Digital Discrimination in Sharing Economy A Requirements Engineering Perspective

Miroslav Tushev, Fahimeh Ebrahimi, and Anas Mahmoud
The Division of Computer Science and Engineering
Louisiana State University
mtushe1@lsu.edu, febrah1@lsu.edu, amahmo4@lsu.edu

***Abstract*—Recent evidence has revealed that Sharing Economy platforms such as Uber, Airbnb, and TaskRabbit, have become active hubs for digital discrimination. This new form of discrimination refers to a phenomenon where a business transaction is influenced by race, gender, age, or any other non-business related characteristic of providers or consumers. Existing research often tackles this problem from a socio-economic and regulatory points of view. However, the research on the design aspects of Sharing Economy software, which enable such complex socio-technical problems to emerge online, is still underdeveloped. To bridge this gap, in this paper, we propose a new perspective on digital discrimination, tackling the problem from a Requirements Engineering point of view. Specifically, we analyze a large dataset of online user feedback as well as synthesize existing literature to identify and classify pervasive discrimination concerns in the Sharing Economy market. Based on this analysis, we devise a crowd-driven domain model to represent these concerns along with their relations to the functional features and user goals of Sharing Economy platforms. This model is intended to provide requirements engineers, working on Sharing Economy software, with systematic insights into the complex types of socio-technical problems that can emerge in the operational environments of their systems.**

I. Introduction

Earlier adaptations of Sharing Economy can be traced back to the early days of the Internet. eBay and Craigslist, both launched in 1995, are prominent examples of platforms that facilitated a collaborative P2P circulation of services and assets. However, the real proliferation of Sharing Economy can be attributed to mobile technology. Using mobile apps as a mediator, services such as Uber, Airbnb, and TaskRabbit were able to penetrate the mainstream culture, enabling consumers to sell, rent, swap, lend, and borrow services and assets at unprecedented scales. According to PwC—the multinational professional services network—Sharing Economy is projected to grow from 15 billion U.S. dollars in 2014 to close to 335 billion U.S. dollars by 2025 [1].

Sharing economy presents opportunities for unemployed, or partially employed, individuals to find employment, generate extra income, increase reciprocity, and access resources that are unattainable otherwise [2]. Despite these proven benefits, recent research has revealed that Sharing Economy solutions tend to be significantly less effective between minorities [3], [4], [5]. Specifically, minority individuals engaging in Sharing Economy are often subject to different forms of discrimination that negatively influence the outcome of their business transactions. For instance, a recent study of ridesharing services found that black riders using Uber waited 30% longer to be picked up [3]. Another study on P2P lodging reported that non-black Airbnb hosts were able to charge 12% more than black hosts [4].

The lion share of existing research on digital discrimination is focused on exploring the social, economic, and legal underpinnings of the problem [3], [4], [6], [7], [8], [9]. In software engineering research, the problem is often tackled *after-the-fact*, where researchers are mainly concerned with formally describing, measuring, and testing for discrimination in existing software [10], [11]. However, an important aspect of the problem is often ignored. This aspect is embodied in the underlying system design decisions which allow established off-line patterns of discrimination to flourish on-line [6], [8]. To bridge this gap, in this paper, we propose a new perspective on digital discrimination, studying the problem through the lens of Requirements Engineering (RE). Our long-term goal is to develop data-driven software design solutions that can be utilized to mitigate concerns of digital discrimination in the Sharing Economy market. The impact of this research will extend to the entire population of Sharing Economy users by targeting barriers holding back minorities, and individuals in marginalized communities, from engaging in P2P economy activities. To achieve our research goals, in this paper, we:

- qualitatively analyze a large dataset of user feedback, collected from the Twitter feeds of eight popular Sharing Economy platforms. Our objective is to examine the distribution and types of digital discrimination in the online feedback of users of Sharing Economy platforms.
- synthesize existing research on digital discrimination in Sharing Economy. Our objective is to identify further empirical evidence on the different types of discrimination concerns typically raised in online user feedback.
- propose a conceptual crowd-driven domain model to represent the synergies and trade-offs between concerns of digital discrimination and the functional features and user goals of Sharing Economy platforms. This model is intended to provide Sharing Economy developers with in-depth insights into the complex realities of their ecosystems at an early stage of the software lifecycle.

The remainder of this paper is organized as follows. Section II provides a background on digital discrimination and justifies the need for our perspective. Section III describes our data collection and analysis process. Section IV proposes a domain model for representing digital discrimination in Sharing Economy. Section V describes a roadmap of future research. Section VI addresses the limitations of our work. Finally, Section VII concludes the paper.

II. Background and Motivation

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 online world, digital discrimination describes an outcome of an online business or service transaction that is influenced by participants' gender, race, age, or any other characteristic that is not directly related to the nature of the transaction. In traditional economy markets, discrimination is countered by imposing anti-discriminatory laws. For instance, in the United States, the Civil Rights Act of 1964 guarantees equal treatment of customers in public accommodations such as hotels or rental property. However, in the online world, discrimination takes a different form, often difficult to detect.

A. Background and Existing Work

A growing body of research has exposed a serious discrimination problem impacting some of the most popular Sharing Economy platforms nowadays. Edelman et al. [4] examined racial discrimination on Airbnb. The authors created guest accounts with canonical African-American and white names. A large number of messages were sent to random hosts inquiring about the availability of their places. The results showed that white guests were accepted 50% of the time, compared to 42% of typical African-American ones. Discrimination in the lodging business was also reported against members of the LGBT community. For example, Ahuja and Lyons [13] analyzed host responses to LGBT accounts in Airbnb. They found that discrimination against male-male pairs existed; hosts were more likely to not reply at all rather than replying "no" to male-male pairs inquiring about room availability.

Similar patterns of discrimination were observed over ridesharing platforms. Ge et al. [3] hired research assistants (RA) of different racial backgrounds to request UberX rides. The authors found that the waiting times for African-American RAs were significantly longer. In addition, more cancellations against African-American riders were observed than their white counterparts. Moody et al. [14] surveyed 1,100 users of UberPOOL and Lyft in order to find the social and economic factors that were prevalent in discriminatory practices. The findings suggested that white passengers that lived in predominantly white communities were more likely to discriminate against other passengers of other races.

