, sexism, and ableism) to transfer online [4], [9]–[11]. In traditional economy markets, discrimination is countered by imposing anti-discriminatory laws [13]. For instance, the U.S. Civil Rights Act of 1964 guarantees equal treatment of customers in public accommodations, such as hotels or rental property. However, in the cyberspace, discrimination takes a different form that is often difficult to detect and deter.

In response to discrimination concerns, SE apps started rolling out a new form of policies for addressing potential issues of discrimination. A policy, in general, serves as a legally binding contract between apps and their end-users [14]. Popular app stores demands apps to provide a privacy policy to specify the types of information they collect about their users and outline how that data is being used, protected, and shared [14]. Similarly, anti-discrimination policies determine the app's stance on discrimination and outline how acts of discrimination are identified and handled. Privacy policies have received singnificant attention in the Software Engineering literature [15], [16]. The objective of this line of research is to asses the quality of such policies as well as gauge best practices for drafting them. However, there is a widespread lack of knowledge about how NDPs can be structured. This can be attributed to the fact that polices are non-code artifacts. Creating and evolving such artifacts is often outside the expertise of developers.

To address this knowledge gap, in this paper, we examine NDPs in the SE market. Our analysis is conducted using a dataset of 108 apps sampled from a broad range of SE application domains. Our goal is to develop a standardized framework for systematically analyzing and evaluating the content of NDPs in such a complex and dynamic software ecosystem. The objectives of our framework are to **a)** help SE app developers draft more comprehensive and less ambiguous NDPs, and **b)** help end-users of SE apps to make more informed decisions in the SE market, either as service providers or receivers.

The remainder of this paper is organized as follows. Section II motivates our work and discusses our research questions. Section III describes our data collection process. Section V describes our NDP quality assessment framework. Section VI discusses our key findings and their impact as well as the main limitations of our study. Finally, Section VII concludes the paper and discusses future work.

II. Motivation and Research Questions

In this section, we provide background information on the problem of digital discrimination, motivate our research, and discuss our research questions.

A. Digital Discrimination

The problem of digital discrimination in online SE markets has been well-documented in recent years. Numerous large-scale surveys and field studies have provided significant evidence on different forms of systematic bias across almost all application domains of SE, including discrimination based on ethnic background (racism), gender or sexual orientation (sexism), and physical appearance (ableism) [9], [10], [17]–[20]. For instance, Ge et al. [9] hired research assistants of different racial backgrounds to request UberX rides. The authors found that the waiting times for Black riders were significantly longer. In addition, more cancellations were observed against Black riders than their White counterparts. In another study, Moody et al. [17] surveyed 1,100 of UberPOOL and Lyft riders. The results showed that White passengers that lived in predominantly White communities were more likely to discriminate against passengers of other races.

Edelman et al. [10] examined racial discrimination over the lodging platform Airbnb. The authors reported that applications from guests with distinctively Black names were 16% less likely to be accepted relative to identical guests with distinctively White names. Discrimination in the lodging business has also been observed against members of the LGBT community. For example, Ahuja and Lyons [18] analyzed Airbnb host responses to LGBT accounts. The results showed that hosts were more likely to not reply at all rather than replying *"no"* to male-male pairs inquiring about room availability. Ableism (discrimination against people with disability) was also reported over Airbnb. For instance, in a randomized field experiment of 3,847 lodging requests, Ameri et al. [21] found that hosts were less likely to approve requests from travelers with blindness, cerebral palsy, dwarfism, or spinal cord injury than to approve travelers without disabilities.

Patters of digital discriminatory have also been observed in the freelancing domain. Thebault et al. [19] surveyed workers on TaskRabbit from the Chicago metropolitan area. The authors found that requests from customers in the socioeconomically disadvantaged South Side area were less likely to be accepted. Hannák et al. [11] analyzed worker profiles on TaskRabbit and Upwork. The results showed that there was a significant bias against White women and Black men on both platforms. In another study, Foong et al. [22] collected self-determined hourly bill rates from the public profiles of 48,019 workers in the U.S. (48.8% women) on Upwork. The authors found that the median woman on Upwork requested only 74% of what the median man requested in hourly bill rate. Another study by Barzilay and Ben-David [20] showed that women's average hourly rates on P2P freelancing platforms were about two-thirds of men's rates. Such gaps persisted even after controlling for experience, educational background, and hours of work.

