# Uber and Lyft Accidentally Shared Drivers' Social Security Numbers with Social Media Companies

Imagine applying to be a driver for Uber or Lyft and having to enter your Social Security number on their website. You'd expect that information to be kept totally private, but researchers just discovered these companies were accidentally sharing this sensitive information with Facebook and TikTok.

## Tracking Pixel Issue Was Responsible

When you visit almost any website today, there are usually invisible tools called "tracking pixels" that collect information about what you're doing on the page.

Tracking pixels allow companies to understand their visitors better. Meta and TikTok provide these tracking pixels to millions of websites.

When new drivers were signing up on Uber and Lyft's websites, they had to enter their Social Security numbers into a form. Even though the companies tried to protect these numbers by scrambling them. the tracking pixels were secretly collecting this information and sending it to Facebook and TikTok.

## The Researchers Decoded The SSN's Using A Massive Table Of Every SSN in Existence

For Uber, they were accidently collecting the Social Security numbers and mislabeling them as phone numbers and sending it on to Meta and TikTok as phone numbers.

"The SSNs were transmitted as an unsalted SHA256 hash of the worker's SSN. One might argue that this level of obfuscation offers privacy because the 9-digit SSN was not itself exposed. However, this would be a naïve interpretation," the researchers explained. They demonstrated that creating a complete lookup table mapping every possible SSN to its hash would take less than 2 minutes on a basic server, making the hashing protection essentially worthless.

For Lyft specifically, the researchers found the issue arose because they named their SSN field "ssn_validation" rather than simply "ssn" - causing it to bypass Meta's filters that were supposed to prevent collection of sensitive data. The tracking pixels then shared these SSN hashes with Facebook and TikTok's servers.

**Read The Whole Research Report on Following Pages**

# Gig Work at What Cost? Exploring Privacy Risks of Gig Work Platform Participation in the U.S.

Amogh Pradeep
Northeastern University
Boston, Massachusetts, USA
amoghbl1@ccs.neu.edu

Johanna Gunawan
Northeastern University
Boston, Massachusetts, USA
gunawan.jo@northeastern.edu

Álvaro Feal
Northeastern University
Boston, Massachusetts, USA
l.feal@northeastern.edu

Woodrow Hartzog
Boston University
Boston, Massachusetts, USA

David Choffnes
Northeastern University
Boston, Massachusetts, USA
choffnes@ccs.neu.edu

## Abstract

In recent years, "gig work" platforms have gained popularity as a way for individuals to earn money; as of 2021, 16% of Americans have at some point earned money from such platforms [13]. Despite their popularity and their history of unfair data collection practices and worker safety, little is known about the data collected from workers (and users) by gig platforms and about the privacy dark pattern designs present in their apps.

This paper presents an empirical measurement of 16 gig work platforms' data practices in the U.S. We analyze what data is collected by these platforms, and how it is shared and used. Finally, we consider how these practices constitute privacy dark patterns. To that end, we develop a novel combination of methods to address gig-worker-specific challenges in experimentation and data collection, enabling the largest in-depth study of such platforms to date. We find extensive data collection and sharing with 60 third parties—including sharing reversible hashes of worker Social Security Numbers (SSNs)—along with dark patterns that subject workers to greater privacy risk and opportunistically use collected data to nag workers in off-platform messages. We conclude this paper with proposed interdisciplinary mitigations for improving gig worker privacy protections. After we disclosed our SSN-related findings to affected platforms, the platforms confirmed that the issue had been mitigated. This is consistent with our independent audit of the affected platforms. Analysis code and redacted datasets will be made available to those who wish to reproduce our findings.

## Keywords

gig work, privacy, tracking, dark patterns

## 1 Introduction

In recent years, "gig work" platforms have gained popularity for individuals as a way to earn money; as of 2021, 16% of Americans have at some point earned money from such platforms [13]. Gig work allows workers to carry out tasks on demand, giving them freedom to earn money with their own devices and on their own time. However, this lack of formality leaves gig workers in a legal grey area, where traditional labor protections are not consistently provided across jurisdictions and platforms strongly oppose regulation [21]. This leaves gig workers vulnerable to harm and exploitation when participating on these platforms [8, 48].

Gig work platforms are particularly important from an online privacy perspective. On the one hand, gig work platforms might need to collect highly sensitive data (e.g., name, location, and driver license information, social security number) from workers to enable dispatching of gig work and remit payment to workers. On the other hand, such sensitive data collection opens the door to privacy abuses such as sharing such highly sensitive information with third parties in ways that are *not* required for the platform to function. We have seen examples of such data misuse by gig platforms in the past, e.g., when The Federal Trade Commission (FTC, US regulator) and Uber reached a settlement because of the company's deceptive data privacy and security practices [32]. However, despite gig work platforms' history of unfair data collection practices, little is known about their general data practices and privacy dark pattern designs.

The key challenge for conducting such rigorous, empirical study of gig work platforms' data practices is that measuring them in depth requires participating on the platform in the real world—in contrast to studies that rely only on limited static [20, 28, 58, 65] or dynamic analysis [19, 34, 44, 76] of code deployed on websites or apps using fictional accounts (or no account at all). For instance, platforms often require physical mobility, e.g., ridesharing gig workers may need to use a mobile app and physically travel, so measuring data practices during such activities in real time while traveling requires custom, portable instrumentation. We further note that identity verification requirements impose constraints on the scale of testing, as they prevent the use of fictional accounts. As a result, prior work on gig worker privacy has largely focused on worker perspectives [74]. Direct observation of gig platforms in-the-wild is uncommon and tends to focus on only one or two platforms at a time [71, 80].

In this paper, we present the largest-scale known empirical study on gig work platforms' data practices in the U.S. To do so, we use a novel combination of methods that include a portable mobile testing infrastructure that allows app testing while on the go, registering for and completing gigs as real workers by completing requisite background checks, and monitoring off-platform communication


*Proceedings on Privacy Enhancing Technologies YYYY(X), 491–510*

https://doi.org/XXXXXXX.XXXXXXX

via text and email. We use these methods to analyze 16 gig work platforms and answer the following research questions:

- **RQ1:** What harmful data sharing and data use practices exist on gig work platforms and how do they vary by modality and user type?
- **RQ2:** How do the requirements for joining gig platforms and accessing work incorporate privacy dark patterns?

In summary, this work presents a multifaceted study of the data and design practices that 16 gig platforms employ at the cost of worker privacy. Reconciling in-the-wild gig platform data against the design and off-platform components of gig work experiences, we examine what personally identifiable information (PII) is collected by gig platforms, how it is collected, and how it is used or abused. Specifically, we first register as workers on 16 platforms, observing what types of information platforms collect from direct user input and whether they send this information via network traffic (and to whom). We additionally collect off-platform SMS and email communications sent by platforms to understand how worker-submitted personal contact information is used. From this data, we relate gig platform registration practices to privacy dark patterns. We attempt to undergo extensive background checks on all platforms, and obtain ready-to-work accounts for 7 platforms. Next, we attempt gig work on platforms that directly facilitate and monitor worker labor to capture real-time data collection practices during a gig. Last, we compare platforms' data sharing practices for worker and non-worker participation to better understand the unique harms workers may face.

Our study reveals that:

- All platforms in our study, except one, i.e., 15 of 16 share worker PII with unrelated third parties. All 15 share Android Advertisement IDs (Ad IDs), 14 share worker full names, and 12 share emails. Two instances include reversible hashes of highly sensitive Social Security Numbers (SSNs) being shared with third parties. In all, we observe 60 third parties; these include social media platforms (Facebook, TikTok, etc.) and even advertisement platforms (DoubleClick, Google Ads, etc.).
- Gig platforms' registration practices often constitute *forced action* dark patterns, leading to more data collection. Some platforms (petsitting/care platforms) burden prospective workers with the financial cost of background checks. Some platforms abuse PII collected at registration through spammy, high-volume SMS and email messages, especially for partially registered workers. Email messages often contain third-party trackers. These constitute *nagging* dark patterns, often arising from *bad defaults*.

Our work provides new insight into the data vulnerability of gig workers, contributes a new methodology for observing gig platform data practices in-the-wild, and includes dimensional[1] analysis of privacy dark patterns.[2] To facilitate additional research in this space, we will make our analysis code and redacted datasets available. In addition, we disclosed our findings regarding unauthorized third-party SSN-related data collection to the companies involved (Uber, Lyft, Tiktok, and Meta). The gig work platforms confirmed that this kind of data collection was unintentional, and they changed their website configurations to mitigate this issue. We conclude this paper with proposed interdisciplinary mitigations for improving gig worker privacy protections.

[1]Referring to Gray et al. [39], who describe dark patterns as "n-dimensional" phenomena requiring holistic analysis across dimensions like differing disciplinary perspectives, time, and other contexts.

[2]We acknowledge and affirm the need for the "dark patterns" moniker to follow suit with other technical term changes in avoiding words with harmful racial connotations. We do not use the popular alternative "deceptive design" as some designs may be transparent or non-deceptive while still resulting in unintended or negative outcomes for end users. In the absence of similarly popular, community-established alternatives that include non-deceptive (e.g. unfair) dark patterns, we retain the term for this work.

## 2 Background and Related Work

This section provides an overview of prior work on gig work platforms and online privacy analysis, then situates our study within this scholarship. We begin by discussing what we mean by "gig platforms" (based on framing from prior work), and provide context for the privacy risks and harms that gig workers may face.

### 2.1 Gig work platforms

Gig work platforms [14, 47] are systems that connect the supply and demand sides of the gig economy by providing a working "commons", whether that be 1:1 algorithmic matchmaking (e.g., rideshare platforms) or by providing a many-to-many marketplace wherein consumers can search through service providers or goods (Etsy, AirBnB, Fiverr, etc.) As such, platforms exert considerable and unique influence over the worker experience, including what data must be shared with the platform to work, as well as other requirements to participate such as background checks. In this paper, we refer to the supply side of the market as *gig workers* and the demand side as *users/consumers*. For example, in a ridesharing platform the driver is the gig worker and the person requesting the ride is the user/consumer. This power dynamic exposes gig workers to potential harms, for example to workers' autonomy, job satisfaction, safety, inequality, well-being, and more [51, 57, 59].

Prior work analyzed different types of gig work platforms. Farooqi et al. studied pay-per-install applications on Android at scale from the perspective of developers and adoption of such systems [27]. Other studies have explored other data asymmetries, such as those between users and workers, on platforms like Amazon Mechanical Turk [50], and Uber [72], and labor laws governing crowdwork [30]. The closest related study to ours [74] explored privacy and power dynamics in gig work platforms from a worker's perspective. Sannon et al. [74] inspected workers' posts on Reddit regarding four work categories (crowdwork, freelancing, ridesharing, and delivery), finding that workers' privacy concerns span not only platform surveillance but risks posed by customers as well.

In summary, prior works have studied gig work platforms but they have often focused on a select few platforms [45, 80]. While Sannon et al. pursued a cross-platform analysis, they did so through workers' accounts[74]. Therefore, we build upon prior scholarship by studying 16 platforms focused on privacy threats present in different areas of worker-platform interaction including: web, mobile, SMS, and email modalities.

### 2.2 Platform privacy

User-facing tech platforms often encompass multiple modalities like web or mobile, both of which have been revealed to contain privacy risks for end users. Extensive prior work revealed problems with

user data exposure on the web, including web tracking [9, 25, 81, 86] and fingerprinting [10, 35, 49]. Similarly, in mobile apps, network traffic studies tied PII flows to advertisement and tracking companies [55, 68, 69, 82]. Other studies demonstrated the prevalence of tracking in apps, regardless of origin [16, 34], purpose [66] or target audience [29, 40, 70]. Privacy scholarship focused on specific platform types also reveal these risks in more "niche" services. For example, work on childcare apps [40] revealed that 40 of the 42 studied Android apps embedded and shared data with third parties. Such findings are in line with other privacy analyses of Android applications targeting children [29, 70]. Vinayaga-Sureshkanth et al. analyzed the privacy risks of e-scooter rental apps [83] and find these apps rely on a dangerous set of permissions to work and that they often contain a significant number of third-parties (a median of 8 per app).

A differentiating feature of gig work platforms is that, in order to participate on these platforms, gig workers must provide various PII including: legal names, home addresses, phone numbers, etc. In many cases, such collection is necessary to allow a gig worker to provide services on the platform, e.g., for background checks. This proves to be a deciding factor in differentiating our study from prior work. Most works on online privacy focus on platforms, websites, and services where disclosure of authentic personal data is optional. As such, prior studies often use fictional accounts and cannot observe how a wide range of highly sensitive and personal data about consumers is shared. Therefore, while we build off analysis techniques widely employed in the literature, our study fills an important gap by focusing on cases of mandatory authentic information disclosure. Mandatory disclosure poses particular risk for gig workers and potentially violates the privacy-enhancing principles of data minimization and purpose limitation. Similar to prior work, we identify cases of data-sharing for secondary purposes (e.g., monetization through ads, tracking or analytics). The key difference is the highly sensitive and authentic nature of the information being shared, including phone numbers and SSNs.