Digital discrimination was also observed in less popular domains of Sharing Economy, such as freelancing. Hannak et al. [5] analyzed worker profiles on TaskRabbit and Fiverr. The reviews of each worker were extracted from their profiles and their gender and race were identified using their pictures. Using regression models, the authors found that there was a significant bias against white women on TaskRabbit and black men on TaskRabbit and Fiverr. In addition, black men often received lower rating scores than other workers with similar attributes. In another study, Foong et al. [15] collected self-determined hourly bill rates from the public profiles of 48,019 workers in the United States (48.8% women) on Upwork, a popular P2P freelancing platform. The authors found that key offline inequalities in pay also existed on Upwork. The median woman on Upwork requested only 74% of what the median man requested in hourly bill rate.

B. Motivation and Perspective

Our brief review shows that digital discrimination in Sharing Economy marketplaces is often studied from socio-economic or regulatory perspectives. In such studies, researchers seek to understand the social aspects of digital discrimination, determine its economic impact on individuals and societies, and introduce legislation to combat this phenomenon and counter its impacts [2], [8], [16], [17].

In software engineering research, digital discrimination, as a research problem, is often tied to the problem of software fairness. The goal of this line of research is to propose formal methods for quantifying discrimination and bias in software and develop algorithmic solutions for fairness testing and preservation, especially in Machine Learning systems where the data can be inherently biased [10], [11], [18], [19], [20]. However, these solutions often ignore the design decisions of the system, reflected in its internal set of functional and non-functional features which enable various established forms of discrimination to flourish in the operational environment of the system.

To bridge the gap in existing research, in this paper, we present a first-of-its-kind RE perspective on digital discrimination. This perspective is captured through a domain model that is constructed using popular requirements modeling notations. Models provide a framework for explicitly describing abstract salient concepts in a specific domain and formally reasoning about these concepts in order to create new knowledge [21]. Once domain knowledge is modeled, it can be preserved, communicated, and effectively used to instantiate and sustain innovation [22], [23]. In the context of RE, domain models are intended to provide concise and abstract representations of complex worlds through a smaller set of well-defined model components along with their relations (synergies and trade-offs) [24]. Our expectation is that a domain model would help software developers to understand the complex dynamics of their ecosystem, and thus, come up with design strategies that can mitigate discrimination concerns in their Sharing Economy systems.

III. Data collection and Qualitative Analysis

Existing research on digital discrimination relies on direct user surveys (questionnaires and interviews) and field studies to identify discrimination concerns among Sharing Economy users [4], [5], [13]. However, these data collection methods are extremely costly, and the sample size, or response rate, are often limited by factors such as the geographical area covered, number of subjects surveyed, and number of platforms studied. To overcome these limitations, in our research, we propose to exploit the social media platform Twitter, as a source of online software user feedback. Previous work showed that Twitter has become a very active channel of communication between software developers and their end-users [25]. This can be particularly observed when the problem is of a social nature. For example, a search for `discriminate AND (Uber OR Airbnb)` returns tweets sighting incidents of discrimination over these platforms. In fact, the hashtag *#AirbnbWhileBlack* has become the main place for reporting and highlighting potential racial bias on the rental app Airbnb. In what follows, we describe our data collection and analysis process in greater detail.

A. Scraping Twitter Data

To conduct our analysis, we collected tweets related to eight main players in the Sharing Economy market. These systems cover the domains of ridesharing (Uber and Lyft), lodging (Airbnb and Couchsurfing), food delivery (Doordash and UberEats), and freelancing (TaskRabbit and Fiverr). Our data collection process extended over the period of two full months, from November 1st to December 31st, 2019.

The data was collected using the Twitter Search API, considering only English tweets that contained the names of any of the eight systems included in our analysis. In total, 667,806 tweets were collected. Fig. 1 shows the number of tweets collected for each system[1].

B. Data Analysis

The main task after collecting our dataset is to locate our specific tweets of interest (discrimination related tweets). Several automated solutions have been proposed in the RE literature for mining Twitter data [25]. However, the majority of these solutions are proposed to classify tweets into generic maintenance tasks, such as feature requests and bug reports, with limited support for detecting issues of special nature, such as discrimination [26], [27]. In fact, finding such specific issues in large amounts of Twitter data has been described as finding a *needle in a haystack* [26].

To overcome this problem, in our analysis, we follow a snowballing approach. Snowballing is a commonly used strategy for exploratory data collection. This strategy starts with identifying an initial set of core strings (seeds) that are used for the first search query. Once the initial set of artifacts is located and examined, the search query is updated with new relevant terms acquired from the set, and another round of search is performed. The process continues until no new or relevant artifacts are found. Snowballing is commonly used in research as a reliable method for achieving high recall rates in tasks such as systematic literature reviews (SLRs) [28], search keyword identification [29], and Twitter data analysis [30].

To collect discrimination-related tweets, we began our search with the seed *discr*, which is the stem of the word *discrimination*. Stemming is necessary to count for the different morphological variants of the word (e.g. *discriminate*, *discriminating*, *discriminated*, and *discrimination*). In addition, the stems *bigot* for *bigotry* and *prejud* for *prejudice* were included since they often appear in English dictionaries as synonyms for the word *discrimination* [31]. Based on these seeds, we located our initial set of tweets.

Three researchers examined these tweets independently, following a systematic coding process to classify them into discriminatory and non-discriminatory tweets and to extract keywords for the subsequent search query [32]. Specifically, an initial meeting was held to discuss the common types of discrimination in today's society. Then, for each tweet, each researcher had to answer three main questions ***a)*** does the tweet describe a discrimination incident? ***b)*** what is the broad type of the discrimination concern raised in the tweet? and ***c)*** are there any other keywords that are strongly associated with the identified concern?. The coding process was carried over multiple sessions to avoid any fatigue issues and to ensure the integrity of the data [28]. A meeting was then held after the end of the coding phase to compile researchers' answers and to resolve coding issues. Such issues included conflicts in the classification of some types of concerns and missing discrimination-indicative words.

The set of identified keywords were picked based on how likely they would indicate discrimination. For example, in the tweet "*... my argument is UberPool should be accessible for all customers*" the keyword *discrimination* was used with the word *accessible*. Since the user is complaining about the lack of accessibility, the stem *accessib* was included in the set of indicator keywords for the next round of search. Extracted indicator words were then used to expand the query and the process was repeated for three rounds, until no more new keywords/tweets were found. At the end of this process, 22 unique discriminatory words or phrases (e.g. *service dog*) were extracted from the dataset. These words are listed in Table I.

It is important to point out that a large percentage of returned tweets were excluded from the final dataset for a variety of reasons. For instance, we did not include any tweets that were not tied to a user's general or specific experience. For example, the tweet *"Airbnb Works To Clean Up Its Reputation For Racial Discrimination In New 3-Year Report https://t.co/MRtHV07jjv"* was not included because the tweet was mainly publicizing a news article about discrimination over Airbnb. Another type of excluded tweets included tweets that were unrelated to discrimination to begin with. For example, the tweet *"@wjxt4 No, that's the parents discretion"* was returned as a possible match due to the presence of the stem *discr* from the word *discretion*, not *discrimination*.