B. Motivation and Research Questions

Policies have long been used by apps as a means to communicate their data collection and sharing practices with their users as well as comply with privacy legislation around the world. These policies, commonly known as privacy policies, have generated considerable attention in recent years [23]. Privacy policy research is mainly focused on detecting violations of the privacy claims in the policy [24], [25], evaluating their readability and comprehensibility [24], [25], and mining their content for software privacy requirements [15], [16], [26]. In general, the research in this domain has revealed that a large percentage of apps were either non-compliant with their stated privacy claims, did not provide a policy, or provided a policy that was ambiguous or incomprehensible to users [23], [27].

In this paper, we take a first-of-its-kind step to analyze NDPs in Sharing Economy. Our work aims to **a)** study the prevalence of NDPs in SE, **b)** propose a framework for systematically analyzing the content of these policies, and **c)** use that framework to assess the quality of existing NDPs of popular SE apps. Our work is intended to spread awareness of digital discrimination and provide app developers, either maintaining SE apps or developing new ones, with systematic guidelines to draft high quality NDPs and evolve such non-code artifacts with minimum overhead. Providing complete and structured NDPs can help SE app users to make more optimized socioeconomic decisions when it comes to navigating the landscape of existing SE solutions. To guide our analysis, we formulate the following research questions:

- ***RQ$_1$:*** *How prevalent are anti-discrimination policies in SE?* Under this research question, we investigate the percentage of anti-discrimination policies in SE. This type of analysis can give an indication of app developers' awareness of the problem of discrimination.
- ***RQ$_2$:*** *Can the quality of existing NDPs be systematically evaluated?* Under this research question, we seek to develop a systematic framework to support the analysis of NDPs content as well as assess their quality.
- ***RQ$_2$:*** *How detailed and informative are existing NDPs in SE?* Under this research question, we examine the quality of information provided in existing SE apps' NDPs. Our objective is to determine a set of quality standards that can be used by other apps to draft their own NDPs.

III. Data Collection

In this section, we describe our data collection process, including selecting apps to be included in our dataset, categorizing these apps, and collecting their NDPs.

IV. DATASET

Our goal is to analyze NDPs across a broad range of SE categories. Therefore, the first step of our analysis is to generate a dataset of SE apps to be investigated in our analysis. Recent statistics estimate that there are thousands of active SE platforms on popular app stores. However, only a handful of these apps are studied in digital discrimination research, including Uber and Lyft from the domain of ride-sharing, Airbnb from the domain accommodation, and UpWork and TaskRabbit from the domain of freelancing. These apps operate in large geographical areas and have massive user bases, thus, discrimination concerns are more likely to manifest over such popular platforms rather than smaller ones. Based on these observations, to for an SE app to be included in our analysis, it has to meet the following criteria:

1) A platform must facilitate some sort of a P2P connection and include the sharing of some sort of a resource, whether an asset (e.g., apartment, car, boat, or a drill, etc.) or a skill (e.g., plumbing, styling hair, coding, etc.).
2) A platform must have an app on Google Play or App Store. Mobile marketplaces provide a convenient way to select popular platforms with larger user bases using the public popularity metrics they provide, such as the number of reviews, stars, and their top charts.
3) A platform must be located and/or have a substantial presence in the US. The U.S. Civil Rights Act of 1964 prohibits discrimination based on race, sex, religion, nationality, or sexual orientation. By focusing on the US market, we ensure that apps operate in a country where discrimination is prohibited by law.

With these criteria in place, we searched for apps to be included in our dataset. Our data collection took place between January and February of 2021. We started by seeding our dataset with of six popular SE apps: Uber, Lyft, Airbnb, Upwork, TaskRabbit, and Fiverr. Existing literature has provided significant evidence of digital discrimination in these apps. We then conducted a Google search using the query: `(sharing OR shared OR gig) AND economy AND (platforms OR app OR system)`. We examined the first 15 pages of the search results and added 72 new platforms that matched our inclusion criteria. We then used the *similar* feature on Google Play and Apple App Store to locate any apps we missed through the Google search. Specifically, we examined the list of similar app resulting from searching app stores for each of our 72 apps. Lightweight snowballing was then used to add any major apps that we potentially missed. Apps were iteratively added until no more new apps that satisfy our inclusion criteria were located. In total, 108 unique apps are included in our dataset. Descriptive statistics of our dataset of apps are provided in Fig **??**.

A. App categorization

App stores classify apps into generic categories of loosely related functionalities. These categories are often left ambiguous (too generic) or straight-up wrong [28], [29]. For example, both Uber and Airbnb are categorized under the *Travel* category in the Apple App Store and DoorDash is classified under the *Food&Drink* category. This type of generic categorization does not provide enough information about the specific application domains of apps. To overcome this limitation, we re-classify apps in our dataset into more fine-grained categories of SE application domains.