## 2.3 Dark patterns

Privacy is additionally implicated in platforms' front-end user experiences, as demonstrated by a growing body of privacy dark patterns scholarship. For example, Bösch et al. [17] provided a popular privacy-oriented taxonomy of dark patterns based on Hoepman [46]'s privacy design strategies, while Gray et al. [39] used dark patterns as a lens by which to critique cookie banner design (note that several studies pertain to consent regimes, e.g. [42, 43, 54, 61, 63, 78]). Waldman [84] and Susser et al. [79] discussed dark patterns and manipulation insofar as they may result in poor privacy outcomes. In fact, privacy and data protections regulations have been championed as early dark patterns enforcement tools, with the GDPR and CCPA being notable examples [11, 12, 52, 64].

Prior dark patterns studies have tended to investigate platforms more horizontally, e.g. targeting entire modalities [23, 41], industries [60], a platform category Schaffner et al. [75], or by inspecting consent as noted in the prior section. Other studies had more vertical scope, focusing on exclusively on single platforms like Facebook [62]. Our work exclusively focuses on the privacy impact of dark patterns. We differ from prior work on dark patterns by focusing exclusively on gig work and situating dark patterns within a user's broader privacy experience in-platform. To our knowledge, our work is the first to inspect gig platforms' privacy dark patterns in this manner, and is motivated by an Organization for Economic Co-operation and Development (OECD) report citing gig workers as a group of consumers that may be disproportionately affected by dark patterns [31].

# 3 Gig Work Platforms Dataset Curation

In this section, we provide our definition of a gig work platform of interest for this study, describe how we collect potential candidates, and how we filter them to get the final set of 16 platforms that we test in our study.

## 3.1 Definition and Selection Criteria

Our goal is to understand privacy risks of gig work platforms for their gig workers. Given finite resources to conduct this study, we focus on platforms where we expect *a priori* that risks will be high; namely, those where workers must interact frequently or over short durations while performing immediate, real-time work (exposing them to real-time monitoring for location and other behavior), and correspondingly require the use of a mobile app (which can expose geolocations, unique identifiers, and other PII to gig work platforms and third parties they include via SDKs). Last, we focus on popular platforms that have a relatively large number of workers in order to identify data practices that potentially harm a large number of people.

Based on these goals, we use the following platform-selection criteria. First, they have to meet our definition of a gig work platform, i.e., they act as intermediaries that facilitate a way for users to find a worker for a given task. More specifically, they allow service-side workers (i.e., gig workers) to earn wages from demand-side users.

Next, to meet the "real-time work" criteria, we choose platforms in which workers interact with the platform on a daily (or more frequent) basis. Thus we consider platforms where the duration of work is under one day of work (8 hours) as these typically entail repeated interactions with the platform for additional work, and thus pose continuous data risks to the workers. Further, platforms that allow longer-term work (multiple days/months) tend to operate as contractor matching services, and contract-based work (e.g., job boards, storefronts, etc.) is out of scope of this study. We also exclude platforms that allow flexible schedule employment as workers in such settings fall under a different set of labor/protection laws that are well established.

Finally, we include only platforms that rely on a mobile app with large numbers of workers to carry out the labor. We are specifically interested in this aspect of gig work as mobile apps provide access to a number of sensors (e.g., GPS) and other PII (e.g., unique identifiers). Mobile apps are known to access such data for secondary purposes, and we are interested in understanding whether data collection for secondary purposes is a requisite for workers to earn money on these platforms. We identify apps with large numbers of workers by using the number of app downloads (as reported by the corresponding app store) as a rough proxy.

## 3.2 Search Results

To obtain a dataset of relevant platforms to study, we query related terms in popular search engines Google and Bing/DuckDuckGo. The terms include: “find work", “find gig work", “gig work", “gig economy companies", and similar combinations of the words “gig", “app", “platforms" and “worker". From the results of these queries, we construct a list of potential candidate platforms to further explore. We further include any search-engine-generated lists (e.g., the Google "knowledge panel") of platforms. Most links are to blog posts, crowdsourced websites (e.g., Wikipedia), and news articles; we found 71 platforms in this step, before filtering as described below.

## 3.3 Filtering

For each of these potential gig work platforms, we visit their website to determine whether they fit our study criteria. Since this study is performed in the U.S., we exclude platforms that are intended for use outside the U.S., and platforms that do not fit our definition of a gig platform (§3.1). We also exclude platforms that involve passive rental of owned physical property (e.g., a vehicle or lodging) and semi-public disclosure of the same (e.g., listing the permanent location of said property) in order to be considered ready-to-work; these include platforms such as AirBnb, Turo, etc. These were excluded to protect the geographic privacy of the researchers and their property.

Gig work platforms often require workers to use a mobile app to participate on the platform; these apps are also vital to understanding data practices as they can collect extensive data about their users. Thus, we exclude platforms that do not have an app on the Google Play Store. We pick the Play Store as it is the largest app store and the Android ecosystem has the most established resources for empirical analysis. We acknowledge that the same apps may behave differently in alternative app stores, and so our analysis of the Play Store provides a lower bound of app-based data collection. Further, the Play Store also provides a download count which we use as a lower-bounds proxy for the number of workers on the platform; to focus on large platforms we include apps that have been downloaded a minimum of 500,000 times.[3]

Our above filter criteria can lead to certain types of gig work (e.g., ridesharing and delivery) being overrepresented in our dataset. To mitigate this and cover a more diverse set of gig work platforms, we augment under-represented industry categories in our corpus with additional, slightly less popular, platforms. For example, the pet care industry category initially contained only one platform (Rover), so we included another platform (Wag!) which is listed as having only 50k+ installs.

After filtering, we include 16 platforms for analysis. This covers a diverse set of platforms and categories including ridesharing, food delivery, grocery delivery, personal services, and more. We provide a complete list in Table 1.

[3]Note that some of the apps are the same for workers and consumers/general users. The large app downloads requirement thus acts as a proxy for the number of gig workers, with the assumption that it is correlated with how many users are supported by gig workers on the platform. Secondly, given iOS operational constraints for our experiments and lack of download/installation numbers, we do not include Apple users of these gig work platforms. Thus we caution that the Play Store popularity measure is a partial proxy and lower-bounds estimate of the number of workers on a given platform.

# 4 Methodology

To understand the potential privacy harms that gig work platforms pose to workers, we sign up and participate as workers on the platform. In this section, we describe the methods and infrastructure used for this purpose and provide details about each of our tests.

## 4.1 Test Infrastructure

Drawing from prior work, we rely on network traffic analysis to understand the data collection practices of platforms. Specifically, we seek to understand what PII—including, but not limited to, birthday, email ID, phone number—is shared from the user devices to other parties over the Internet. To this end, we use a multifaceted approach where we investigate data practices when interacting with platforms via a web browser, mobile app, SMS, and email. This section describes our experimental infrastructure and methods for collecting such information.

*4.1.1 User Personas.* During the tests we use different “personas”. Each persona is configured with PII such as a birthday, email address, phone number, and physical address. The email addresses and phone numbers that we use are valid, and receive emails and text/calls respectively. We use two types of personas: fictional and legitimate. Fictional personas are throw-away personas used for certain tests where we do not go through additional identity verification checks. Legitimate personas use two authors’ real information (including SSNs, legal names, addresses, etc.) for the purpose of successfully passing background checks and approvals for working, but otherwise use fresh email addresses and phone numbers, to prevent contamination with authors’ pre-existing personal accounts.

All personas, regardless of type, were created and used between September 2023 and February 2024 and associated email addresses and phone numbers were kept active for at least one month following their creation. During this time, the mailboxes receive emails, and the phone numbers receive both phone calls and text messages. We collect a copy of each type of communication; emails are retrieved in .mbox format via Google Takeout [6], and calls and texts are extracted as JSON files through a custom-built Android app.

*4.1.2 Web Infrastructure.* To understand the data practices of platforms when a worker accesses the platform’s website, we rely on a web browser. Particularly, we use the Google Chrome browser (multiple versions released in 2023) running on a MacBook to visit platform websites and interact with them. We use Chrome’s DevTools [2] to observe the types of PII that platforms collect and the destinations that this data is sent to. DevTools allows us to collect all HTTP traffic[4] generated during our interactions with these platforms in the form of HAR files. To prevent contamination between successive platform tests using the same browser, we clear all data stored on the browser (cache, cookies, logins, etc.) before each test.

Additionally, we collect screen recordings using MacOS’s default recording tool. This allows us to identify dark patterns in user interfaces. It further enables us to correlate observed data sharing with web page context and disclosures.

[4]Note that as a TLS endpoint, Chrome DevTools provides access to decrypted HTTP requests and responses without needing TLS interception.

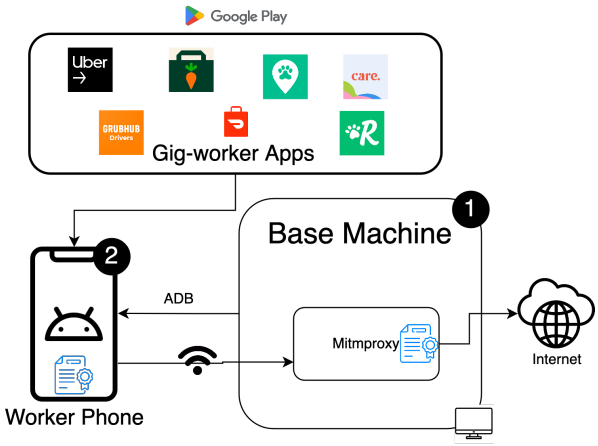



**Figure 1: Overview of setup used for testing gig work apps on Android phones.**

*4.1.3 Android Infrastructure.* To understand the PII a platform collects and dark design patterns platforms apps exhibit, we present our Android infrastructure (Figure 1). We install Android apps from the Google Play Store on a worker phone which is connected to the Internet via a base machine as described below.

***Base Machine*:** The base machine ( Figure 1 ❶) is a Linux machine[5] that acts as a Wi-Fi hotspot and gateway between the worker phone and the Internet. All traffic on this hotspot is forwarded to an instance of mitmproxy [7] running on the same machine. This enables us to look at all HTTP and HTTPS traffic a platform app generates in the form of mitmproxy dump files. Lastly, the base machine is connected to the worker phone via Android Debug Bridge (adb) [1] to collect screen recordings during tests.

***Worker Phone*:** The worker phone ( Figure 1 ❷) connects to the Internet via the base machine's Wi-Fi hotspot. We run only one app on a phone at any one time by killing all other running applications to ensure that all traffic we collect is associated to the test app. Installing apps from the Google Play Store requires the phone to be logged in to a Google account; for this, we use the persona's email address. Last, for mitmproxy to successfully intercept all HTTPS traffic, we add a custom certificate to each worker phone's root store. We use two rooted Pixel 3 Android smartphones running Android 11 as worker phones.

**Certificate Pinning:** Certificate Pinning is a known technique [67] by which an app accepts only a limited number of app-specified TLS certificates for HTTPS connections, as opposed to using the device's root store. Such techniques prevent us from collecting HTTPS traffic with mitmproxy as described above. To address this, we use Frida [4] and a pinning circumvention script (from Pradeep et al. [67]) which disables pinning checks via the adb interface between the base machine and worker phone, and allows us to collect pinned HTTPS traffic with mitmproxy.

[5]We use either a desktop or a Raspberry Pi connected to a standalone power source (power bank), enabling the setup to be portable for certain tests.

## 4.2 Tests Performed

We now discuss how we conduct experiments on our infrastructure. In our experiments, we interact with each platform as both prospective and approved gig workers; all tests we conduct are performed in the U.S., following all applicable regulations.

Note that we concentrate on secondary data sharing practices that are not required for the gig workers' immediate participation on a platform. We especially focus on the following PII: name, home address, birthday, phone number, email address, location (GPS), and social security number (SSN) where applicable. Gig workers can access the platform through different modalities; we look for said PII in both web and mobile app interactions between worker and platform. Mobile apps introduce further platform-specific, persistent PII such as Android Ad IDs [5]. We present the different tests run on each platform in Table 1, split by type of worker-consumer pairing (marketplace or matchmaking)

*4.2.1 Experiment 1: Registration Attempts with Fictional Personas.* To understand the data requirements of participation on a platform, and potential dark patterns that might exist, even before a functional account is obtained, we perform sign ups with fictional personas. We use multiple personas and sign up as both a worker on both web and mobile.

For web, we visit each platform website (N=16), navigate to the worker registration form, and attempt to complete the sign-up process as far as the website would permit. If presented with Privacy Policies or Terms of Services, we accept them; in all cases we notice that we are unable to proceed further with registration if we do not accept these terms. In some cases, platforms require us to download an app, in this case, we stop. For mobile, we download each platform app (N=16) from the Google Play Store, and attempt to complete the sign up process via the app.