[1] Data is available at https://seel.lsu.edu/data/re20.zip

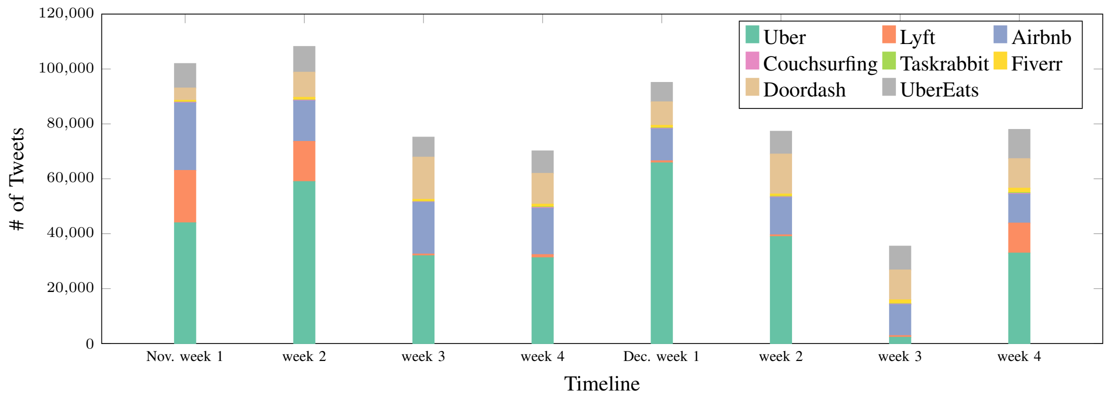

Fig. 1: The number of tweets collected for each Sharing Economy system during the data collection period.

TABLE I: The results of our snowballing analysis, showing the number of tweets **(included)**, o classified as actual discrimination-related concerns, after each round of snowballing.

Round	Terms	# of tweets collected (included)							
		Uber	Lyft	Airbnb	CouchSurfing	TaskRabbit	Fiverr	DoorDash	UberEats
1st	*discr (discrimination), prejud (prejudice), bigot (bigotry)*	280(**22**)	16(**10**)	104(**21**)	2(**0**)	0(**0**)	0(**0**)	8(**6**)	18(**2**)
2nd	*car seat, raci (racism, racist), race, deaf, accessib, gender, disab (disability, disable), service animal, dog, gay, sex (sexism, sexist), lgbt, elder, wheelchair, religi (religious, religion)*	1816(**127**)	228(**48**)	470(**59**)	4(**0**)	9(**1**)	5(**0**)	157(**11**)	184(**1**)
3rd	*minor, handicap, trans, infant*	126(**0**)	17(**1**)	8(**4**)	1(**0**)	1(**0**)	0(**0**)	7(**0**)	7(**1**)

C. Results

In general, four types of discrimination were detected in our data: racism, sexism, ableism, and parental. These types can be described as follows:

- **Racism:** Racism refers to any incidents of discrimination that are based on the perceived ethnicity (race) of users. Such concerns were observed in tweets such as *"@Airbnb you have racist host users who deny stay to guests that are not clearly white."*
- **Ableism (Disability):** Ableism is discrimination against people with disabilities, or who are perceived to have disabilities. Such concerns were observed in tweets such as *"@DoorDash please tell your drivers to consider disabled customers before telling them come out to retrieve their orders. It's highly offensive."*
- **Sexism:** Sexism is prejudice or discrimination based on a person's perceived sex or gender. Such concerns were observed in tweets such as *"@gem_zam @doxiebaby I literally was refused an Uber 2 weeks ago b/c the driver didn't want a gay passenger. Many lgbtq can not hide behind anything some can, and there is privilege in that, but many if not most of us cannot."*
- **Parental:** these concerns refer to cases where a consumer experienced discrimination due to being with a child. For example, *"@Uber_Support it's really important my account is opened. It's the only way I have to travel in Canada. Especially after I just gave birth! Is there a discrimination over woman having infants? Of course I have a car seat! But that sounds not the issue!!"*. Even though only few tweets were detected for this type, we decided to categorize these tweets under their own category.
- **Other:** under this category, we list any discrimination concerns that did not fit any of the previous categories, including tweets that reported discrimination incidents without specifying exactly what type of discrimination took place, such as *"@Airbnb @AirbnbHelp Why close my complaint on discriminatory behavior by host without a proper resolution? After accepting payment, host cancels the booking on discriminatory grounds. Is this what one has to expect from #AirBnB?."*

In terms of platforms, we observed that the ridesharing services, Uber and Lyft, suffered from the most cases of discrimination, followed by the lodging service Airbnb. In fact, these results were expected given the popularity of these services over other services such as Fiverr or Couchsurfing. We also observed that food delivery platforms had instances of discrimination, however, such tweets were not as common as in ridesharing data. These observations suggest that user data for these platforms should be collected over longer periods of time in order to increase the chances of capturing discrimination-related tweets. A breakdown of discrimination-related tweets (number of tweets) per platform is provided in Table II.

TABLE II: Number of discrimination tweets per each Sharing Economy platform.

Type	Uber	Lyft	Airbnb	Couchsurfing	TaskRabbit	Fiverr	DoorDash	UberEats
Racism	84	33	31	0	1	0	4	3
Ableism (Disability)	29	7	33	0	0	0	1	0
Sexism	24	15	11	0	0	0	5	0
Parental	3	1	0	0	0	0	1	0
Other	9	3	9	0	0	0	6	1

In summary, the results of the first phase of our analysis show that discrimination concerns do exist and they often get reported over social media. However, these concerns tend to be scarce and buried within large amounts of irrelevant tweets as well as vary in their quantity and intensity among different platforms. For instance, given the commercial nature of Sharing Economy platforms, their Twitter feeds tend to be overloaded with spam. Furthermore, some of the popular Sharing Economy services have become household names, or even used as verbs (e.g., *"Im going to Uber to work today"*). Therefore, isolating tweets that actually raise discrimination concerns among these tweets that simply mention the name of the service can be a very laborious task. We further noticed that a large number of tweets were very brief in describing incidents of discrimination with no details about the incident (e.g. *"My Uber driver mad sexist Jesus Christ"*). This can be attributed to the nature of Twitter as a micro-blogging service that does not allow messages longer than 280 characters.