While automated app classification techniques are available [28], [29], given the relatively small size of our dataset, we conducted the classification manually. Specifically, three judges (authors) independently examined the description of each of our app available on both app stores as well as the app's official Web-page. Categories of apps were recorded as they emerged in the text. We used memoing to keep track of the reasoning behind each suggested category. Axial coding was then used to consolidate individual categorizations into more abstract categories. For example, the categories of *food delivery* and *grocery delivery* were merged into a single *food delivery* category and *boat-sharing* and *bike-sharing* into *vehicle-sharing*. Generated categories were then iteratively revised until no more categories were found. By the end of our classification process, six main categories of SE apps, shown in Fig. 1 have emerged. In general, these categories can be described as follows:

- **Skill-based:** These apps facilitate the sharing of personal skills (hiring labor). Specific examples include baby sitting (Sittercity and Urbansitter), tutoring (Verbling, Codementor, and Classgap), and freelancing platforms (Fiverr and Upwork).
- **Delivery:** Under this category, we include apps which enable users to utilize their vehicles for delivery to other users. Examples of apps in this category include grocery and food delivery (UberEats, Grubhub, and Shipt) and other delivery tasks (DriveMatch, uShip, and Dolly).
- **Ride-sharing:** This category include apps which allow their users to share rides, this include carpooling and driver/rider connections. Examples of apps in this category include traditional ride-sharing services, such Uber, Lyft, and Via, as well as more specialized platforms, such as HopSkipDriver for children transportation, Veyo for medical transportation, and Wingz for hiring a driver.
- **Vehicle-sharing:** Under this category, we include any app which enables users to lend their means of transportation. This category is different from the above category in the sense that the resource being shared in the vehicle not the driver. Examples of apps under this category include car sharing platforms (Turo and HyreCar), boat sharing (Get-MyBoat and Boatsetter), bike sharing (Spinlister), and RV sharing (RVezy and Outdoorsy).
- **Logding:** This category includes renting and short-term accommodation services such as Airbnb, Vrbo, and Misterbnb as well as space-sharing for storage (Neighbor), events and work (Splacer and LiquidSpace), and even parking (ParqEx).
- **Other:** Although our objective was to classify all apps

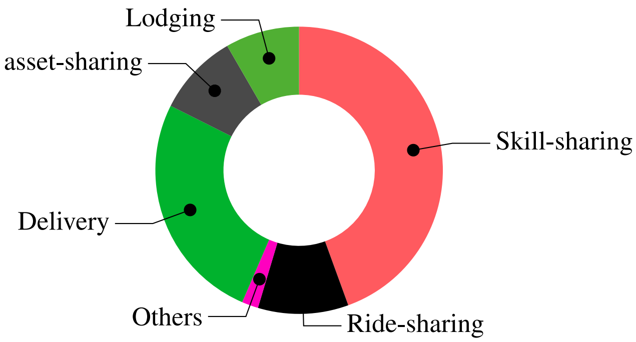

Fig. 1: Distribution of our apps over application domains.

TABLE I: Most common types of discrimination.

Type	Discrimination against:
Racism	Ethnicity, color, or nationality.
Sexism	Gender or sexual orientation.
Ableism	Physical, sensory, or intellectual disability.
Parental	Parents with children or pregnant women.
Ageism	Older or younger people.
Religious	Perceived religion or a set of beliefs.
Classism	Particular social class.

into the main general categories, two apps in our dataset were too niche-oriented to warrant a creation of a separate category. These apps are Prosper for lending and borrowing money and Kickstarter, a platform for crowdfunding various projects.

B. Policy collection

To answer our first research question, we collect the NDPs for our set of apps. Unlike privacy policies, app stores do not enforce NDPs, therefore, locating such policies can be a challenging task. For instance, most privacy policies are often titled *Privacy Policy*, however, NDPs are titled differently, such as *Non-discrimination*, *Anti-discrimination*, or *Inclusion Statement*. To locate such policies, we explore the website of each app as well as the mobile app itself. Any statement that addresses discrimination or any of its types is collected as a potential NDP. Table I lists the main acts of discrimination as described by the U.S. Equal Employment Opportunity Commission. These acts commonly appear in diversity and social justice literature [**?**].