Across modalities, we abort sign ups if asked to complete background checks, provide valid IDs, provide banking details like credit cards or routing numbers, or provide tax information. Providing these would require verifiable documentation or information that we do not fabricate for our fictional personas. Lastly, some platforms are locked for our region, due to a saturated market of workers; we are unable to test these further. This approach resulted in zero working accounts with our fictional persona, provided limited insight into platform behavior, and motivates the need for using legitimate personas (unlike most prior work).

*4.2.2 Experiment 2: Registration Attempts with Verified Personas.* One goal of our study is to understand the data collection involved when workers perform tasks on these platforms. To perform that test, we need legitimate personas with verifiable information to obtain fully valid, "ready-to-work" accounts we can use for tests. To this end, we create new accounts using two legitimate, verifiable personas associated with two authors of this paper; these can be used to perform gig work on platforms. When required, the authors submit their real documentation as required to pass background checks and fully register as valid gig workers (such as Social Security Numbers (SSNs), driver's licenses, etc.). We use fresh email addresses and phone numbers to prevent linking these

with authors' personal accounts.[6] We elaborate more on our ethics and compliance considerations in § 4.4. Last, we pay for whatever verification is necessary, as some platforms charge to verify (N=4) identities while others do so for free (N=4). We are unable to obtain ready-to-work accounts for some due to location locks (N=3). Some platforms locked us out of our accounts for unknown reasons (N=2), while others unexpectedly required payment to list worker availability or bid on gigs (which was out of scope for this paper) (N=2). Lastly, N=1 platform allows us to sign up with no verification necessary but does not provide a working account.

We perform sign ups on both web and mobile similar to our initial registration attempts, following the same steps and accepting terms when necessary. In some cases, we are required to switch platforms to perform some verification step; we do so when necessary and collect all associated data.

*4.2.3 Experiment 3: Gig Completion Attempts with Verified Personas.* To understand the data sharing practices of gig work platforms when a worker performs tasks, we participate as gig workers on 4 platforms with legitimate worker accounts. We observe that regardless of what modality a worker uses to sign up, 9 require a mobile app to perform work while 7 allow workers to use both modalities.

Depending on the platform, these tests differ; for rideshare platforms, we pose as clients using fictional user accounts and request rides while waiting for rides using our legitimate worker account. We do so to prevent collecting data of other users on the platform who have not consented to participating in this study. Overall, we have two types of gig work platforms in our study; marketplace and matchmaking. Marketplace platforms are those where a customer can view all workers in a particular area and pick them to complete a gig. These are easy to test as we can pick our worker accounts directly. Matchmaking platforms on the other hand determine which user is matched with which worker, making it harder to match with our worker account. These platforms are of two categories: Delivery and Rideshare; in the case of Rideshare, we control for this by driving to a remote location where we repeatedly ask for rides to match successfully. We thus attempt gigs on 2 platforms in this experiment.

## 4.3 Manual Content Analysis

In this section we describe our manual content analysis and qualitative data coding [73] methods for live (recorded) interaction data and communications received from our gig platforms.

Given the small number of platforms, two authors first coded each platform (using all available screen recordings) for the requirements asked of workers during registration or gig completion interactions as a set of *a priori* codes (provided in Table 8). That is, we look for things workers must disclose (e.g., PII and other types of information) or do (e.g., create accounts, download apps, etc.) to successfully register for or access work on a platform, as these constitute potential privacy/information, labor, or other costs for prospective gig workers—as well as potential *Forced Action*$^{H}$ dark patterns. (In this paper, we superscript dark patterns with the letter corresponding to their granularity level in the Gray et al. [38] ontology; $^{H}$ for High – general dark pattern strategies (also used as broader dark pattern categories), $^{M}$ for Meso – context-agnostic methods that describe how users' interaction expectations are subverted, and $^{L}$ for Low – situated, contextually-dependent means of dark pattern execution.) We use spreadsheets to manage codes and author notes, and resolve any discrepancies in discussions between the two authors towards full consensus. Next we use these codes as a baseline for developing a dark patterns codebook, creating a list of relevant patterns from the Bösch et al. [17] privacy taxonomy and *forced actions* from the Gray et al. [37, 38] categories of dark patterns. We conduct a second coding round using this codebook and again discuss new codes towards consensus.

For analyzing off-platform communications content, we first develop a preliminary set of codes based on potential nudges or nags, according to the perceived immediate purpose of the content: to call users to encourage them to complete registration or work actions, or market customer-side services to the worker persona. To facilitate human labeling of messaging content and reduce redundancy, we first manually inspect our fictional and verified personas Gmail inboxes to record the binary presence of each purpose within each platform's subset of messages. We view subject lines directly in the Gmail interface rather extracting them from *.mbox* dumps in order to correctly view encoded items (such as emojis or other special characters) in subject lines, access email preview text for interpreting ambiguous subject lines, and review emails as presented in-the-wild. As SMS messages were more uniform in format and comparatively free of encoding issues, we review SMS directly from JSON dumps. We again use spreadsheets to manage codes and resolve discrepancies in discussions for full agreement.

Our final set of observed dark patterns in the gig platform experience (including patterns later evidenced through platforms' data sharing practices) is presented with descriptions in Table 9.

**Table 1: Platforms in our study marked with the tests we perform and split by type (Marketplace and Matchmaking respectively).**

| Platform | Category | Verified reg. | Gig |
|---|---|---|---|
| Airtasker | Tasks | ✔ | |
| Care | Caregiving | ✔ | |
| Fiverr | Tasks | ✔ | |
| Peopleperhour | Tasks | ✔ | |
| Rover | Caregiving | ✔ | ✔ |
| Taskrabbit | Tasks | ✔ | |
| Thumbtack | Tasks | ✔ | |
| Upwork | Tasks | ✔ | |
| Urbansitter | Caregiving | ✔ | |
| Wag | Caregiving | ✔ | ✔ |
| Doordash | Delivery | ✔ | |
| Grubhub | Delivery | ✔ | |
| Instacart | Delivery | | |
| Lyft | Rideshare | ✔ | ✔ |
| Shipt | Delivery | ✔ | |
| Uber | Rideshare | ✔ | ✔ |

[6] We acquired and completed any additional paperwork requested by the platform or local regulations where possible (e.g., driver history records or local supplementary car inspections).

### 4.4 Ethical considerations

By nature, any labor-related interactions conducted in our gig platforms involve real users when matching clients to workers. Similarly, by conducting live tests, our experiments are subject to the policies of each platform. We now discuss how we addressed ethical considerations in our study.

All registrations were completed in compliance with local regulations, including passing and obtaining mandatory vehicle inspections specific to rideshare work in our region. We additionally reviewed platform policies and conducted all interactions according to these policies, taking special care to adhere to user terms of service. We did not modify or otherwise interfere with the execution of any work or service requests made through the platforms, did not modify the platform websites or apps themselves, and appropriately sent payment for any gigs successfully completed through the platforms. While completing gigs, we faithfully self-assigned tasks to the best of our ability as facilitated by the platforms and reject all assignments with non-author users. Further, any non-author data collected during our tests (such as client information during the matching process) is promptly deleted, and any author PII is deleted after tests and analysis.

Any verifiable information or official documentation was provided to platforms solely for the purposes of truthfully completing registration, passing any background or safety checks, or providing payment as required by a platform. These were provided in compliance with all parties involved with processing background checks to the best of our knowledge.

As we will further discuss in the next section, we identified unintended disclosures of reversible SSN hashes to third parties. Given the extremely sensitive nature of such data, we disclosed this to the parties involved and engaged in discussions with relevant stakeholders to facilitate remediation. At the time of writing, the observed cases of SSN disclosure to third parties have been removed from the affected websites.

## 5 Data Sharing Practices in Gig Platforms

This section presents our findings on data sharing practices gleaned from mobile app permission analysis and network traffic captures while testing the 16 platforms described in the previous sections. We further consider how tracking is conducted when workers open emails sent to them by gig work platforms.

As part of the sign-up process, gig workers are expected to provide various types of PII that include names, addresses, and Social Security Numbers (SSNs). Understanding what platforms do with PII is therefore important in understanding the data risks faced by workers on these platforms. To this end, we analyze the network traffic data recorded during our sign-up process described in § 4.1.

### 5.1 PII Sharing Practices

We begin by noting that gig work platforms collect PII from workers as a *necessary* part of the sign-up process—this is required for tax compliance, platform safety, and many other reasons. However, when such data is opaquely shared with third parties, it is often neither necessary to perform the duties of a worker nor expected. Furthermore, such data collection is arguably even more unavoidable for workers, who rely on the platforms to make money and have strong incentives to accept whatever the platform imposes.

We find PII sharing with third parties by gig work platforms is *pervasive*: 15 of the 16 platforms that we study share at least one form of PII with one or more third parties across modalities. Table 2 lists platforms sorted by the total number of third-parties they send PII to along with the types of PII they send. We find that Advertisement IDs (Ad ID) are shared by every platform that shares any PII.[7] Names and emails are the next most commonly shared PII, by 14 and 12 respectively.

We next analyze the set of third parties that receive worker PII, finding that gig work platforms send PII to a large set of destinations serving a wide range of purposes. Table 3 lists the third parties receiving PII from the largest number of platforms; the complete dataset is listed in Appendix A. Grouping these parties by their purposes, we see that Web APIs are the most popular (15), followed by analytics (14), social networks (13), and advertising (6). Most of these endpoints are not required for gig workers to sign up, and thus represent unnecessary privacy risks for workers. The fact that social media and advertisers appear in the list is concerning, as it constitutes an additional form of exploitation (monetizing worker data via targeted advertising) for a vulnerable workforce.

While collectively there is extensive sharing with third parties, we find that even individual platforms share PII with multiple third parties. In fact, all data sharing platforms share data with multiple third-parties, and some platforms (Fiverr and DoorDash) do so with more than 20 third parties. We observe that both hashed and unhashed PII is sent to these third parties; we must note here that in some cases (names, phone numbers, SSNs) hashing does not provide any significant privacy protections. Hashing is only useful for privacy when the number of possible inputs is high enough to prevent computing a dictionary of all hashes— in this case, it is not considering the fields' relatively small numbers and commonality of names.

**Social Security Numbers:** Perhaps the most alarming PII sharing came from two platforms (Lyft and Uber), which shared reversible hashes of highly sensitive Social Security Numbers (SSNs) with third-parties. In the US, the SSN is a 9 digit identifier that is permanently linked to a person. Exposure of this governmental identifier can lead to serious risks to the user, including identity theft [77]. Lyft shared unsalted hashes of worker SSNs with Facebook and TikTok, while Uber shared unsalted hashes of SSNs with Facebook. Upon further manual analysis of the requests responsible for this sharing, we find that tracking pixels found on sign-up web pages are responsible. We find no evidence of SSN collection on the mobile app counterparts of these platforms.

The SSNs were transmitted as an unsalted SHA256 hash of the worker's SSN. One might argue that this level of obfuscation offers privacy because the 9-digit SSN was not itself exposed. However, this would be a naïve interpretation: given the relatively small size of all possible SSNs (all 9-digit numbers), calculating a mapping between every possible SSN and its corresponding hash is trivial. In fact, we built such a mapping of the entire SSN space in less than 2 minutes on a commodity 32-core server. As such, the hashing

[7] Android documentation states that an Ad ID is a "unique, user-resettable ID for advertising".

**Table 2: Total number of third-parties (grouped by second level domains) every platform sends PII to. Parentheses indicate PII sent without hashing. Lyft and Uber share unsalted hashes of SSNs with third parties (marked in red).**

| Platform | Ad. ID | Name | Email | Phone | Loc. | Add. | Total |
|---|---|---|---|---|---|---|---|
| Fiverr | (2) | (27) | 5 | 2 | | 1 | 30 |
| DoorDash | (4) | (21) | 7 (4) | (18) | (1) | (1) | 23 |
| Rover | (5) | 15 (14) | 5 (3) | | (2) | 2 (1) | 18 |
| Wag | (4) | (9) | 8 (7) | (1) | (1) | | 11 |
| Care | (5) | 3 (2) | 7 (1) | | | | 10 |
| Thumbtack | (4) | 4 (3) | 6 (2) | 3 (1) | (2) | | 9 |
| Upwork | (2) | (8) | 4 (2) | | (1) | | 9 |
| Lyft | (2) | 3 (2) | 5 (1) | 5 (1) | | 1 | 7 |
| Taskrabbit | (4) | 4 (2) | 4 (1) | 2 | | | 7 |
| Airtasker | (4) | 3 (2) | (1) | 1 | | (2) | 6 |
| Grubhub | (2) | 3 (2) | 3 (2) | 2 (1) | (2) | | 6 |
| Urbansitter | (6) | | | | (2) | (1) | 6 |
| Uber | (1) | (1) | | | | 1 | 4 |
| Instacart | (2) | 1 | 2 | 2 | | | 3 |
| Shipt | (2) | (1) | | | | | 3 |
| Total | (15) | 14 (13) | 12 (10) | 9 (5) | (7) | 7 (4) | 15 |

used to obfuscate SSNs provides no practical protection for users because it is trivially reversible.