Finally, we observed that the majority of discrimination-related tweets reported the experience of consumer (e.g., renters or riders), with only a small percentage reporting issues from the service provider side (e.g., hosts or drivers). This emphasizes the need for using other types of data collection methods (e.g., surveys and field studies) in order to capture the concerns of all types of users.

IV. A Domain Model of Digital Discrimination

Our perspective is captured through a conceptual crowd-driven model that represents concerns of digital discrimination along with their relations to user goals and system features in the different application domains of Sharing Economy. According to Eric Yu [24], *"conceptual modeling frameworks aim to offer succinct representations of certain aspects of complex realities through a small number of modeling constructs, with the intent that they can support some kinds of analysis."*

A. Model Notation

To represent our model, we adapt the notation of the hybrid domain modeling technique known as Feature-Goal Interdependency Graph (F-SIG). F-SIGs enable a comprehensive qualitative reasoning about the complex interplay between the functional and non-functional features (softgoals) of a domain, allowing developers to record design rationale and evaluate different design choices early in the process [33]. In F-SIGs, functional features are represented using a rectangle shape and softgoals are represented using a cloud shape. The edges of the graph represent the interrelationships among the softgoals and features of the domain. These relationships are typically represented as arrows accompanied with plus and minus signs to indicate the type and direction of impact among softgoals and features: `(-)` for `hurts`, `(--)` for `breaks`, `(?)` for `unknown` or unresolved, `(+)` for `helps`, and `(++)` for `makes`. To represent digital discrimination concerns (or anti-goals), we add a new component to the diagram in the shape of a shaded cloud. Furthermore, we slightly alter the semantics of the F-SIG notations to fit our narrative. Specifically, we use an arrow with an empty circle head (◯) to indicate a *mitigates* relationship (e.g., a certain feature mitigates a certain concern) and an arrow with a crossed-circle head ($\oplus$) to indicate a *leads to* relationship (e.g., a certain concern leads to a certain action). Our model along with our notations key is shown in Fig. 2. In what follows, we describe our model generation procedure along with its entities and their relationships in greater detail.

B. A Procedure for Model Generation

To generate our domain model, we use inductive reasoning, specifically, we iterate through the classified tweets to identify evidence points that could be used to establish our perspective. This grounded-theory based approach [34] is commonly used in RE research to interpret human-generated input in tasks such as mining rationale from app store reviews [35] or resolving ambiguities during requirements elicitation sessions [36]. To build our model, we go through our list of tweets, identifying the list of user concerns, goals, and features as they appear in the data. We further look for tweets connecting these different components to establish their relations.

In addition to evidence found in the data, we seek any empirical evidence in the literature related to digital discrimination in Sharing Economy. Such evidence is located using an *ad hoc* (exploratory) literature review strategy. An exploratory literature review starts by forming an initial search activity using appropriate terminology as a search query with no predefined research questions. The results of the search are then iteratively used to add more terms to the query, explore more venues, or more research groups. The process stops whenever a sufficient evidence is located. An *ad hoc* literature review does not guarantee scientific rigor as a systematic literature review. However, for exploratory studies, it can be sufficient to establish an initial perspective on the topic of interest (e.g., [37]). To identify our related studies, we search digital libraries using derivations of the following search query:

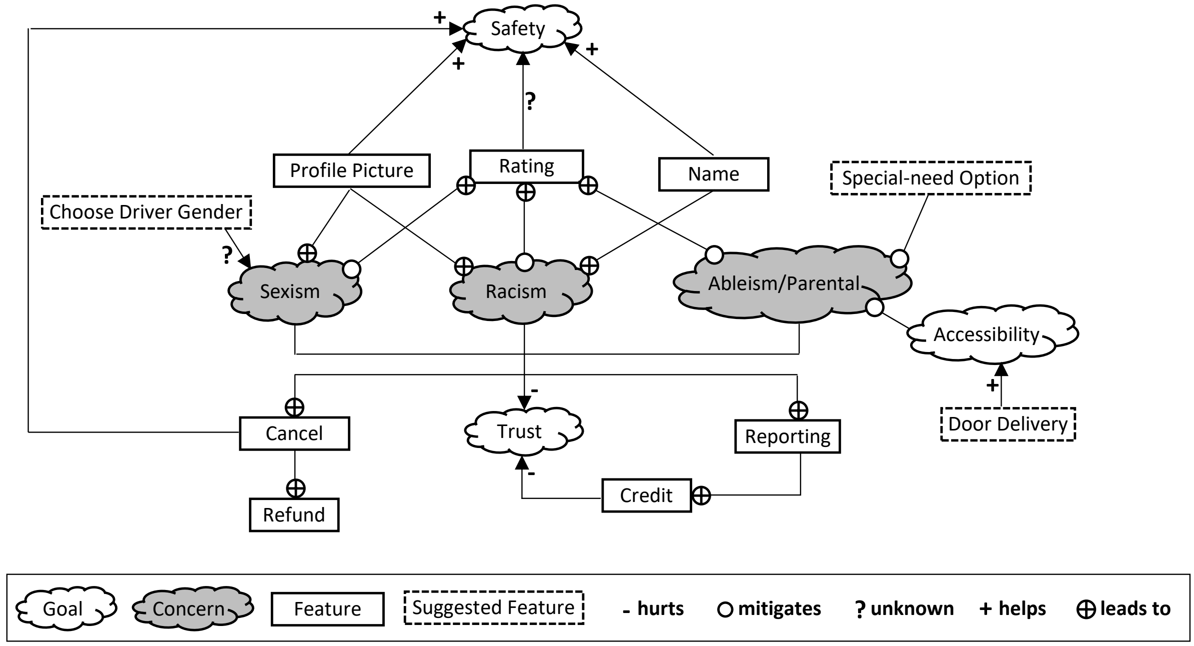

Fig. 2: A suggested domain model of digital discrimination in the Sharing Economy market.

((Digital discrimination) ***OR*** *discrimination* ***OR*** *racism* ***OR*** *sexism* ***OR*** *ableism* ***OR*** *disability* ***OR*** *accessibility)* ***AND*** *(ridesharing* ***OR*** *Uber* ***OR*** *Lyft* ***OR*** *Airbnb* ***OR*** *Couchsurfing* ***OR*** *Doordash* ***OR*** *UberEats* ***OR*** *TaskRabbit* ***OR*** *Fiverr* ***OR*** *(Sharing Economy)* ***OR*** *(gig economy))*

Our search was conducted over Google scholars, the ACM Digital Library, and IEEE Xplore. Papers were included in our analysis if they were long (4+ pages), written in English, and explicitly addressed an issue of digital discrimination over one of our subject platforms.