To identify these policies, we utilized Google's search operators to search apps' websites directly using the query `site: <app website> AND (discrimination OR <discrimination types (Table I)>)`. For the apps we could not locate a policy, we manually searched the website for such a policy. In general, our search has exposed three categories of apps in our dataset when it comes to NDPs. These categories include:

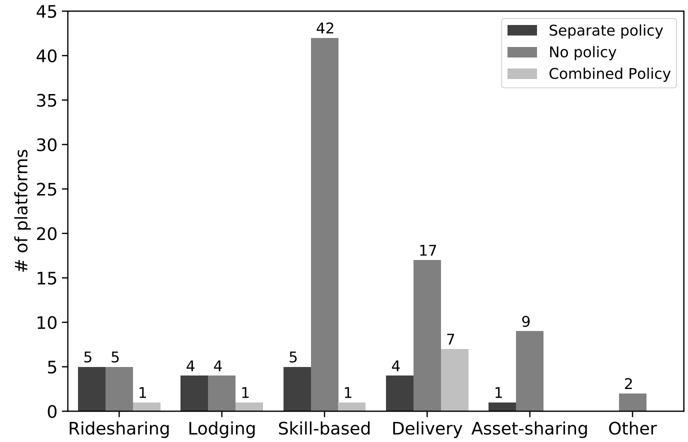

Fig. 2: GE app categories determined by our classification, by presence

- **Separate policy:** This category include apps which have a separate NDP provided on a separate page. In total 19 apps had a separate NDP policy.
- **Combined policy:** In 10 apps, NDPs were combined with another policy. Examples include combining a NDP with a sexual harassment policy, a zero-tolerance policy, community guidelines, code of conduct, or Terms of Service.
- **No policy:** In majority of our apps (79), we were not able to locate a policy beyond a standard statement in the ToS agreement with generic language in relation to discrimination. For example.

A distribution of these three categories of NDPs over our categories of SE application domains in shown in Fig. 2. In general, to answer $\boldsymbol{RQ_1}$, we can safely say that the majority of apps included in our dataset does not provide NDPs. Even the ones that do, often include their NDPs with another policy and not as a separate policy.

V. Quality Assessment of NDPs

Quality analysis of software policies aims at providing a systematic, comprehensive, and structured assessments of policy content [30], [31]. This type of analysis is is commonly applied to privacy policies of mobile apps and Websites to provide guidelines for evaluating exiting policies and aid in the creation of new ones [30], [32]–[34]. To answer $\boldsymbol{RQ_2}$, in this section, we propose a protocol for assessing the quality of NDPs and apply this protocol to the set of NDPs in our dataset.

A. A framework for NDP assessment

The process of policy assessment is typically conducted manually, following a systematic process that checks the content of the policy against a set of quality measures. These measures range from simple quantitative metrcis, such as length [35], to more complex qualitative measures, such as readability or compliance with existing regulations [36]. After such measures are defined, a group of policy evaluators are asked to assess the quality of the policy using these measures. Typically, evaluators are provided with a set of questions to help them throughout the process [36], [37]. Following these guidelines, we compile a set of measures to assess NDPs in our dataset. These measures, adapted from existing privacy policy evaluation protocols, along with their associated set of questions, are provided in Table II.

B. Results and Analysis

Once our review protocol was defined, we printed out the NDPs collected for our apps. Each of the three authors went through each NDP independently, answering questions related to criteria from (3-7) in Table II. Results were then compiled and summarized in Table III. Overall, given the specific nature of our questions, only few coding errors (inaccuracies in answering some of the questions) were detected and corrected. In what follows, we discuss our findings in greater detail.

*1) **Length and Readability**:* The Flesch Reading Ease (FRE) [38] is a popular metric used to assess readability of a text. FRE ranges from 0 to 100, where higher score indicates texts that are easier to read. The metric is calculated by the following formula:

$$206.835 - 84.6 \times \frac{\text{\# of syllables}}{\text{\# of words}} - 1.015 \times \frac{\text{\# of words}}{\text{\# of sentences}} \quad (1)$$

The core idea behind FRE is that longer words and longer sentences are more difficult to comprehend. Therefore, FRE penalizes texts with a high number of syllables per word and a high number of words per sentence. FRE metric is commonly used in policy readability research [35], [39], [40]. It is important to note that the metric is only suitable for longer texts due to high variability of shorter texts. Therefore, we calculated FRE only for NDPs with 100 words and more.

*2) **Types**:* Our annotation shows that most polices list a large number of discrimination types in their NDPs. TaskRabbit, in particular, refers to 17 different types of discrimination including racism, color, ancestry, national origin, religion, creed, age, sex, gender, physical or mental disability, medical condition, genetic information, marital or civil partner status, military or veteran status. The policy even provide more sub-types of discrimination, such as *gender (including pregnancy, childbirth, breastfeeding or related medical conditions)*.