Based on our manual analysis network traffic from the TikTok and Meta Pixel trackers, the reason SSN hashes were shared is that the trackers were incorrectly interpreting SSNs as phone numbers (even though they have different numbers of digits in the US).[8] Further, Lyft and Uber configured these trackers to collect information from forms, which was optional and arguably should have been disabled. Given the severity of this data sharing practice, we manually checked the privacy policies of these platforms and confirmed that such collection was not disclosed in the privacy policy of either Lyft or Uber.

We disclosed this finding to the relevant parties (Lyft, Uber, Facebook, and TikTok) to facilitate remediation and (we hope) disgorgement the data. Lyft and Uber both removed the sharing of SSN hashes from their websites after discussions with our team. Meta stated that data transmitted by the Pixel is not recorded on Meta servers, a claim that we cannot independently verify. They also claim that there are filters that should prevent accidental SSN data collection; however, we found that these filters did not work as intended for Uber and Lyft. This is because the filter looks for field names matching "ssn" but this is not always the name of the SSN field. For example, Lyft used "ssn_validation" for SSNs. TikTok could not reproduce our finding (because Lyft had removed the tracker after our disclosure) and claims that their tracking code will now prevent SSN collection—something we were not able to independently verify due to the removal of tracking by Lyft.

**Web vs App PII Sharing:** An important question is whether gig workers are exposed to different PII sharing over one modality versus another, and what are the implications for privacy across modality. To compare the data sharing practices of platforms on websites and apps, we present Jaccard similarities of PII types and

[8] In observed network traffic, the name of the field containing the hashed SSN is an abbreviation for phone number (e.g., "ph").

**Table 3: Companies that receive PII from at least 2 gig work platforms. Parentheses indicate PII sent without hashing. Red indicates companies that receive unsalted hashes of SSNs.**

| Domain | Ad. ID | Name | Email | Phone | Loc. | Add. | Total |
|---|---|---|---|---|---|---|---|
| Google - API | (11) | (11) | (6) | | (4) | (2) | 15 |
| Facebook | (8) | 11 (3) | 10 | 8 (1) | | | 13 |
| Google - Se. | (3) | (6) | 4 (2) | 3 (2) | (3) | | 10 |
| TikTok | | (2) | 7 | 6 (1) | | 1 | 8 |
| Firebase | (6) | | | | | | 6 |
| Doubleclick | | (5) | 2 (1) | (1) | | | 6 |
| Google - An. | | (5) | 2 (1) | 2 (1) | | | 6 |
| Branch | (6) | (1) | | | | | 6 |
| Amplitude | (2) | (3) | | (1) | (1) | | 4 |
| Twilio Segment | (3) | (3) | (3) | (2) | | | 4 |
| Google - Ads | | (3) | 2 (1) | (1) | | | 4 |
| Bing | | (4) | | (1) | | | 4 |
| Linkedin | | (3) | (1) | (1) | | | 3 |
| AppsFlyer | (3) | | | | | | 3 |
| NewRelic | | (3) | (2) | (1) | (1) | | 3 |
| Impact | | (1) | 1 | | | | 2 |
| Snapchat | | (1) | 2 | (1) | | | 2 |
| Adjust | (2) | | | | | | 2 |
| CloudFlare | | (2) | (1) | (1) | | | 2 |
| Twitter | | (2) | | (1) | | | 2 |
| Yahoo | | (2) | 1 | (1) | | | 2 |
| Onetrust | | | | | | 2 | 2 |
| Pinterest | | (1) | 2 | | | | 2 |
| theTradeDesk | | (2) | | | | | 2 |
| Braze | (1) | (1) | 1 | | | | 2 |
| Stripe | | (2) | | | | (1) | 2 |
| Iterable | (2) | (1) | (1) | | | | 2 |

PII type–3rd party pairs in Table 4. We also present the number of unique PII types shared with third parties on web and app; we see that a majority of platforms (7) share more PII with third parties on their apps while 3 platforms share more on their websites. This reconfirms that looking at both modalities in combination is crucial to understand the overall privacy impact a worker faces while participating on these online platforms. Further, this mixed picture means that gig workers cannot rely on one modality versus another to reduce their privacy exposure.

## 5.2 Permission analysis

While network traffic analysis provides hard evidence of PII exposed by platforms, runtime analysis of apps represents a lower bound as the behaviors observed are bound to the tests performed (which do not cover all possible app behaviors). For mobile apps, a complementary approach is to analyze the Android permissions requested by apps, which can be an estimate of the type of data that apps have access to at runtime. To that end, we use static analysis to parse the permissions used by each app and categorize according to Android's official documentation [22] (i.e., normal, dangerous and custom permissions).

When looking at differences between worker and user apps, we find that, on average user apps request fewer permissions (33 vs 35). When focusing on dangerous permissions, which give apps access to restricted data or features, we see that the differences on average are slightly smaller (7 by user apps and 8 by worker apps). Android apps

**Table 4: Platforms in our study sorted by Jaccard similarities of PII type and PII-type/third-party pairs, along with number of unique PII types shared per modality. We observe low similarity between data sharing practices on platform websites and mobile apps.**

| Platform | PII | PII-3rd | # Web | # App |
|---|---|---|---|---|
| DoorDash | 0.5 | 0.69 | 3 | 6 |
| Thumbtack | 0.6 | 0.25 | 3 | 5 |
| Wag | 0.6 | 0.16 | 3 | 5 |
| Rover | 0.6 | 0.1 | 3 | 5 |
| Care | 0.67 | 0.0 | 2 | 3 |
| Taskrabbit | 0.5 | 0.14 | 3 | 3 |
| Grubhub | 0.6 | 0.0 | 3 | 5 |
| Upwork | 0.5 | 0.0 | 2 | 4 |
| Fiverr | 0.4 | 0.05 | 4 | 3 |
| Airtasker | 0.2 | 0.0 | 3 | 3 |
| Lyft | 0.17 | 0.0 | 5 | 2 |
| Uber | 0.0 | 0.0 | 2 | 2 |
| Instacart | 0.0 | 0.0 | 3 | 1 |

**Table 5: Platforms in our study sorted by Jaccard similarities of PII type and PII-type/third-party pairs, along with number of unique PII types shared between worker and customer apps. We observe a high similarity, and 15/16 platforms share more worker data than customer data.**

| Platform | PII | PII-3rd | # Worker | # Customer |
|---|---|---|---|---|
| Lyft | 1.0 | 0.67 | 2 | 2 |
| Fiverr | 0.67 | 0.6 | 3 | 2 |
| Care | 0.67 | 0.5 | 3 | 2 |
| DoorDash | 0.83 | 0.33 | 6 | 5 |
| Taskrabbit | 0.67 | 0.43 | 3 | 2 |
| Wag | 0.67 | 0.35 | 5 | 5 |
| Urbansitter | 0.67 | 0.33 | 3 | 2 |
| Rover | 0.6 | 0.37 | 5 | 3 |
| Airtasker | 0.33 | 0.5 | 3 | 1 |
| Uber | 0.5 | 0.25 | 2 | 1 |
| Thumbtack | 0.4 | 0.28 | 5 | 2 |
| Instacart | 0.5 | 0.14 | 1 | 2 |
| Grubhub | 0.4 | 0.18 | 5 | 2 |
| Shipt | 0.0 | 0.0 | 2 | 0 |

are known to be permission hungry, and as such, it is not surprising that most of the user apps access the same set of permissions as the worker apps. However, we still find some exceptions in which the differences across app type are clearer. Namely, the Shopper app (5 vs 9 dangerous permissions), Taskrabbit (7 vs 10) and Grubhub (6 to 9). Although the differences across app types are not significant, we find a large number of custom permissions (9 in average for both types). Custom permissions are declared by apps and are often undocumented, which makes it harder to attribute them to a specific data type or feature [33, 56]. As previous work reported, the most common permissions appear to be related to Google services (e.g., Google Cloud-to-Device messaging and Google Mobile Services).

## 5.3 Email Data Sharing Practices

In addition to apps or websites, e-mails can also contain tracking URLs [24]. Though we primarily collect off-platform communications for the purpose of understanding what platforms do with the PII they collect from gig workers at registration, we also inspect emails for tracking behavior. Specifically, we look at all URLs embedded in all emails our worker accounts receive from platforms and match them with Easylist and Easyprivacy [3]. We find that 8 platforms embed trackers in emails; 5 of these include third party trackers while the other 3 include first party trackers only. The former include marketing companies such as Movable Ink (contacted by two platforms), Exact Target (one platform) and Litmus (one platform) and Google (one platform).

We note that, while in this paper we have focused mainly in third-party tracking, the existence of tracking behavior in email communications, even from first parties, is worrisome. This can allow gig work platforms to silently track when and if users read these communications.

## 5.4 Worker vs Consumer Data Sharing

As part of necessary identity checks, workers often have to input more personal information when signing up to the platforms. A key question is whether this translates into more data being collected by these apps during the sign-up process, and/or sharing with third parties. Workers are required include substantial personal information during signup, and we argue that this data being shared with third parties for commercial purposes can be considered unexpected by workers (if not an abuse of the platforms power and position). To that end, we calculate Jaccard similarities of PII types shared with 3rd parties and PII type–3rd party pairs in Table 5 for mobile app worker and customer sign ups.

It is important to note that 5 of the platforms provide the same app for both workers and consumers, while the remaining 11 provide different apps. While the differences are not stark, we still observe that all platforms except 1 (Instacart) share more data types for workers than for customers. More concretely, we find that email is the most popular extra data shared with third parties (N=10), followed by names (N=5), and locations (N=3). Despite some third-party services receiving a certain type of data only from worker accounts, we find no evidence of third parties present only on worker or consumer signups (i.e., the set of third parties remains stable).

# 6 Privacy Dark Patterns in Gig Platforms

This section presents the results of the content analyses described in § 4.3. Specifically, we focus on deceptive user interface designs (i.e., dark patterns) that impact or exploit gig worker privacy.

**Table 6: Table describing the number of platforms in our study for which we noted each dark pattern. As all 16 platforms required some provision of PII at registration, we count only those platforms requiring password creation and thus establishing user accounts for *Forced Registration*$^{M}$, and for *Forced Disclosure*$^{M}$ we count only platforms which mandated additional information beyond PII used for formal background or identity verifications.**

| Dark Pattern | № Platforms |
|---|---|
| Forced Registration$^{M}$ | 14 |
| Forced Modality Switching$^{M}$ | 5 |
| Forced Communication/Disclosure$^{M}$ | 8 |
| Nags$^{H}$ to Register | 4 (Email), 5 (SMS) |
| Nags$^{H}$ to Work | 11 (Email), 6 (SMS) |
| Nags$^{H}$ to Consume | 7 (Email), 0 (SMS) |

## 6.1 Dark Patterns in Worker Platform Interfaces

Here we enumerate the privacy dark patterns we observed based on the requirements for joining gig platforms.

*6.1.1 Forced Registration$^{M}$ Before Work-Readiness.* Prospective workers are not guaranteed to pass platforms' verification checks, but are asked to provide PII in the process of seeking platform approval. Submitting information is necessary, but temporary. Account creation, on the other hand, is more involved and opens up the potential for long-term data relationships between a prospective worker and a gig platform—regardless of whether the worker is approved or not. Our data shows that 14 of our 16 platforms request a password in one or both modalities, thus creating user accounts. The remaining two platforms, Shipt and GrubHub, placed us on regional waitlists to work and appeared to have submitted our application without creating accounts.

On one hand, "placeholding" an account for a user can help workers later finish signing up if they were interrupted during the process for any reason; this is convenient if a potential worker intends to return to the registration process later on, but privacy-disadvantageous if a person decides against gig work partway through their application. When an account is created separately from worker verification, this may also subvert user expectations of what a platform can or will do with already-submitted (or partially-submitted) account creation PII, particularly if a prospective worker abandons the sign-up process partway through. This issue is often further exacerbated by off-platform communications received by our personas in § 6.2.

Platforms' web and mobile app experiences were additionally inconsistent in the extent to which they communicated the separation between workers and consumers at registration, with Airtasker and Rover instructing workers to register for general user accountsprior to selecting worker roles, and five platforms (Airtasker, Fiverrr, PeoplePerHour, Urbansitter, and Rover) offering only one centralized app to both consumers and workers. Such designs force worker-only users to create platform-wide accounts, which blurs the distinctions between workers and consumers and potentially subjects workers to unsolicited marketing or communications unrelated to their immediate purpose in joining the platform (working). We take this lack of separation as evidence of worker misclassification as well as abusive re-purposing of the PII provided to the platform by workers.