C. Model Entities

F-SIG models have four main components, functional features, user goals, user concerns, and their relationships. Generally speaking, a user concern (or anti-goal) can be defined as any functional or non-functional behavior of the system that might negatively impact its users experience or their overall well-being [38]. In the context of our domain, user concerns are mainly related to discrimination issues. These concerns are shown in gray cloud shapes in our model (Fig. 2). Ableism and parental discrimination are represented as one entity as incidents of these two types of discrimination are frequently accompanied with similar accessibility concerns or requests for some sort of special accommodation (e.g., car seats).

D. User Goals

A user goal can be described as any abstract objective that the system under consideration should achieve [39]. User goals in Sharing Economy can range from economic growth to building up social capital [2]. In our model, we identify three main goals that are directly related to concerns of digital discrimination. These goals can be described as follows:

- **Trust:** Trust is a key user goal in Sharing Economy [40], [41]. The concept itself stems from the fact that conducting business transactions with strangers involves inherent risk, therefore, providers and consumers at both ends of the P2P connection need to establish a certain level of mutual trust before a transaction can take place [42]. For example, riders need some sort of trust before they can get in a stranger's car and hosts need to trust that guests will not damage their property. Without such a minimum level of trust, it can be hard for any Sharing Economy business to flourish [43].
  The relationship between trust, as a user goal, and discrimination, as a concern, has long been established in the literature [44], [45]. In general, any form of discrimination leads to a decline in trust between users and in the platform in general [46]. In fact, recent evidence has identified the lack of trust as a main reason for the poor participation of minorities, or users in disadvantaged communities, in Sharing Economy [2], [40], [47]. Even though the word *trust* was rarely mentioned in our discrimination-related tweets, the notion of trust was common in the data, with tweets such as *"As a black person I know not to trust AirBnB as they had a problem a few years back where racists would cancel the listing or raise the rates when someone of color tried to rent. It got so bad that there is a Airbnb type company for people of color"*. Given these observations, we establish a `hurts` relationship between all forms of discrimination and users' Trust in the platform.
- **Accessibility:** Accessibility is a major user goal for people with disability as well as parents. In general, user accessibility complaints revolve around the services being inaccessible or not being accommodating of their special

needs, such as the lack of accommodation for service dogs or wheelchairs. These concerns appeared in tweets such as *"...I am a wheelchair user and unfortunately, the host misrepresented the accessibility of their house..."* and *"...The driver overcharged them because they had service dogs..."* In general, accessibility decreases concerns of Ableism and Parental discrimination. Therefore, we establish a `mitigates` relationship between accessibility and the concern representing these two types of discrimination in the model.
- **Safety:** As mentioned earlier, Sharing Economy is inherently risky. Therefore, safety is another major goal of Sharing Economy users [16], [48]. This user goal is enforced through multiple functional features (users' pictures, names, ratings, cancellation, etc.) that are designed to enhance the safety of the app for its users.

E. Functional Features

A functional feature can be defined as any tangible behavior of the system. To build our model, we only include features of Sharing Economy systems that are related, according to our data, to digital discrimination, such as ratings, user profiles (pictures and names), and canceling transactions. A major goal of our domain model is to expose the synergies and trade-offs between discrimination concerns and these features. To expose such relations, we scan through our set of discrimination tweets, looking for tweets that connect any functional features to discrimination incidents. We further synthesize existing literature to seek evidence supporting or refuting our established relations. The features present in our data are:

- **Cancel:** Cancellation is directly related (`helps`) to safety, users can cancel a transaction if they do not feel safe to receive of provide a service. In our data, cancellation is the most common action associated with discrimination incidents. Such incidents were detected in tweets such as *"So an @Uber driver cancels my trip because I wasn't physically able to walk and meet her"* and *"I've now had four drivers cancel on me in a row outside a gay venue @uber."* Given these strong signals in the data, we assign a `leads to` relationship between all types of discrimination and the feature Cancel.
- **Refund:** refund is often associated with service cancellation. This relationship appeared in tweets such as, *"@Uber doesn't want to refund me for a racist driver canceling my trip"* and *".. the cabin refunded and canceled us when they found out about my service dog which is very illegal, especially since it's an Airbnb"*. Therefore, we assign a `leads to` relationship between Cancel and Refund.
- **Profile Picture:** profile pictures are a necessity for safety. Pictures of users are meant to reduce anonymity as well as facilitate identification offline [43], thus, we establish a `helps` relationship between Profile Picture and Safety. However, profile pictures have been recently identified as a main enabler of digital discrimination. Basically, service providers or consumers can decline or cancel a transaction based on the physical characteristics of the other party as they appear in their profile pictures. Concerns about pictures appeared in tweets such as *"@denvercoder I wouldn't use AirBnB anyway because they want a pictures of you. It's none of their business what race I am. The only reason I can see why they want a photo is so people can discriminate based on race."*
In the literature, discrimination based on profile pictures was observed by Ert et al. [43] who found that more trustworthy-looking Airbnb hosts charged higher prices for similar apartments. In a follow up study of 1,020 Airbnb listings, Jaeger et al. [49] reported that photo-based impressions of hosts' attractiveness significantly influenced their rental prices. The authors also reported that compared to white hosts, black hosts charged lower prices for their apartments. In general, both studies suggested that people were willing to pay more for a similar apartment if the host was perceived to be more attractive, trustworthy, or white in their profile pictures. Based on these observations, a `leads to` relationship is established between Profile Picture and concerns of Racism and Sexism.
- **Name:** Similar to profile pictures, user names were also found to be a main trigger of racial discrimination. Such incidents were reported in tweets such as *"speaking from experience always keep a friend with a white-sounding name close if you planning on getting an @Airbnb or holla at @Uber in NYC anytime soon"*. In the literature, multiple independent field studies confirmed the link between user name and discrimination. For instance, Cui et al. [50] found that requests from Airbnb guests with African American–sounding names were 19.2% less likely to be accepted than those with white-sounding names. Edelman et al. [4] reported that guests with African American–sounding names were 16% less likely to be accommodated relative to identical guests with white-sounding names. In a field study, Ge et al. [3] observed more frequent cancellations by Uber drivers against passengers when they used African American-sounding names. Given these observations, we establish a `leads to` relationship between Name and Racism. However, as mentioned earlier, more information shared on the online platform about users is necessary for safety reasons. Therefore, the relationship `helps` is established in our model between Name and Safety.
- **Rating:** ratings and reviews, or reputation systems, are commonly used in Sharing Economy platforms to establish trust [51], [52]. Users often resort to the provided rating system to rank, often on a 1-5 star scale, their experiences with other users (providers or consumers). In our dataset, tweets about this feature often detected using words such as *star*, *rating*, and *review*. In general, most of these tweets described situations where users used lower ratings to express their anger towards discrimination. For example, *"My Uber driver said you choose to be gay, so why get offended if somebody says he doesn't like you?*

It's just an opinion. 1 star rating mate" and *" ... and he told her that she had to chill b/c his brother was gay, so that kind of talk wasn't going to be tolerated. She told him she was gon give him 1 Star ..."*. Based on these tweets, we establish a `leads to` relationship between concerns of discrimination and Ratings.