*3) **Examples**:* Our manual annotations shows that examples are not as common in NDPs. Airbnb, Turo, and Neighbor providing the most set of example. For instance, Airbnb policy states that *"Airbnb hosts may not decline a booking from a guest based on gender identity unless the host shares living spaces (for example, bathroom, kitchen, or common areas) with the guest"*. Neighbor one of the best examples in terms of quantity and quality. The app provide specific examples of what is considered discriminatory behavior, such as *"Posts that assume someone is suspicious because of their race or ethnicity"*. Turo is another app which provides through examples of discriminatory behavior that might affect the service. For example, *"Turo hosts may not make assumptions about the guest's ability to operate their vehicle"* and *"Turo hosts may not decline, cancel, or impose any different terms or conditions for a reservation based on familial status"*.

*4) **Legal**:* Providing references to specific counter-discrimination legislation. GoShare's NDP, for example, states that *"A variety of federal, state, and local laws strictly prohibit such forms of discrimination, including Title VII of the Civil Rights Act of 1964, the Age Discrimination Act of 1967, and the Americans with Disabilities Act of 1990."*. Some other apps provide a generic legal statement. For example, TaskRabbit states that *"or any other basis protected by applicable laws in jurisdictions in which TaskRabbit operates (collectively referred to as a protected class)"*.

*5) **Enforcement and Ramifications**:* In response to discrimination claims, adapted minor feature changes to mitigate, or counter, different forms of systematic bias on their platforms. For instance, in order to control for discrimination based on profile pictures or names, apps such as Airbnb and Uber delay the exposure to users' information till after the transaction is confirmed [**?**]. To control for bias in reviews, Airbnb rolled out a design change to ensure that hosts and guests can see the reviews only after both parties have submitted their reviews. According to Airbnb, *"Both hosts and guests may worry that if they leave an honest review that includes praise and criticism, they might receive an unfairly critical review in response. To address this concern, reviews will be revealed to hosts and guests simultaneously."* [**?**].

VI. Discussion and Impact

Given the general shift in society towards more equality and prosperity, we anticipate that NDPs are going to become mandated by law in the near future. However, in the absence of a standardized format and the lack of regulations, drafting such polices remains a challenging and time-consuming task during software maintenance. To help overcome these challenges, the framework presented in this paper provides developers with a systematic protocol for evaluating their policies based on their intrinsic characteristics and by comparing them to existing high quality NDPs. This framework can also help developers to keep NSPs in-check during code evolution. This can be particularly important for start-ups, where it is often financially infeasible to hire a third-party (e.g., a law firm) to take care of the policy as the system evolve and as we learn more about the problem of digital discrimination.

Our work in this paper bridges an important gap in software maintenance and evolution research by focusing on the evolution of non-code artefacts. Maintaining software polices is a good example of adaptive maintenance tasks, where an artifact has to constantly change to adapt to external factors, such as

TABLE II: Our measures for NDP content assessment.

No	Measure	Description	Questions for evaluators
1	Length (L)	The length of the policy (number of words) can be used as a basic measure of its quality. Intuitively, longer policies are assumed to be more detailed [35]	Measured automatically.
2	Readability (FRE.)	Readability is another meausre that is commonly used to assess quality [35], [39], [40]. The more readable the policy, the more accessible it is for the casual user.	Calculated automatically using the Flesch Reading Ease (FRE) [38]
3	Types (T)	Discrimination in SE apps can take many forms (TableI). Therefore, a policy that excitability mentions more of these types is considered higher in quality, or more comprehensive.	How many specific types of discrimination does the policy mention? How many types does the policy define?
4	Examples (Ex.)	A policy that provides examples of specific types of discriminatory behavior that might affect the app is considered to be higher in quality. Examples are used as to demonstrate what actions might be classified as discriminatory. In policy analysis, examples are considered an important instrument to communicate policy practices with the casual user [32].	Does the policy provide any examples of discriminatory behavior? How many examples are provided?
5	Legislation (Lg.)	This criterion assess whether a policy contains references exiting anti-discrimination regulations in the judicial area the app operates. Internet privacy policies are often assessed based on their compliance with existing privacy regulations [31], such as the Federal Trade Commission's Fair Information Practices guidelines [32].	Does the policy refer to any exiting legislation?
6	Enforcement (En.)	A policy which lists the set of measures (functional or non-functional) that are taken by the app to mitigate discrimination is considered higher in quality. Enforcement mechanisms are commonly used in the content analysis of privacy policies [32]. These types of mechanisms also include methods for reporting incidents of violation	Does the policy list any features that the app implements to mitigate discrimination? Is there a reporting mechanism in place?
7	Ramifications (Rmf.)	A policy which mentions the ramifications for violating the standards is considered more comprehensive [].	Does the policy mention the types of actions (penalties) to be imposed on policy violators?