*6.1.2 Forced Modality Switching$^{M}$.* Platforms differed in the extent to which mobile apps or multimodal steps were required for registration. Four platforms (Grubhub, Instacart, Shipt, and Taskrabbit) allowed minimal registration from web landing pages, with DoorDash partially requiring the use of a separate mobile device to take identity verification photos (DoorDash clarified that no app download was required to complete this step). All five collected names, email addresses, phone numbers, and general location information via web browser before requiring mobile app use for the remaining registration steps. Though all of these registration attempts constituted partial sign-ups, Doordash and Taskrabbit additionally required password creation and were the only two of the five to send messages to our persona email address. Conversely, Fiverrr was the only platform in our study to require a switch to web in order to complete registration (specifically, to add profile information) started in its mobile app.

We consider this to be *forced modality switching*, which has been noted in prior work [53, 75] for the way this design increases user labor for completing a given task. Forced multi-modal use additionally exposes the user to cross-modality tracking or data exposure, and we find this of particular concern for gig workers who may use personal devices for work conducted on platforms that require real-time app use. Grubhub and Shipt forced us to continue registering via mobile app, but both placed us on a worker waitlist and effectively prevented any further registration. It is unclear why region saturation checks were not conducted in the browser, given that we submitted PII there first before being modality-blocked.

*6.1.3 Forced Communication or Disclosure$^{M}$.* To some extent, the gig work context necessitates certain disclosures or data collection, like verifying identities for background checks or requiring valid driver's licenses for rideshare platforms. This worker information can be used to ensure platform safety for both consumers and workers, improve worker-consumer matches, and otherwise ensure smooth operation of gig work on the platform. Some platforms often required, or seemed to require, much more. For example, 7 platforms required minimum-character-count biographies for worker profiles. Further, task platforms tended to ask for work history or skills for building a worker's profile.

The more information is collected on a worker, the more vulnerable they are to privacy risks. Thus we highlight the forced nature of profile information requests and question whether all platform details are necessary in initial registration flows, as opposed to optional (opt-out) from the start with the opportunity to add more information later.

Of the platforms we were able to successfully complete (pet-sitting) work on (Wag, Rover), both asked us to collect community endorsements at some point during registration and presented these as necessary for our profile as shown in Figure 4 in the Appendix. This behavior implicates *forced disclosure$^{M}$* sub-patterns like *Friend Spam$^{L}$* and *Social Pyramid$^{L}$*, which collect information on other, external users through unwanted contact or by directly recruiting them to the platform. However, neither platform actually checked whether these endorsements were successfully filed or whether the

required number was met (minimum one for Rover and five for Wag) as we did not note any immmediate obstacles to accessing gigs while we had fewer than the required endorsements.

**Shadow User Profiles?** When forced disclosure is accompanied with the data sharing practices in Table 5.4, worker data is additionally exposed to the broader tracking economy. These results relate to unintended data sharing described in the Bösch et al. [17] *Shadow User Profiles*$^{L}$ dark pattern insofar as "affected individuals [may not be] aware of personal data records they have accidentally created." We also note that the author whose SSN was shared with TikTok's tracking pixel did not have a TikTok account, relating to the Bösch et al. [17] discussion of service providers non-transparently collecting records about non-users. As the original intention of this pattern in Bösch et al. [17] focuses on bystander data being shared with platforms through a first-party user, we do not emphasize this dark pattern for our study. However, we provide this discussion to illustrate the dark patterns implications of tracking behaviors which lack up-front transparency, or bury informed consent in overbroad privacy policy disclosures.

## 6.2 Off-Platform Communications

In this section we examine the worker experience of receiving messaging communications from platforms (e.g., calls, texts, and emails). We find highly variable volumes of messaging, representing a wide range of worker communications strategies across platforms. The sheer volume of messages and disparities between platforms or personas suggests the potential of submitted PII being abused to overcommunicate to workers. Such behavior can opportunistically utilize information disclosed to a platform by a worker, and presents a form of *Bad Default* dark patterns for default-on, built-into-registration communications settings.

Within the first two weeks after registering for a platform (starting registration from the website), we received 8 phone calls from Doordash alone, 43 texts from 10 platforms, and 578 emails from all 16 platforms.[9] Platforms that did not send emails to our fictional persona (GrubHub, Instacart, and Shipt) correspond to the platforms that required downloading a mobile app to complete registration. DoorDash additionally sent SMSes appearing to be from real people (meaning that the text body was written in first-person and provided the first name of a platform representative). These calls and texts were received from multiple numbers, but all received texts appeared to be from the same identity. Table 10 presents the average number of messages sent per day, per platform in the first two weeks for direct comparison across personas.

Care.com sent relatively few emails (N=3) to our fictional persona in the first two weeks, at a rate of 0.03 emails per day. Two of these three prompted registration completion ("Did you forget something?" and "Finish enrolling for Care.com now"), with the third marketing their premium subscription service, which includes annual background checks in the listed price ("Premium today, opportunities tomorrow."). Conversely, they sent a rate of 1.21 emails per day (N=17) to our real persona in the first two weeks, with emails now including (some templated) advertisements for jobs in our area and worker advice.

[9]Over the entire duration of our study, we received 1,670 total email items spanning three personas registering for all 16 apps, but this includes longitudinal data from multiple rounds of registration so we normalize inbox counts to the first two weeks rather than all emails received.

From the 1,670 total emails across all worker persona inboxes for the entire duration of the study, we noted 140 unique subject lines after controlling for templates and duplicates, yielding a rate of 91.6% of emails coming from duplicate or templated subjects. To demonstrate the volume of non-unique emails, only 10 of the 1037 total emails sent by Thumbtack for the entire duration of our study across two personas were truly unique and otherwise untemplated. All other Thumbtack emails (N=1027) took some form of the template "<FirstName L.> needs <Task Name> in <City, State>." All of the non-templated items constituted the first 10 consecutive emails sent to our persona. Care's uptick in emails sent (from 3 to our fictional persona in the first two weeks, to 80 to our verified persona's first two weeks) included templates like "Your New Job Match in <City>: Up to $/hr." Uber, on the other hand, was comparatively more "creative" with 51 unique subject lines across 222 total emails. High-volume messaging facilitated by templating or automation exacerbate abuses of submitted PII, even more so for gig workers that may be submitting personal email accounts or phone numbers.

*6.2.1 Types of Nags in Off-Platform Communications.* Across all platforms, 9 platforms urged us to complete registration if we had not done so yet, 11 encouraged us to go work, and 6 wanted to reach us as consumers of the platform's services. Only Shipt did not send call-to-action messages, sending only one email to let us know that we were waitlisted. While we do not relate these to in-platform interface nags (e.g., interruptive pop-ups or interfaces that redirect expected functionality [37]), high-volume, repetitive messaging presents an off-platform relation to nagging dark patterns.

**Nags to Register:** Reminders to complete registration typically described missing documents and used terms like "Action Needed" or "Complete your registration" to convey alerts. While generally useful for a registrant actively seeking to gain worker approval and start completing gigs, nearly all platforms sent multiple reminders, often with duplicate content and at regular intervals. The few exceptions were Shipt and Airtasker for verified personas, and Airtasker, Instacart, Shipt, and Grubhub for our fictional persona. These platforms only sent status information (e.g., informing us that we had uploaded information to the platform or that we would be put on a waitlist) or no emails at all (likely due to registration being cut off by mandatory modality switching); all five sent three or fewer emails when they did not nag us to register.

**Nags to Work:** Thumbtack and Care.com's templated deluge of emails were primarily used to prompt us to work by highlighting seemingly personalized, human-connected opportunities in our markets. We were able to directly link the task variables in Thumbtack's "someone needs <task> done in your area" emails to tasks we selected during registration when prompted for the services our personas intended to provide. Advertising opportunities to workers is not necessarily deceptive or manipulative and can benefit users seeking work, but the sheer volume of these templated messages raise our suspicions on the authenticity of these opportunities. Coupled with a lack of opt-outs during registration, we consider this behavior to be opportunistic and spammy.

**Customer Marketing:** Six platforms (Uber, Lyft, Rover, Taskrabbit, Thumbtack, and Fiverrr) addressed our personas not only as workers, but as consumers of their services. This may be a quirk of

*Bad Defaults*$^{M}$ for notifications, but we argue that these nags affirm the need for clearer gig worker classification (and subsequently stronger employee-like protections). We received consumer-side marketing for UberEats and Uber rides services, as well as Uber One memberships. These marketing emails often included discounts or other promotions. For example, Uber cheerfully promoted its rideshare, food delivery, and subscription services to us with subject lines like "$25 off = FREE meal. Want it? Have it.," "60% off Uber One: our biggest membership sale!," and "Don't stress about how to get there, ride with the Uber app." Thumbtack offered "home inspiration for days" and tips on how to "care for [our] home like a pro," with these two emails being sent by an address specifically targeting consumers: do-not-reply@customer.thumbtack.com. Rover in particular used 26 unique subject lines and sent 40 total emails, with subject lines including content like provocatively vague queries (e.g., "We have a question for you/and your dog..."), dog behavior-specific questions (e.g., "Does your dog ever sigh?") or dog owner advice (e.g., "How to find a great vet for your dog"), general pet content (e.g., "Dog people news: top trending stories"), and direct marketing for Rover's services (e.g., "Book holiday care for your pet" and "Book a sitter [over the long weekend]").

In Lyft's case, we noticed differences in how incomplete registrations were addressed. The personas that did not complete registration (either a fictional persona, or one with verified PII but no driver documentation) began to receive heavily templated customer offers a few weeks after attempted registration, when the weeks prior had focused on sending registration reminders instead. These subject lines told us to "claim [our] offer, <Persona Name>," "claim <NN>% off of [our] first <N> rides, <Persona Name>," or to "try Lyft and earn $<NNN> for <N> rides." Comparatively, our fully-registered persona received only one customer-marketed email ("Party the <FirstName> way with Lyft,"), with the remainder of received emails being operational messages. Considering that gig workers join these platforms for the purpose of earning money and completing work, unsolicited marketing messages constitute privacy dark patterns in their abuse of PII submitted for worker registration.

# 7 Discussion

In this paper, we show how gig work platforms collect and share PII from their workers. Further, their apps often employ privacy dark patterns that amplify data collection, and platforms often take advantage of registration information to nag potential workers. These findings lead us to the following observations and recommendations.

## 7.1 Workers are Distinctly and Inherently Vulnerable

Gig workers are incentivized to be transparent when registering to work on gig platforms. Transparency and disclosure of personal information allows prospective workers to pass background checks necessary for approval, as well as to be desirable candidates for clients' various needs. However, this transparency comes at a privacy cost; the nature of highly-sensitive data collection makes workers distinctly vulnerable to privacy risks as compared to consumers who may be registered to the same platform. To use a rideshare service or hire a gig worker, consumers are not asked to pass background checks or verify their identities to the same extent. Furthermore, extra information required from workers can lead to more data collection by these platforms as shown in § 5.4. Most egregiously we noted Uber and Lyft failing to treat SSNs submitted to web forms differently from changeable PII like emails or phone numbers, sending them to Meta and TikTok.

In terms of communications from these platforms, we take the example of DoorDash which made four calls and two texts to our fictional persona (which only partially registered, as we did not verify identity for this persona). Human-touch outreach may be seen as considerate amidst a sea of highly automated messages, but contacting prospective workers across multiple modes is potentially invasive, regardless of whether platform Terms of Use and privacy policies include these communications in blanket statements. Furthermore, privacy policies (and legal notices) have been proven to be convoluted and hard to understand for the average user [15, 18, 36]. Both DoorDash texts and two calls were sent during working hours, with the other two calls made around 6 PM. As gig workers often use gig work to supplement income from other jobs, calling during the workday is potentially disruptive. Calling after typical work hours, on the other hand, encroaches upon workers' personal time. In either event, such contact from gig platforms *that the applicant doesn't work for (yet)* may constitute unwanted solicitation for the prospective worker and thus encroaches upon the privacy "right to be let alone [85]."

## 7.2 Mitigations Should Make Protections Proportional to Vulnerability

Mitigating workers' privacy harms requires platforms to recognize them as distinctly vulnerable and adjusting their data practices accordingly. If gig platforms do not self-regulate effectively, then legal protections that start with proper classification of gig workers as employees is a start. From there, privacy principles for vulnerable classes must be upheld.

Our results additionally affirm the ubiquity of platforms sharing data with third-party trackers. Though this practice is common and often covered[10] by platform privacy policies, we are not convinced that worker data should be treated in the same as general user data, in no small part due to the inherent vulnerabilities that workers take on at registration. Interestingly, 15 platforms shared worker data with more parties than consumer data—so gig worker data is exposed to more privacy risks than those from consumers. This is particularly concerning because gig workers have limited-to-no choice but to provide highly sensitive data to gig platforms (often more than once) if they want to earn money.

Global privacy regulations [11, 64] already commonly incorporate Fair Information Practice Principles (FIPs), which include limitation principles for data collection and use, as well as purpose specification for that data.