In the literature, ratings and reviews were found to help countering discrimination. For example, in their field experiment, Cui et al. [50] reported that positive reviews posted on Airbnb guests' pages significantly reduced discrimination towards guests' with African-sounding names. Another study by Brown et al. [53] reported that rider ratings may reduce proxy discrimination by drivers who can use star ratings to infer how safe or considerate a rider may be. Based on these observations, we add a `mitigates` relationship between Rating as a feature and concerns of discrimination.

On the flip side, several tweets in our data revealed that rating systems could be abused. Some users reported getting a negative rating based on discriminatory reasons. Examples of such tweets included, *"so not only was my Airbnb host racist, but she left a false review about me saying I threw a party in her Airbnb and broke multiple rules"* and *"Wowwww I just read my review for the Airbnb I stayed at In NY. My host was actually racist asf!!!! Never doing @Airbnb again"*. Furthermore, ratings seem to be raising safety issues for some users who reported being concerned about the ramifications of leaving a bad review for other users, for example, *"@DoorDash and @DoorDash_Help one of your dashers is harassing me after I gave him a bad review for being racist and begging for a 5 star rating"*. Based on these data points, we establish in our model a `unknown (?)` relationship between Rating as a feature and Safety as a user goal.

- **Reporting:** Reporting refers to any feature that enables users to report discriminatory behavior to the platform. In general, people resort to reporting as a way of fighting back discrimination or violations of anti-discrimination policies. For example, in the tweet *"@AskLyft Cool! This will be my second report this month. It's also my second lyft ride this month. What is lyft doing to ensure their drivers understand that accommodating service animals is not optional"*, the user reported Lyft drivers who refused to accommodate their service animal in a clear violation of Lyft's service-animal policy[2]. Based on these observations, we establish a `leads to` relationship between the different concerns of discrimination in our domain and Reporting as a feature.
- **Credit:** in several incidents, users reported getting monetary credit as a token of courtesy to re-establish trust in the platform. However, the general feeling towards this feature seems to be negative. This was reflected in tweets such as *"@DoorDash As a disabled person I rely on sis for food ... I'm tired of having to struggle to get to my lobby when delivery drivers don't want to come to door. I'm tired of minuscule credits when providing feedback"* and *"... we're sorry drivers don't care that you're disabled, here's $3 in credit for future trips ..."*. Based on these tweets, we establish a `leads to` relationship between Reporting and Credit and a `hurts` relationship between Credit and Trust.

[2]https://help.lyft.com/hc/en-us/articles/115013080048

F. User Suggested Features

In addition to the features discussed above, users in our data also suggested multiple features to counter discrimination over specific platforms. These features, represented using dashed rectangles in our model, include:

- **Special-Need:** this feature appeared in a tweet of a disabled person suggesting a handicap option for riders with special needs, *"@Uber why do you not have a handicap option??? People in wheelchairs need #uber too."* Our expectation is that this option could be useful in cities where there is no special accommodation option, such as Uber's Wheelchair-Accessible Vehicles (WAVs). In fact, declaring special-needs before booking (e.g., a baby seat, wheelchair, or service dog) could minimize after-the-fact cancellations such as when a driver cancels after they arrive and realize the person is using wheelchair or with a service dog. Given these observations, a `mitigates` relationship is established between Special-Needs and Ableism/Parental.
- **Door Delivery:** a tweet suggested that DoorDash should add a feature for door delivery. This feature is intended to make the app more accessible to disabled users *"@DoorDash please add a door-delivery option to help disabled customers who cant come out to retrieve their orders"*. Based on this tweet, a `helps` relationship is established between Door Delivery and Accessibility.
- **Choosing Driver Gender:** an interesting suggestion by some users is to have an option for selecting the gender/race of the driver. The assumption is that this option would minimize discrimination, for instance, *"@Uber can we pls weed folks out before you pick them up. let me choose my driver's gender. The point is that we should all have options that make us feel comfortable with"*. However, other users are concerned that this option would actually lead to more discrimination, for instance *"@Alt_Deadpool @AskLyft @Uber_Support I have a problem with you choosing based on gender because its sexist. Not to mention impractical business-wise because women drivers represent UNDER 10% of drivers yet 50% of passengers. are you ready to pay double/triple the fares for your gender-based employment discrimination"*. Given these tweets, we assign an `unknown (?)` relationship between the Choosing Driver Gender and Sexism.

G. Validation

A summary of the primary studies included in our analysis, the model entities they study, and sample tweets related to these entities are shown in Table III. While our model provides an initial overview of a domain at a very abstract level, it is important to point out that it only reflects an instance that is weighted based on the current analysis of the data. In other words, our model is an incomplete first step to capture a complex domain. These limitations stem from the fact that the data input (a space of time on Twitter) is incomplete and the literature review is ad-hoc rather than systematic. However, our expectation is that more data (online feedback, direct surveys, and evidence from the literature) will help us to identify more domain entities as well as resolve their interdependency relations, thus, enhance the completeness, correctness, and consistency of the proposed model. A crowd-driven model will enable app developers to better understand the main end-user concerns in their domain of operation as well as identify design tactics that can enhance their user goals and mitigate their concerns.

V. ROADMAP

The perspective developed in this paper is a first step of a long-term research plan that is aimed at addressing concerns of digital discrimination in Sharing Economy. To achieve our short and long term visions, our work will be extended along the two following directions of future work:

- **Data collection and analysis:** Our data collection process will be expanded in terms of time, scale, and source of data. Specifically, discrimination-related data will be harvested over longer periods of time, covering a very broad range of Sharing Economy platforms. Our objective is to study how different patterns of digital discrimination change in terms of intensity and frequency over time and over different application domains. Other sources of online user feedback, such as app store reviews, will be included in our analysis. In addition, a set of optimized search queries (e.g., targeting active hashtags of digital discrimination), text processing (e.g., spam filtering), and systematic data coding techniques will be utilized to enable further investigations of the problem.
- **Extrinsic Evaluation:** Extrinsic evaluation is concerned with criteria relating to the model function, or role, in relationship to its purpose (e.g., validation through experience). To conduct such analysis, our generated models will be provided to selected groups of Sharing Economy software developers to be used as an integral part of their development activities. Evaluation data will be collected through surveys that will measure the impact of such models on idea formulation and the success or failure of handling discrimination concerns among the end-users of these platforms.