changing regulations. Furthermore, such polices can be used to monitor the evolution of the system by monitoring changes to the NDP policy. Existing research suggests that important information about the technical implementation from modifications to privacy policies [36]. Users often find themselves having to choose from among hundreds of SE solutions in one the fastest growing markets in the world []. Making the right decisions is critical to maximize the personal, social, and economic gains in the SE market. To that extent, informative, comprehensive, and accessible NDPs can serve as an important input to help them make an informed-decision.

To summarize our findings, we revisit our research questions:.

- ***RQ$_1$:*** *How prevalent are anti-discrimination policies in SE?* Our analysis of 109 SE apps show that NDPs are not common in SE. Most apps either do not provide a NDP at all or provide a very brief and generic statement. Only few apps maintain a separate NDP. Such polices appear under different names which might negatively impact their discoverability.
- ***RQ$_2$:*** *Can the quality of existing NDPs be systematically evaluated?* Existing frameworks and protocols for evaluating the content of software privacy policies can be adapted to NDPs. Specifically, NDPs can be evaluated based on a set of measures that can be extracted directly from the policy. These measures include quantitative metrics, such as length, readability, number of examples, and types of discrimination acknowledged in the policy as well as qualitative measures, such as whether the policy describe any measures taken to mitigate discrimination and how cases of violation can be reported and handled.
- ***RQ$_2$:*** *How detailed and informative are existing NDPs in SE?* Our analysis shows that the majority of the proposed polices are

Finally, in terms of limitations, a main threat to the external validity of our study might stem from the fact that only 109 popular SE platforms were considered in our analysis. However, as mentioned earlier, discrimination issues are more likely to manifest over these platforms rather than smaller platforms which typically target homogeneous populations of users. Furthermore, our search process utilized multiple search strategies as well as inclusion criteria to locate a representative sample. Generally speaking, the size of the dataset is aligned with datasets typically used in policy analysis research [35]–[37]. Another threat might stem from the fact that our policy evaluation protocol was conducted manually. Nonetheless,

TABLE III: NDP content assessment results. The Flesch Reading Ease (FRE) metric is calculated for NDPs with L $\geq$ 100

Domain	App	NDP type	L	FRE	T	Ex	Lg	En	Rmf
Ridesharing	Uber	Separate	134	18.78	9	1	✗	✗	✓
	Lyft	Separate	97	-	10	0	✗	✓	✓
	Via	Separate	102	30.77	8	0	✗	✓	✓
	HopSkipDrive	Combined	37	-	11	0	✗	✗	✗
	Veyo	Separate	189	8.30	11	0	✓	✓	✗
	Wingz	Separate	88	-	9	1	✗	✗	✓
Asset-sharing	Turo	Separate	770	23.78	13	15	✗	✓	✓
Delivery	Doordash	Combined	64	-	14	0	✗	✗	✗
	Grubhub	Combined	60	-	16	0	✗	✗	✗
	uShip	Combined	153	35.17	9	0	✗	✗	✗
	GoShare	Combined	1424	13.00	11	0	✓	✓	✓
	Postmates	Separate	72	-	11	1	✗	✗	✓
	Roadie	Combined	480	12.10	13	0	✗	✓	✓
	Instacart	Combined	35	-	11	0	✗	✗	✗
Lodging	Airbnb	Separate	2012	34.70	8	22	✗	✓	✓
	Misterb&b	Separate	420	52.40	5	5	✗	✓	✓
	Vrbo	Separate	91	-	0	0	✗	✗	✓
	Neighbor	Combined	572	30.90	8	15	✗	✗	✓
	Spareroom	Separate	399	44.33	14	5	✓	✗	✗
Skill-based	Taskrabbit	Separate	403	0	17	4	✗	✓	✓
	Upwork	Separate	308	21.90	10	7	✗	✗	✗
	Thumbtack	Separate	152	18.60	11	3	✗	✓	✓
	Jobstack	Separate	412	0	21	0	✗	✓	✓
	Sittercity	Separate	132	40.93	0	0	✗	✗	✗
	Withlocals	Combined	83	-	0	0	✗	✗	✗

manual inspection of policy the common practice in these kind of studies. This threat can be mitigated by using a systematic review process and a well-defined review protocol.