Purpose limitations prevent further data processing beyond "specified, explicit, and legitimate purposes" that workers should be made aware of (beyond clickwrap Terms of Use) when submitting

[10] Debatably covered; usable security and privacy scholarship robustly investigate problems with fine-print clickwrap agreements and often note that technical compliance does not equate to conspicious and transparent disclosure.

that data. This mitigates opportunistic behavior, and would disallow platforms from advertising consumer-side services to worker registrants without informed consent. Purpose limitations may also urge platforms to reconsider their architecture, especially if these limitations are tailored for protecting gig workers from misclassification-related vulnerabilities. In Rover's case, in which petsitters have no option but to use an app or website primarily designed for pet owners, clearer separation of in-platform interfaces may help preserve worker privacy better than the present design.

The GDPR additionally includes a data "right to erasure," which could help mitigate *Immortal Accounts*$^{L}$ resulting from *forced registration*$^{M}$, especially for prospective workers who later decide against joining a platform or for extant workers choosing to "retire."

## 7.3 Limitations and Future Work

**U.S. Scope:** Our study was conducted in the United States, in a region unprotected by strong state privacy laws. As such, some of the interfaces and practices we observed may be contingent on where we ran our experiments; that is, GDPR requirements may already mitigate the behaviors we saw and make for different results for the same experiment conducted in the EU. Comparing across privacy jurisdictions was out of scope for our study, but our work demonstrates the necessity of future work comparing the effects of privacy regulations on resultant user privacy experiences.

**Registration vs. Work Interactions:** Despite our best efforts to collect live data and conduct real-world gigs, we were only able to attempt work on a subset (N=4) of our platforms (and actually complete gigs on even fewer; N=2). This is in part due to technical constraints. For example, DoorDash documentation notes that the app may not function correctly on jailbroken devices, which ours were, and we were unable to "go live" on Lyft with either of our study devices (an external test on a personal, non-jailbroken device confirmed that we were able to "go live" on Lyft). Furthermore, the use of Frida for certificate pinning bypassing could raise anti-analysis flags by these apps, leading to them not working on our test devices [26]. Operational constraints within our methodology also prevent gig completion in some cases. When attempting to drive for Uber, our methods required rejecting all potential rides except ride requests coming from our other test device, but Uber did not connect our driver persona to our rider persona in the hour we waited for an acceptable rideshare match. We expect that future work might overcome such limitations and provide even further insight into gig worker surveillance while completing tasks.

**Other Deceptive Designs and Dark Patterns in Gig Work:** Beyond privacy, gig platforms may include other types of dark patterns towards the purpose of increasing worker labor, changing worker behavior, or otherwise steering workers towards platform-desired outcomes. We did not annotate our dataset for such other dark patterns within the scope of this study as we kept our codebooks purposely narrow in scope to best relate to platforms' data sharing practices at a higher-level, so our results should be considered only as lower-bound estimates of potential dark patterns deployed in gig platforms.

Upon further inspection of the Lyft driving operational issues mentioned earlier in § 7.3, we determined that the most likely cause was either our jailbroken Pixel 3 or our man-in-the-middle setup. Such requirements may have security motivations (e.g., Doordash explicitly notes on its site that their app is not likely to run on jailbroken devices and had operational issues immediately at app log-in). However, given the bring-your-own-device (BYOD) nature of gig work, device constraints constitute potential *forced actions*$^{H}$ in the control demanded over what might be a worker's personal device. This is in comparison to traditional labor, in which employees are often provided with work devices that have mutual privacy benefits (employers can secure their own networks and assets, while employees can keep personal devices free from employer monitoring) and autonomy benefits in the sense that personal devices can be kept free from employer control as well. However, more research is needed to understand the extent to which such requirements constitute *forced actions*$^{H}$, or how to mitigate against them if so.

# 8 Conclusion

This paper presented a rigorous, multifacted, and multimodal analysis of privacy issues faced by gig workers. Specifically, we find that gig workers are subjected to extensive PII collection and sharing when interacting with gig work platforms, including dozens of third-party endpoints and transmission of highly sensitive information like SSNs. Further, we identify numerous privacy-related dark patterns in gig platform user interfaces and interactions, including forced communication and information disclosure, as well as extensive nagging and cross-marketing via phone and email communication. Our findings suggest the need for better gig worker privacy protections, better separations between worker and consumer sides of the platforms, and improved autonomy for gig workers when it comes to controlling how their information is used. While we revealed interesting behavior from the perspective of gig workers, we identify numerous areas for future work, including additional visibility into data collection while performing work, and comparisons of platform behavior across jurisdictions. To facilitate future progress in these areas by the community, we will make our code and redacted datasets available.

# Acknowledgments

We thank the reviewers for their valuable feedback and suggestions for improving our paper. This research was partially funded by the NSF under the ProperData Frontier award (SaTC-1955227).

# References

[1] 2024. *Android Debug Bridge (adb) | Android Studio | Android Developers.* https://developer.android.com/tools/adb
[2] 2024. *Chrome DevTools | Chrome for Developers.* https://developer.chrome.com/docs/devtools/
[3] 2024. EasyList - Overview. https://easylist.to/
[4] 2024. *Frida • A world-class dynamic instrumentation toolkit.* https://frida.re/
[5] 2024. *Get a user-resettable advertising ID.* https://developer.android.com/training/articles/ad-id
[6] 2024. *Google Takeout.* https://takeout.google.com/settings/takeout
[7] 2024. *mitmproxy - an interactive HTTPS proxy.* https://mitmproxy.org/
[8] Jason Abbruzzese. 2021. Uber riders and drivers share fears about safety after company releases assault numbers. *NBC News* (2021). https://www.nbcnews.com/tech/tech-news/uber-riders-drivers-share-fears-about-safety-after-company-releases-n1097446
[9] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. 2014. The web never forgets: Persistent tracking mechanisms in the wild. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security.* 674–689.
[10] Gunes Acar, Marc Juarez, Nick Nikiforakis, Claudia Diaz, Seda Gürses, Frank Piessens, and Bart Preneel. 2013. FPDetective: dusting the web for fingerprinters. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security.* 1129–1140.
[11] California Consumer Privacy Act. 2020. California Consumer Privacy Act (Final Text of Proposed Regulations). https://www.oag.ca.gov/sites/all/files/agweb/pdfs/privacy/oal-sub-final-text-of-regs.pdf

[12] California Privacy Rights Act. 2020. California Privacy Rights Act. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

[13] Monica Anderson, Colleen McClain, Michelle Faverio, and Risa Gelles-Watnick. 2021. The State of Gig Work in 2021. https://www.pewresearch.org/internet/2021/12/08/the-state-of-gig-work-in-2021/

[14] Ira Anjali Anwar, Michaelanne Thomas, Kentaro Toyama, and Julie Hui. 2022. Gig Platforms as Faux Infrastructure: A Case Study of Women Beauty Workers in India. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 409 (nov 2022), 25 pages. https://doi.org/10.1145/3555134

[15] Annie I Antón, Julia Brande Earp, Qingfeng He, William Stufflebeam, Davide Bolchini, and Carlos Jensen. 2004. Financial privacy policies and the need for standardization. *IEEE Security & privacy* 2, 2 (2004), 36–45.

[16] Eduardo Blázquez, Sergio Pastrana, Álvaro Feal, Julien Gamba, Platon Kotzias, Narseo Vallina-Rodriguez, and Juan Tapiador. 2021. Trouble over-the-air: An analysis of fota apps in the android ecosystem. In *2021 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1606–1622.

[17] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. 2016. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proc. of PETS* 2016, 4 (2016), 237–254. https://content.sciendo.com/view/journals/popets/2016/4/article-p237.xml

[18] Rochelle A Cadogan et al. 2004. An imbalance of power: the readability of internet privacy policies. *Journal of Business & Economics Research (JBER)* 2, 3 (2004).

[19] Shauvik Roy Choudhary, Alessandra Gorla, and Alessandro Orso. 2015. Automated test input generation for android: Are we there yet?(e). In *2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE, 429–440.

[20] Andrea Continella, Yanick Fratantonio, Martina Lindorfer, Alessandro Puccetti, Ali Zand, Christopher Kruegel, Giovanni Vigna, et al. 2017. Obfuscation-Resilient Privacy Leak Detection for Mobile Apps Through Differential Analysis.. In *NDSS*, Vol. 17. 10–14722.

[21] Samantha Deloua. 2023. Doordash, Grubhub and Uber sue New York City over minimum wage law. *CNN* (2023). https://www.cnn.com/2023/07/06/business/doordash-grubhub-uber-sue-new-york-minimum-wage-law/index.html

[22] Android Developers. 2009.

[23] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376600

[24] Steven Englehardt, Jeffrey Han, and Arvind Narayanan. 2018. I never signed up for this! Privacy implications of email tracking. *Proc. Priv. Enhancing Technol.* 2018, 1 (2018), 109–126.

[25] Steven Englehardt and Arvind Narayanan. 2016. Online tracking: A 1-million-site measurement and analysis. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. 1388–1401.

[26] Nathan S Evans, Azzedine Benameur, and Yun Shen. 2015. All your root checks are belong to us: The sad state of root detection. In *Proceedings of the 13th ACM International Symposium on Mobility Management and Wireless Access*. 81–88.

[27] Shehroze Farooqi, Álvaro Feal, Tobias Lauinger, Damon McCoy, Zubair Shafiq, and Narseo Vallina-Rodriguez. 2020. Understanding incentivized mobile app installs on google play store. In *Proceedings of the ACM internet measurement conference*. 696–709.

[28] Parvez Faruki, Ammar Bharmal, Vijay Laxmi, Manoj Singh Gaur, Mauro Conti, and Muttukrishnan Rajarajan. 2014. Evaluation of android anti-malware techniques against dalvik bytecode obfuscation. In *2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*. IEEE, 414–421.

[29] Álvaro Feal, Paolo Calciati, Narseo Vallina-Rodriguez, Carmela Troncoso, Alessandra Gorla, et al. 2020. Angel or devil? a privacy study of mobile parental control apps. *Proceedings of Privacy Enhancing Technologies (PoPETS)* 2020 (2020).

[30] Alek Felstiner. 2011. Working the crowd: employment and labor law in the crowdsourcing industry. *Berkeley J. Emp. & Lab. L.* 32 (2011), 143.

[31] Organization for Economic Co-operation and Development. 2021. Roundtable on Dark Commercial Patterns Online: Summary of discussion. https://one.oecd.org/document/DSTI/CP(2020)23/FINAL/En/pdf

[32] FTC. 2017. Uber Settles FTC Allegations that It Made Deceptive Privacy and Data Security Claims. https://www.ftc.gov/news-events/news/press-releases/2017/08/uber-settles-ftc-allegations-it-made-deceptive-privacy-data-security-claims

[33] Julien Gamba, Álvaro Feal, Eduardo Blazquez, Vinuri Bandara, Abbas Razaghpanah, Juan Tapiador, and Narseo Vallina-Rodriguez. 2023. Mules and permission laundering in android: Dissecting custom permissions in the wild. *IEEE Transactions on Dependable and Secure Computing* (2023).

[34] Julien Gamba, Mohammed Rashed, Abbas Razaghpanah, Juan Tapiador, and Narseo Vallina-Rodriguez. 2020. An analysis of pre-installed android software. In *2020 IEEE symposium on security and privacy (SP)*. IEEE, 1039–1055.

[35] Alejandro Gómez-Boix, Pierre Laperdrix, and Benoit Baudry. 2018. Hiding in the crowd: an analysis of the effectiveness of browser fingerprinting at large scale. In *Proceedings of the 2018 world wide web conference*. 309–318.

[36] Mark A Graber, Donna M D Alessandro, and Jill Johnson-West. 2002. Reading level of privacy policies on internet health web sites. *Journal of Family Practice* 51, 7 (2002), 642–642.

[37] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174108

[38] Colin M. Gray, Cristiana Santos, and Nataliia Bielova. 2023. Towards a Preliminary Ontology of Dark Patterns Knowledge. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 286, 9 pages. https://doi.org/10.1145/3544549.3585676

[39] Colin M. Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 172, 18 pages. https://doi.org/10.1145/3411764.3445779

[40] Moritz Gruber, Christian Höfig, Maximilian Golla, Tobias Urban, and Matteo Große-Kampmann. 2022. "We may share the number of diaper changes": A Privacy and Security Analysis of Mobile Child Care Applications. *Proceedings on Privacy Enhancing Technologies* 3 (2022), 394–414.

[41] Johanna Gunawan, Amogh Pradeep, David Choffnes, Woodrow Hartzog, and Christo Wilson. 2021. A Comparative Study of Dark Patterns Across Web and Mobile Modalities. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 377 (oct 2021), 29 pages. https://doi.org/10.1145/3479521

[42] Johanna Gunawan, Cristiana Santos, and Irene Kamara. 2022. Redress for Dark Patterns Privacy Harms? A Case Study on Consent Interactions. In *Proceedings of the 2022 Symposium on Computer Science and Law* (Washington DC, USA) *(CSLAW '22)*. Association for Computing Machinery, New York, NY, USA, 181–194. https://doi.org/10.1145/3511265.3550448

[43] Frode Guribye, Oda Elise Nordberg, Are Nyhammer, Marija Slavkovik, and Than Htut Soe. 2021. Dark patterns in cookie consent notices: new definitions and mitigation strategies. *Position Papers of CHI'22 "What Can CHI Do About Dark Patterns?* (2021). https://drive.google.com/file/d/1V-P98rH_wqlzLaqY1HZZPDGfdHcCKWBx/view

[44] Yuyu He, Lei Zhang, Zhemin Yang, Yinzhi Cao, Keke Lian, Shuai Li, Wei Yang, Zhibo Zhang, Min Yang, Yuan Zhang, et al. 2020. Textexerciser: Feedback-driven text input exercising for android applications. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1071–1087.