VI. THREATS TO VALIDITY

The study conducted in this paper suffers from several methodological constraints that might jeopardize the validity of our findings. A main threat to the validity of our study stems from the fact that our data was collected from Twitter and only for a relatively limited period of time. A recent report by the Pew Research Center has shown that most Twitter users rarely tweet, and the most prolific 10% create 80% of traffic among adult users. In the U.S., only 22% of American adults use Twitter, and this segment tends to be younger, more highly educated and wealthier than the general public. The report also states that Twitter users are more likely to be sensitive to issues of racial discrimination [54]. Twitter might also conceal sampling bias given that the demographic of users (e.g., gender, age, and location) is unknown. However, as mention earlier, our goal in this paper is to develop a preliminary perspective of the problem. Twitter, as a social media platforms, is expected to provide a low-cost preliminary evidence given that discrimination is inherently a social problem. Nonetheless, we acknowledge the fact that data collected over longer periods of time and from other channels of feedback, such as app store reviews, online blogs, and direct user surveys, are necessary to achieve a better coverage of the problem and eliminate sampling bias, especially for smaller platforms that do not typically receive a large number of tweets.

Furthermore, only a few popular Sharing Economy platforms were considered in the analysis. This might generate some external validity threats given that there are hundreds of such platforms operating in different geographical areas. In our analysis, we are interested in platforms operating in large geographical areas with the biggest market share, thus, we narrowed down the market to its most *fit* elements from a user perspective. Popular platforms receive significantly more crowd feedback in comparison to smaller platforms [55]. Furthermore, selecting mature platforms gives smaller and newcomer platforms a chance to learn from the mistakes of the big players in the market [56].

Another threat might stem from the fact that our review of existing literature was exploratory rather than systematic, focusing only on studies related to discrimination issues raised in our tweets. However, the majority of these studies included large scale field studies which were conducted using large numbers of data points (e.g., users profiles, survey participants, and experimental subjects). Therefore, we believe that our review process was sufficient to provide a preliminary evidence on our inferred model relations. Nonetheless, a systematic literature review is still necessary to generate a more mature preview of existing work.

Finally, the method for classifying our data and generating our diagram relied on our own interpretation of the data. While this method followed the general guidelines of grounded-theory, some methodological compromises had to be made. For example, the large size of our dataset and the extreme sparsity of discrimination tweets prevented us from using techniques such as random or stratified sampling to analyze the data. Furthermore, subjectivity concerns could be raised about our manual classification of discrimination tweets. These concerns were partially mitigated by using three judges and majority voting for conflict resolution.

TABLE III: Primary studies identified through our literature review and the model entities they relate to along with sample tweets.

Primary Study	Model Entity	Sample tweet
[40], [41], [42], [43], [44], [45], [46], [47]	Trust	*"As a black person I know not to trust AirBnB as they had a problem a few years back where racists would cancel the listing or raise the rates when someone of color tried to rent."*
[16], [48]	Safety	*"@transportforall @Uber Can you remind us how many wheelchair accessible vehicles #uber actually have? #london #blacktaxi are 100% wav. Disgraceful and extremely irresponsible of you to advertise #uber considering their serious safety issues."*
[43], [49]	Profile Picture	*"@denvercoder I wouldn't use AirBnB anyway because they want a pictures of you. It's none of their business what race I am. The only reason I can see why they want a photo is so people can discriminate based on race."*
[3], [4], [50]	Name	*"speaking from experience always keep a friend with a white-sounding name close if you planning on getting an @Airbnb or holla at @Uber in NYC anytime soon".*
[50], [53]	Rating	*"My Uber driver said you choose to be gay, so why get offended if somebody says he doesn't like you? It's just an opinion. 1 star rating mate"*

VII. Conclusion

In this paper, we developed an RE perspective of the problem of digital discrimination in Sharing Economy. Our perspective is captured through a conceptual domain model which was generated based on analyzing online user feedback and synthesizing existing empirical evidence. Our model provides a first-of-its-kind conceptual understanding of the problem, including the main concerns of discrimination along with their interactions with the user goals and functional features of the different Sharing Economy platforms. Based on our model, we proposed a research agenda aimed at establishing an interdisciplinary effort for understanding and mitigating issues of digital discrimination in the Sharing Economy market.

References

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

[2] 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.

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

[4] 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.

[5] 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.

[6] M. Chenga and C. Foley, "The sharing economy and digital discrimination: The case of Airbnb," *International Journal of Hospitality Management*, vol. 70, pp. 95–98, 2018.

[7] 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.

[8] K. Levy and S. Barocas, "Designing against discrimination in online markets," *Berkeley Technology Law Journal*, vol. 32, 2017.

[9] N. Leong and A. Belzer, "The new public accommodations: Race discrimination in the platform economy," *Georgetown Law Journal*, vol. 105, no. 5, pp. 1271–1322, 2017.

[10] S. Friedler, C. Scheidegger, S. Venkatasubramanian, S. Choudhary, E. Hamilton, and D. Roth, "A comparative study of fairness-enhancing interventions in machine learning," in *Conference on Fairness, Accountability, and Transparency*, 2019, pp. 329–338.

[11] S. Galhotra, Y. Brun, and A. Meliou, "Fairness testing: Testing software for discrimination," in *Joint Meeting on Foundations of Software Engineering*, 2017, pp. 498–510.

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

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

[14] 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.

[15] E. Foong, N. Vincent, B. Hecht, and E. M. 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.

[16] B. Rogers, "The social costs of Uber," vol. 82, no. 6, 2015.

[17] 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.

[18] R. Angell, B. Johnson, Y. Brun, and A. Meliou, "Themis: Automatically testing software for discrimination," in *ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2018, pp. 871–875.

[19] F. Calmon, D. Wei, B. Vinzamuri, K. Ramamurthy, and K. Varshney, "Optimized pre-processing for discrimination," in *Advances in Neural Information Processing Systems 30*, 2017, pp. 3995–4004.

[20] F. Kamiran and T. Calders, "Data preprocessing techniques for classification without discrimination," *Knowledge and Information Systems*, vol. 33, pp. 1–33, 2012.