VII. Conclusions

Our work in this paper will be extended across two main directions:

- Automation: Another important step of our analysis is to automated the policy analysis process. This step will include using text mining and modeling techniques to automatically learn the structure of NDPs, the main topics they discuss, and eventfully generate an over all quality score. This analysis will require a large number of policies to provide enough information for text mining algorithms to make accurate predictions. A fully automated prototype will be made publicly available to help app developers around the world draft high quality NDPs.
- User studies: Automated metrics such as readability can provide an indication of NDPs' accessible. However, to get a more objective assessment, user studies must be conducted. Such studies will involve recruiting large samples of SE users (providers and receivers) and deigning questionnaires to assess their level of understanding of NDPs.

Acknowledgment

A place holder for acknowledging the funding agency.

References

[1] C. Martin, "The sharing economy: A pathway to sustainability or a nightmarish form of neoliberal capitalism?" *Ecological Economics*, vol. 121, pp. 149–159, 2016.

[2] T. Dogru, M. Mody, and C. Suess, "Adding evidence to the debate: Quantifying Airbnb's disruptive impact on ten key hotel markets," *Tourism Management*, vol. 72, pp. 27–39, 2019.

[3] G. Quattrone, D. Proserpio, D. Quercia, L. Capra, and M. Musolesi, "Who benefits from the "sharing" economy of airbnb?" in *International Conference on World Wide Web*, 2016, pp. 1385–1394.

[4] T. Dillahunt and A. Malone, "The promise of the sharing economy among disadvantaged communities," in *Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2285–2294.

[5] PwC, "The sharing economy: Consumer intelligence series," *PricewaterhouseCoopers LLP*, 2015.

[6] J. Hamari, M. Sjöklint, and A. Ukkonen, "The sharing economy: Why people participate in collaborative consumption," *Journal of the Association for Information Science and Technology*, vol. 67, no. 9, pp. 50–63, 2016.

[7] G. Zhu, K. So, and S. Hudson, "Inside the sharing economy: Understanding consumer motivations behind the adoption of mobile applications," *International Journal of Contemporary Hospitality Management*, vol. 29, no. 9, pp. 2218–2239, 2017.

[8] S. Shaheen, N. Chan, A. Bansal, and A. Cohen, "Mobility and sharing economy: Impacts synopsis," 2015.

[9] Y. Ge, C. Knittel, D. MacKenzie, and S. Zoepf, "Racial and gender discrimination in transportation network companies," no. NBER Working Paper No. 22776, 2017.

[10] B. Edelman, M. Luca, and D. Svirsky, "Racial discrimination in the sharing economy: Evidence from a field experiment," *American Economic Journal: Applied Economics*, vol. 9, no. 2, pp. 1–22, 2017.

[11] A. Hannák, C. Wagner, D. Garcia, A. Mislove, M. Strohmaier, and C. Wilson, "Bias in online freelance marketplaces: Evidence from TaskRabbit and Fiverr," in *ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 1914–1933.

[12] M. Bertrand and E. Duflo, *Handbook of Field Experiments*. Elsevier, 2016, ch. Field Experiments on Discrimination, p. 310.

[13] L. Ratti and N. Countouris, "The sharing economy and eu anti-discrimination law," *The Cambridge Handbook of the Law of the Sharing Economy*, pp. 486–498, 2018.

[14] X. Wang, X. Qin, M. Bokaei, R. Slavin, T. Breaux, and J. Niu, "Guileak: Tracing privacy policy claims on user input data for Android applications," in *Inter. Conf. on Software Engineering*, 2018, pp. 37–47.

[15] S. Blenner, M. Köllmer, A. Rouse, N. Daneshvar, C. Williams, and L. Andrews, "Privacy policies of android diabetes apps and sharing of health information," *The Journal of the American Medical Association*, vol. 315, no. 10, pp. 1051–1052, 2016.

[16] J. Bhatia, T. Breaux, and F. Schaub, "Mining privacy goals from privacy policies using hybridized task recomposition," *ACM Transactions on Software Engineering and Methodology*, vol. 25, no. 3, p. 22, 2016.

[17] J. Moody, S. Middleton, and J. Zhao, "Rider-to-rider discriminatory attitudes and ridesharing behavior," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 62, pp. 258–273, 2019.

[18] R. Ahuja and R. Lyons, "The silent treatment: LGBT discrimination in the sharing economy," Trinity College Dublin, Department of Economics, Dublin, Tech. Rep., 2017.

[19] J. Thebault-Spieker, L. Terveen, and B. Hecht, "Avoiding the south side and the suburbs: The geography of mobile crowdsourcing markets," in *ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2015, pp. 265–275.