[45] Richard Heeks. 2019. Digital Economy and Digital Labour Terminology: Making Sense of the 'Gig Economy', 'Online Labour', 'Crowd Work', 'Microwork', 'Platform Labour', Etc.. In *Development Informatics Working Papers*. Global Development Institute. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3431728

[46] Jaap-Henk Hoepman. 2014. Privacy Design Strategies. *IFIP Advances in Information and Communication Technology* 428 (2014). https://doi.org/10.1007/978-3-642-55415-5_38

[47] Shan Jiang, Le Chen, Alan Mislove, and Christo Wilson. 2018. On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 863–872. https://doi.org/10.1145/3178876.3186134

[48] Ja'han Jones. 2021. Apps like Uber and Lyft exploit their workers. We can't ignore it any longer. *MSNBC* (2021). https://www.msnbc.com/the-reidout/reidout-blog/uber-lyft-worker-exploitation-rcna5865

[49] Marc Juarez, Sadia Afroz, Gunes Acar, Claudia Diaz, and Rachel Greenstadt. 2014. A critical evaluation of website fingerprinting attacks. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. 263–274.

[50] Thivya Kandappu, Vijay Sivaraman, Arik Friedman, and Roksana Boreli. 2013. Exposing and mitigating privacy loss in crowdsourced survey platforms. In *Proceedings of the 2013 workshop on Student workhop*. 13–16.

[51] Sangmi Kim, Elizabeth Marquis, Rasha Alahmad, Casey S. Pierce, and Lionel P. Robert Jr. 2018. The Impacts of Platform Quality on Gig Workers' Autonomy and Job Satisfaction. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 181–184. https://doi.org/10.1145/3272973.3274050

[52] Jennifer King and Adriana Stephan. 2021. Regulating Privacy Dark Patterns in Practice - Drawing Inspiration from the California Privacy Rights Act. *Georgetown Law Technology Review* 5 (2021). Issue 250.

[53] Monica Kowalczyk, Johanna T. Gunawan, David Choffnes, Daniel J Dubois, Woodrow Hartzog, and Christo Wilson. 2023. Understanding Dark Patterns in Home IoT Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 179, 27 pages. https://doi.org/10.1145/3544548.3581432

[54] Chiara Krisam, Heike Dietmann, Melanie Volkamer, and Oksana Kulyk. 2021. Dark Patterns in the Wild: Review of Cookie Disclaimer Designs on Top 500 German Websites. In *European Symposium on Usable Security 2021* (Karlsruhe, Germany) *(EuroUSEC '21)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3481357.3481516

[55] Anh Le, Janus Varmarken, Simon Langhoff, Anastasia Shuba, Minas Gjoka, and Athina Markopoulou. 2015. AntMonitor: A system for monitoring from mobile devices. In *Proceedings of the 2015 ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big (Internet) Data*. 15–20.

[56] Rui Li, Wenrui Diao, Zhou Li, Shishuai Yang, Shuang Li, and Shanqing Guo. 2021. Android custom permissions demystified: A comprehensive security evaluation. *IEEE Transactions on Software Engineering* 48, 11 (2021), 4465–4484.

[57] Toby Jia-Jun Li, Yuwen Lu, Jaylexia Clark, Meng Chen, Victor Cox, Meng Jiang, Yang Yang, Tamara Kay, Danielle Wood, and Jay Brockman. 2022. A Bottom-Up End-User Intelligent Assistant Approach to Empower Gig Workers against AI Inequality. In *2022 Symposium on Human-Computer Interaction for Work* (Durham, NH, USA) *(CHIWORK 2022)*. Association for Computing Machinery, New York, NY, USA, Article 9, 10 pages. https://doi.org/10.1145/3533406.3533418

[58] Davide Maiorca, Davide Ariu, Igino Corona, Marco Aresu, and Giorgio Giacinto. 2015. Stealth Attacks: An Extended Insight into the Obfuscation Effects on Android Malware. *Computers & Security* (2015).

[59] Elizabeth B. Marquis, Sangmi Kim, Rasha Alahmad, Casey S. Pierce, and Lionel P. Robert Jr. 2018. Impacts of Perceived Behavior Control and Emotional Labor on Gig Workers. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 241–244. https://doi.org/10.1145/3272973.3274065

[60] Arunesh Mathur, Gunes Acar, Michael J. Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 81 (nov 2019), 32 pages. https://doi.org/10.1145/3359183

[61] Celestin Matte, Nataliia Bielova, and Cristiana Santos. 2020. Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. In *Proc. of IEEE Symposium on Security and Privacy*.

[62] Thomas Mildner and Gian-Luca Savino. 2021. Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article

464, 7 pages. https://doi.org/10.1145/3411763.3451659
[63] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376321
[64] European Parliament and Council of European Union. 2016. EU General Data Protection Regulation (GDPR). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN.
[65] Sebastian Poeplau, Yanick Fratantonio, Antonio Bianchi, Christopher Kruegel, and Giovanni Vigna. 2014. Execute this! analyzing unsafe and malicious dynamic code loading in android applications.. In *NDSS*, Vol. 14. 23–26.
[66] Amogh Pradeep, Álvaro Feal, Julien Gamba, Ashwin Rao, Martina Lindorfer, Narseo Vallina-Rodriguez, David Choffnes, et al. 2023. Not Your Average App: A Large-scale Privacy Analysis of Android Browsers. In *Privacy Enhancing Technologies Symposium (was International Workshop of Privacy Enhancing Technologies)*.
[67] Amogh Pradeep, Muhammad Talha Paracha, Protick Bhowmick, Ali Davanian, Abbas Razaghpanah, Taejoong Chung, Martina Lindorfer, Narseo Vallina-Rodriguez, Dave Levin, and David Choffnes. 2022. A comparative analysis of certificate pinning in Android & iOS. In *Proceedings of the 22nd ACM Internet Measurement Conference*. 605–618.
[68] Abbas Razaghpanah, Rishab Nithyanand, Narseo Vallina-Rodriguez, Srikanth Sundaresan, Mark Allman, Christian Kreibich, Phillipa Gill, et al. 2018. Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem. In *The 25th Annual Network and Distributed System Security Symposium (NDSS 2018)*.
[69] Jingjing Ren, Ashwin Rao, Martina Lindorfer, Arnaud Legout, and David Choffnes. 2016. Recon: Revealing and controlling pii leaks in mobile network traffic. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. 361–374.
[70] Irwin Reyes, Primal Wijesekera, Joel Reardon, Amit Elazari Bar On, Abbas Razaghpanah, Narseo Vallina-Rodriguez, Serge Egelman, et al. 2018. "Won't somebody think of the children?" examining COPPA compliance at scale. In *The 18th Privacy Enhancing Technologies Symposium (PETS 2018)*.
[71] Hilary C Robinson. 2017. *Making a digital working class: Uber drivers in Boston, 2016-2017*. Ph. D. Dissertation. Massachusetts Institute of Technology.
[72] Alex Rosenblat and Luke Stark. 2016. Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International journal of communication* 10 (2016), 27.
[73] Johnny Saldana. 2013. *The Coding Manual for Qualitative Researchers (2nd Ed.)*.
[74] Shruti Sannon, Billie Sun, and Dan Cosley. 2022. Privacy, surveillance, and power in the gig economy. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15. https://doi.org/doi/pdf/10.1145/3491102.3502083
[75] Brennan Schaffner, Neha A. Lingareddy, and Marshini Chetty. 2022. Understanding Account Deletion and Relevant Dark Patterns on Social Media. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 417 (nov 2022), 43 pages. https://doi.org/10.1145/3555142
[76] Shiwangi Singh, Rucha Gadgil, and Ayushi Chudgor. 2014. Automated Testing of Mobile Applications using Scripting Technique: A Study on Appium. *International Journal of Current Engineering and Technology (IJCET)* (2014).
[77] USA Social Security Administration. 2005. Identity Theft and Your Social Security Number. https://www.ssa.gov/pubs/EN-05-10064.pdf
[78] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. *Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3419249.3420132
[79] Daniel Susser, Beate Roessler, and Helen Nissenbaum. 2019. Online Manipulation: Hidden Influences in a Digital World.
[80] Austin Toombs, Colin Gray, Guoyang Zhou, and Ann Light. 2018. Appropriated or Inauthentic Care in Gig-Economy Platforms: A Psycho-Linguistic Analysis of Uber and Lyft. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188657
[81] Pelayo Vallina, Álvaro Feal, Julien Gamba, Narseo Vallina-Rodriguez, and Antonio Fernández Anta. 2019. Tales from the porn: A comprehensive privacy analysis of the web porn ecosystem. In *Proceedings of the Internet Measurement Conference*. 245–258.
[82] Narseo Vallina-Rodriguez, Srikanth Sundaresan, Abbas Razaghpanah, Rishab Nithyanand, Mark Allman, Christian Kreibich, and Phillipa Gill. 2016. Tracking the trackers: Towards understanding the mobile advertising and tracking ecosystem. *arXiv preprint arXiv:1609.07190* (2016).
[83] Nisha Vinayaga-Sureshkanth, Raveen Wijewickrama, Anindya Maiti, and Murtuza Jadliwala. 2022. An investigative study on the privacy implications of mobile e-scooter rental apps. In *Proceedings of the 15th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 125–139.
[84] Ari Ezra Waldman. 2020. Cognitive biases, dark patterns, and the 'privacy paradox'. *Current Opinion in Psychology* 31 (Feb. 2020), 105–109.
[85] Samuel Warren and Louis Brandeis. 1890. The Right to Privacy. In *Harvard Law Review*, Vol. IV. Issue 5. https://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html
[86] Zhonghao Yu, Sam Macbeth, Konark Modi, and Josep M Pujol. 2016. Tracking the trackers. In *Proceedings of the 25th International Conference on World Wide Web*. 121–132.

# A Data Sharing Endpoint Categories

Table 7 presents the categories of third parties that receive PII from platforms in our study.

# B Content Analysis Appendices

Table 8 provides the set of initial codes used to label platforms in our dataset. When analyzing these resultant codes, we later relate gig platform behavior to dark patterns in Table 9.

Here we produce our final dark patterns codebook and provide additional descriptions or context for codebook development, including any extensions made to the Gray et al. [38] ontology using their syntactical structure..

## B.1 Content Screenshots and Examples

Here we provide screenshots from our study, used to contextualize results.

# C Obstacles to Gig Completion

Here we describe why we could not complete gigs on some platforms within this study. We provide this information in the hopes of helping future studies overcome these obstacles.

**Methodological Match Constraints:** Our methods required only accepting work from our own experiment personas. If we were successfully able to register for a platform and attempt work, but the platform did not match us to our study persona within a reasonable amount of time, we stopped waiting for our intended match and ended our "shift." Future studies might leverage multiple consumer-user personas to increase the chances of successfully matching with study accounts, though this would scale up mobile test infrastructure requirements.

**Technical Constraints:** We used jailbroken devices in order to facilitate our data collection, but some platforms (and mobile apps generally) may not function correctly on jailbroken devices. DoorDash documentation mentioned potential functionality issues with jailbroken devices, with our DoorDash apps crashing when attempting to log in. Lyft allowed us to log in and use the app on our test device up until starting our "shift" (going online to begin accepting rides). External tests on personal, unjailbroken devices showed that this was not an issue with our accounts but rather with the device itself, despite our best efforts to verify the test device with Lyft Support.

**Other Constraints:** We were unable to complete Instacart and GrubHub gigs within our study duration, as Instacart would not let us log into our account when we attempted to work (without explanation), and GrubHub placed us on a worker waitlist in our region.