[21] J. Holt, S. Perry, and M. Brownsword, *Model-Based Requirements Engineering*. The Institution of Engineering and Technology, 2011.

[22] O. Glassey, "Method and instruments for modeling integrated knowledge," *Knowledge and Process Management*, vol. 15, no. 4, pp. 247–257, 2008.

[23] M. Lubars, C. Potts, and C. Richter, "A review of the state of the practice in requirements modeling," in *International Symposium on Requirements Engineering*, 1993, pp. 2–17.

[24] E. S. Yu, "Conceptual modeling: Foundations and applications," A. T. Borgida, V. K. Chaudhri, P. Giorgini, and E. S. Yu, Eds. Springer-Verlag, 2009, ch. Social Modeling and i*, pp. 99–121.

[25] G. Williams and A. Mahmoud, "Mining Twitter feeds for software user requirements," in *International Requirements Engineering Conference*, 2017, pp. 1–10.

[26] E. Guzman, R. Alkadhi, and N. Seyff, "A needle in a haystack: What do Twitter users say about software?" in *International Requirements Engineering Conference*, 2016, pp. 96–105.

[27] N. Jha and A. Mahmoud, "Using frame semantics for classifying and summarizing application store reviews," *Empirical Software Engineering*, vol. 23, no. 6, pp. 3734—3767, 2018.

[28] C. Wohlin, P. Runeson, M. Höst, M. Ohlsson, B. Regnell, and A. Wesslèn, *Experimentation in Software Engineering*. Springer, 2012.

[29] X. Du, J. Bian, and M. Prosperi, "An operational deep learning pipeline for classifying life events from individual tweets," in *Annual International Symposium on Information Management and Big Data*, 2018, pp. 54–66.

[30] T. Denskus and D. E. Esser, "Social media and global development rituals: A content analysis of blogs and tweets on the 2010 MDG summit," *Third World Quarterly*, vol. 34, no. 3, pp. 405–422, 2013.

[31] E. Loper, E. Klein, and S. Bird, *Natural language processing with Python*. O'Reilly, 2009.

[32] J. Saldana, *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd, 2015.

[33] S. Jarzabek, B. Yang, and S. Yoeun, "Addressing quality attributes in domain analysis for product lines," *IEE Proceedings - Software*, vol. 153, no. 2, pp. 61–73, 2006.

[34] J. Corbin and A. Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE, 1998.

[35] Z. Kurtanoviċ and W. Maalej, "Mining user rationale from software reviews," in *International Requirements Engineering Conference*, 2017, pp. 61–70.

[36] Y. Elrakaiby, A. Ferrari, P. Spoletini, S. Gnesi, and B. Nuseibeh, "Using argumentation to explain ambiguity in requirements elicitation interviews," in *International Requirements Engineering Conference*, 2017, pp. 51–60.

[37] R. Rosario and G. Widmeyer, "An exploratory review of design principles in constructivist gaming learning environments," *Journal of Information Systems Education*, vol. 20, no. 3, pp. 289–300, 2016.

[38] G. Williams and A. Mahmoud, "Modeling user concerns in the app store: A case study on the rise and fall of Yik Yak," in *International Requirements Engineering Conference*, 2018, pp. 64–75.

[39] A. van Lamsweerde, "Goal-oriented requirements engineering: A guided tour," in *IEEE International Symposium on Requirements Engineering*, 2001, pp. 249–262.

[40] F. Hawlitschek, B. Notheisen, and T. Teubner, "The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy," *Electronic Commerce Research and Applications*, vol. 29, pp. 50–63, 2018.

[41] M. ter Huurne, A. Ronteltap, R. Corten, and V. Buskens, "Antecedents of trust in the sharing economy: A systematic review," *Journal of Consumer Behaviour*, vol. 16, no. 6, pp. 485–498, 2017.

[42] C. van der Cruijsen, M. Doll, and F. van Hoenselaar, "Trust in other people and the usage of peer platform markets," *Journal of Economic Behavior & Organization*, vol. 166, pp. 751–766, 2019.

[43] E. Ert, A. Fleischer, and N. Magen, "Trust and reputation in the sharing economy: The role of personal photos in Airbnb," *Tourism Management*, vol. 55, pp. 62–73, 2016.

[44] K. Berg Nodtvedt, H. Sjastad, S. Skard, H. Thorbjornsen, and J. Van Bavel, "Racial bias in the sharing economy and the role of trust and self-congruence," *Social Science Research Network*, 2020.

[45] S. Jhaver, Y. Karpfen, and J. Antin, "Algorithmic anxiety and coping strategies of Airbnb hosts," in *CHI Conference on Human Factors in Computing Systems*, no. 412, 2018, pp. 1–12.

[46] L. Einav, C. Farronato, and J. Levin, "Peer-to-peer markets," *Annual Review of Economics*, vol. 8, no. 1, pp. 615–635, 2016.

[47] R. McLachlan, C. Opila, N. Shah, E. Sun, and M. Naaman, "You can't always get what you want: Challenges in p2p resource sharing," in *CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2016, p. 1301–1307.

[48] B. Malin and C. Chandler, "Free to work anxiously: Splintering precarity among drivers for Uber and Lyft," *Communication, Culture & Critique*, vol. 10, no. 2, pp. 382–400, 2017.

[49] B. Jaeger, W. Sleegers, A. Evans, M. Stel, and I. van Beest, "The effects of facial attractiveness and trustworthiness in online peer-to-peer markets," *Journal of Economic Psychology*, vol. 75, pp. 102–125, 2019.

[50] R. Cui, J. Li, and D. Zhang, "Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb," *Management Science*, 2019.

[51] M. Luca, "Designing online marketplaces: Trust and reputation mechanisms," *Innovation Policy and the Economy*, no. 17, pp. 77–93, 2017.

[52] T. Ye, R. Alahmad, C. Pierce, and L. Robert Jr., "Race and rating on sharing economy platforms: The effect of race similarity and reputation on trust and booking intention in Airbnb," in *International Conference on Information Systems*, 2017.

[53] A. Brown, "Prevalence and mechanisms of discrimination: Evidence from the ride-hail and taxi industries," *Journal of Planning Education and Research*, 2019.

[54] S. Wojcik and A. Hughes, *Sizing Up Twitter Users*. Pew Research Center, 2019.

[55] S. Mcilroy, W. Shang, N. Ali, and A. Hassan, "User reviews of top mobile apps in apple and google app stores," *Communications of the ACM*, vol. 60, no. 11, pp. 62–67, 2017.

[56] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in *International Requirements Engineering Conference*, 2013, pp. 125–134.