[20] A. Barzilay and A. Ben-David, "Platform inequality: Gender in the gig-economy," *Seton Hall Law Review*, vol. 47, no. 2, pp. 393–431, 2017.

[21] M. Ameri, S. Rogers, L. Schur, and D. Kruse, "No room at the inn? disability access in the new sharing economy," *Academy of Management Discoveries*, vol. 6, no. 2, pp. 176–205, 2020.

[22] E. Foong, N. Vincent, B. Hecht, and E. Gerber, "Women (still) ask for less: Gender differences in hourly rate in an online labor marketplace," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 53:1–53:21, 2018.

[23] F. Ebrahimi, M. Tushev, and A. Mahmoud, "Mobile app privacy in software engineering research: A systematic mapping study," *American Economic Journal: Applied Economics*, vol. 133, no. 2, p. 106466, 2021.

[24] L. Yu, X. Luo, C. Qian, S. Wang, and H. Leung, "Enhancing the description-to-behavior fidelity in Android apps with privacy policy," *IEEE Transactions on Software Engineering*, vol. 44, no. 9, pp. 834–854, 2018.

[25] A. Aydin, D. Piorkowski, O. Tripp, P. Ferrara, and M. Pistoia, "Visual configuration of mobile privacy policies," in *Inter. Conf. on Fundamental Approaches to Software Engineering*, 2017, pp. 338–355.

[26] J. Young, "Commitment analysis to operationalize software requirements from privacy policies," *Requirements Engineering*, vol. 16, no. 1, pp. 33–46, 2011.

[27] S. Zimmeck, Z. Wang, L. Zou, R. Iyengar, B. Liu, F. Shaub, S. Wilson, N. Sadeh, S. Bellovin, and J. Reidenberg, "Automated analysis of privacy requirements for mobile apps," in *AAAI Fall Symposium Series*, 2016, pp. 286–296.

[28] G. Berardi, A. Esuli, T. Fagni, and F. Sebastiani, "Multi-store metadata-based supervised mobile app classification," in *Annual ACM Symposium on Applied Computing*, 2015, pp. 585–588.

[29] A. Al-Subaihin, F. Sarro, S. Black, L. Capra, M. Harman, Y. Jia, and Y. Zhang, "Clustering mobile apps based on mined textual features," in *International Symposium on Empirical Software Engineering and Measurement*, 2016, pp. 1–38.

[30] T. Dehling, F. Gao, and A. Sunyaev, "Assessment instrument for privacy policy content: Design and evaluation of ppc," in *the Pre-ICIS Workshop on Information Security and Privacy*, 2014.

[31] R. Ryker, E. Lafleur, B. McManis, and K. C. Cox, "Online privacy policies: An assessment of the fortune e-50," *Journal of Computer Information Systems*, vol. 42, no. 4, p. 2002, 2002.

[32] D. A. Bradbard, C. Peters, and Y. Caneva, "Web accessibility policies at land-grant universities," *The Internet and Higher Education*, vol. 13, no. 4, pp. 258–266, 2010.

[33] P. Savla and L. Martino, "Content analysis of privacy policies for health social networks," in *IEEE International Symposium on Policies for Distributed Systems and Networks*, 2012, pp. 94–101.

[34] S. Rains and L. Bosch, "Privacy and health in the information age: A content analysis of health web site privacy policy statements," *Health Communication*, vol. 24, no. 5, pp. 435–446, 2009.

[35] C. Jensen and C. Potts, "Privacy policies as decision-making tools: An evaluation of online privacy notices," in *the SIGCHI Conference on Human Factors in Computing Systems*, 2004, p. 471–478.

[36] B. Miller, K. Buck, and J. Tygar, "Systematic analysis and evaluation of web privacy policies and implementations," in *International Conference for Internet Technology and Secure Transactions*, 2012, pp. 534–540.

[37] J. Zimmerle and A. Wall, "What's in a policy? evaluating the privacy policies of children's apps and websites," *Computers in the Schools*, vol. 36, no. 1, pp. 38–47, 2019.

[38] R. Flesch and A. Gould, *The art of readable writing*, 1949, vol. 8.

[39] A. McDonald, R. Reeder, P. Kelley, and L. Cranor, "A comparative study of online privacy policies and formats," in *International Symposium on Privacy Enhancing Technologies Symposium*, 2009, pp. 37–55.

[40] A. Powell, P. Singh, and J. Torous, "The complexity of mental health app privacy policies: A potential barrier to privacy," *JMIR mHealth and uHealth*, vol. 6, no. 7, 2018.