**Table 7: Popular categories of third parties receiving at least one type of PII from platforms in this study. Parentheses indicate PII sent without hashing.**

| Category | SSN | Ad. ID | Email | Name | Phone | Add. | Loc. | Total |
|---|---|---|---|---|---|---|---|---|
| web API | | (11) | (6) | (11) | | (2) | (4) | 15 |
| analytics | | (11) | 8 (6) | (10) | (4) | (1) | (3) | 14 |
| social network | 2 | (8) | 10 (1) | 11 (4) | 8 (1) | 1 | | 13 |
| search engine | | (3) | 5 (2) | (6) | 3 (2) | | (3) | 10 |
| marketing | | (4) | 2 (1) | (3) | (1) | | | 6 |
| advertising | | | 2 (1) | (5) | (1) | | | 6 |
| deep linking | | (6) | | (1) | | | | 6 |
| privacy compliance | | | 1 | | | 2 | | 3 |
| payments | | | | (3) | | (1) | | 3 |
| web infrastructure | | | (1) | (2) | (1) | | | 2 |
| miscellaneous - identity verification | | | (1) | (1) | | (1) | | 1 |
| miscellaneous - experiments | | | (1) | (1) | | | | 1 |
| e-commerce | | | | (1) | (1) | | | 1 |
| web accessibility | | | 1 | (1) | | | | 1 |
| miscellaneous - visualization | | | | (1) | | | | 1 |
| miscellaneous - computing | | | | (1) | | | | 1 |
| miscellaneous - cybersecurity | | | | (1) | | | | 1 |
| fraud prevention | | | | (1) | | | | 1 |
| miscellaneous - surveys | | | | | | 1 | | 1 |
| miscellaneous - technology | | | | (1) | | | | 1 |
| customer communication | | | | (1) | | | (1) | 1 |

**Table 8: Codebook of gig platforms' registration requirements, relevant behaviors, and types of off-platform content used to understand privacy-erosive design strategies or dark patterns in our gig platform interactions.**

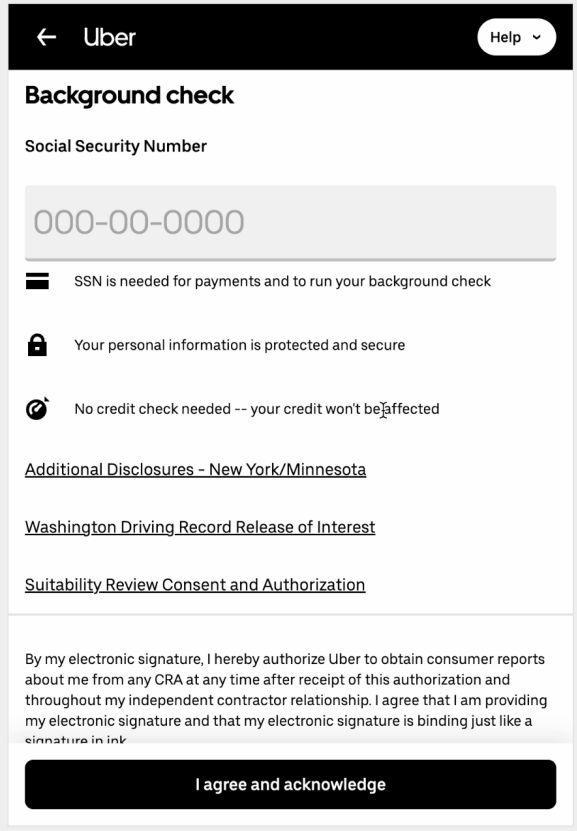
| Category | Code | Description |
|---|---|---|
| Required Worker Information | Name | PII. Worker's first and last name. |
| | Phone Number | PII. Worker's phone number. |
| | Email Address | PII. Worker's email address. |
| | Age | PII depending on granularity. Worker's age information, collected as date of birth, age in years, or attestment to age requirement. |
| | Location | PII. Worker's location information, collected at any granularity (e.g., ZIP code, state, city, street address), inferred by platform, or requested in permissions. |
| | Password | New password of choice for logging into plaftorm. |
| | Proof of Identity | PII. Any information (Social Security Number, documentation, photograph) required to pass background checks or identity verification. |
| | Profile Photo | PII. Photograph of worker's face for their worker or freelancer profile, not for identity verification purposes but to be shown to customers. |
| Registration Blockers | App required | When a registration procedure initiated in a browser site requires use of the mobile app to continue. |
| | Payment required | When a registration procedure requires payment to continue, including flat fees, recurring fees, and billed-later payment information. |
| | Identity verification required | When a registration procedure requires some form of identity verification to continue, including submission of highly identifying information or uploading relevant documents. |
| | Waitlist | When market saturation prevents further registration and instead places the partially-registered worker on waitlist. |
| | Work Endorsements | Requests for external endorsements or testimony of a worker's background. |
| Platform Behavior | Account creation disclosure | Whether a platform made explicitly clear that an account was being created, through password requirements or via language in the interface. |
| | Partitioned registration procedure | Whether a registration procedure is clearly intended for workers and distinct from general user registration. |
| | Partitioned mobile application | Whether a platform's mobile app for workers is distinct and separate from the general user app. |
| | SMS communications | Whether SMS communications (beyond verification codes) were sent to worker persona phones. |
| | Email communications | Whether email communications (beyond verification codes) were sent to worker persona inboxes. |
| Off-Platform Communication Nudges | Customer content | Whether off-platform communications sent by a platform address the worker recipient as a customer-side user. |
| | Work nudges | Whether off-platform communications sent by a platform nudge workers to complete gigs. |
| | Registration nudges | Whether off-platform communications sent to pre-verified workers nudge workers to complete registration. |

**Table 9: Codebook used for labeling dark patterns in our gig platform interactions. Pattern names and descriptions are adapted primarily from the Gray et al. [38] ontology, incorporating other taxonomies when necessary. Bolded items are new meso- or low-level patterns we appended to the ontology according to the methods described by Gray et al. [38], and patterns are superscripted with their pattern level: high$^{H}$, meso$^{M}$, and low$^{L}$.**

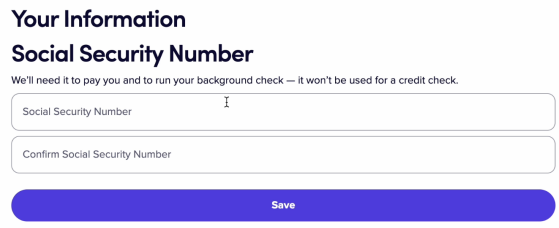
| High-Level Dark Pattern | Meso- or Lower-Level Dark Pattern | Description |
|---|---|---|
| Forced Action$^{H}$ | Forced Registration$^{M}$ | Designs that create user accounts (login information or a similar record) prior to and despite incomplete worker registration, or that force non-worker account features or centralization upon gig workers. These subvert user expectations that the process of submitting PII gaining platform approval is separate from account creation. |
| | Forced Communication and Disclosure$^{M}$ | Designs subverting user expectations that a platform will only request the information needed to complete their applications, and instead tricking workers into additional disclosures. We additionally note *Social Pyramid*$^{L}$ and *Friend Spam*$^{L}$ patterns. |
| | Forced Modality Switching$^{M}$ [53, 75][11] | Subverts user expectations that a given task can be completed in the modality they are currently using, instead blocking task completion until the user performs additional labor and connects to the system through an additional modality. |
| | Pay to Work | Designs that use *Forced Action*$^{H}$ to offload the costs of doing business onto gig workers, making the worker pay to overcome barriers to gaining platform approval. Payment may be kept until the last step with varying transparency in prior steps, in which this pattern may also be considered under *Sneaking*$^{H}$. |
| Sneaking$^{H}$ | Shadow User Profiles$^{L}$ [17] (Hiding Information$^{M}$) | Technical designs in which service providers track information about individuals beyond the borders of the service itself, without users' informed knowledge. Bösch et al. [17] describe this primarily through incoming tracking information (the service receiving information from external sources), but we additionally interpret this pattern to include platforms' exfiltration of user information to tracking third-parties. |
| Interface Interference | Bad (Messaging) Defaults$^{M}$ [17] | Subverts user expectations that communications and messaging settings (as applied to collected contact information) will be used in their best interest. |

**Table 10: Average number of messages (SMS and emails) sent per day, per-platform during the first two weeks after first attempting to register for a platform, separated by fictional [F] and verified [V] personas. Averages of over one message/day are bolded.**

| Platform | Avg. SMS/Day [F] | Avg. SMS/Day [V] | Avg. Emails/Day [F] | Avg. Emails/Day [F] |
|---|---|---|---|---|
| Uber | 0.429 | 0.286 | **1** | 0.643 |
| Lyft | 0.286 | 0.357 | 0.643 | 0.929 |
| Wag | 0.214 | 0.071 | 0.286 | 0.429 |
| Doordash | 0.143 | 0.5 | 0.071 | 0.643 |
| Airtasker | 0.071 | 0 | 0.214 | 0.143 |
| Grubhub | 0.071 | 0.143 | 0 | 0.143 |
| Instacart | 0.071 | 0.071 | 0 | 0.5 |
| Rover | 0.286 | 0.286 | 0.571 | **1.286** |
| Thumbtack | 0 | 0 | **8.286** | **20.571** |
| Upwork | 0 | 0 | 0.071 | 0.071 |
| Peopleperhour | 0 | 0 | 0.214 | 0.429 |
| Fiver | 0 | 0 | 0.5 | 0.5 |
| Taskrabbit | 0 | 0 | 0.429 | 0.429 |
| urbansitter | 0 | 0 | 0.357 | 0.5 |
| Care.com | 0 | 0 | 0.214 | **1.214** |

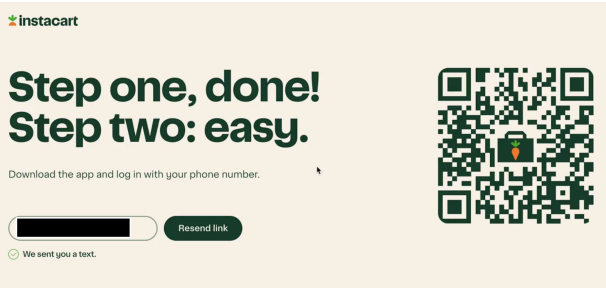



(a) Uber's SSN form submission interface, on desktop. The modal appears to be designed for responsive screen design in the event that registrants use a mobile browser.

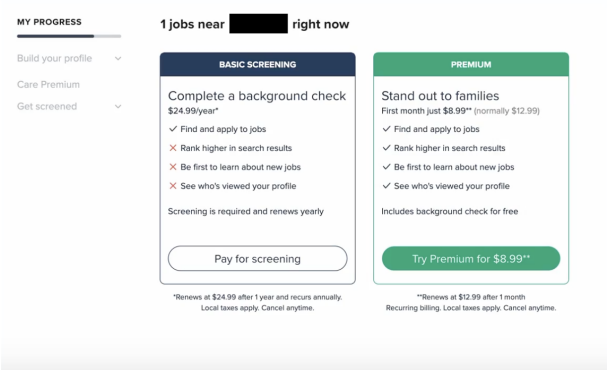



(b) Lyft's SSN form submission interface, on desktop.

Figure 2: Screenshots from the SSN desktop submission pages for (a) Uber and (b) Lyft.

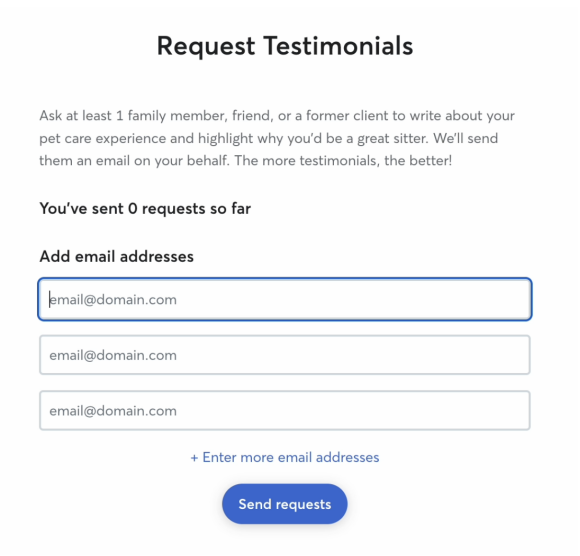



(a) The final browser page provided by Instacart before requiring the mobile app to complete registration.

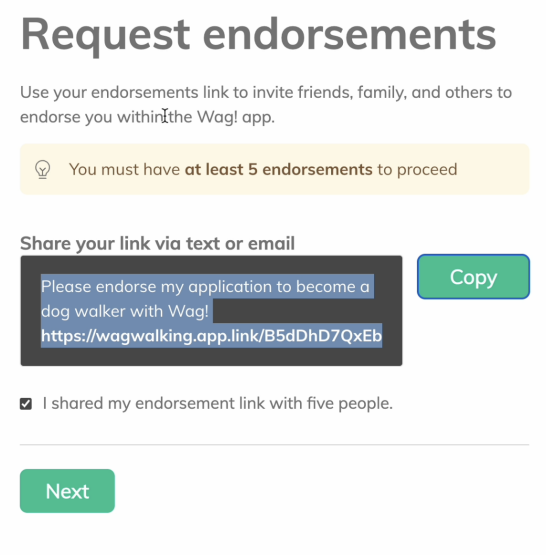



(b) Mandatory background check payment options for Care.com, including a "free" background check if subscribed to a monthly premium service.

Figure 3: Screenshots of (a) the last accessible step for Instacart browser registration and (b) background check payment options for Care.com. In Experiment 1, we halted registration when encountering such steps, even after having submitted some PII in previous registration steps.




(a) Rover's endorsement request pages, on desktop.




(b) Wag's endorsement request interfaces, on desktop.

**Figure 4: Screenshots from the endorsement request pages for (a) Rover and (b) Wag.**