Is Your Inseam a Biometric? Evaluating the Understandability of Mobile Privacy Notice Categories

Rebecca Balebako, Richard Shay, and Lorrie Faith Cranor

July 17, 2013

CMU-CyLab-13-011

CyLab
Carnegie Mellon University
Pittsburgh, PA 15213

Is Your Inseam a Biometric? Evaluating the Understandability of Mobile Privacy Notice Categories

Rebecca Balebako
Carnegie Mellon University
Pittsburgh, PA
balebako@cmu.edu

Richard Shay
Carnegie Mellon University
Pittsburgh, PA
rshay@cmu.edu

Lorrie Faith Cranor
Carnegie Mellon University
Pittsburgh, PA
lorrie@cs.cmu.edu

ABSTRACT

The National Telecommunications and Information Administration (NTIA) has proposed a set of categories and definitions to create a United States national standard for short-form privacy notices on mobile devices. These notices are intended to facilitate user decision-making by categorizing both smartphone data to be shared and the entities with which that data is shared. In order to determine whether users consistently understand these proposed categories and their definitions, we conducted an online study with 791 participants. We found that participants had low agreement on how different data and entities should be categorized. We also compared our online results with those provided by four anonymous NTIA stakeholders, finding that even the stakeholders did not consistently categorize data or entities. Our work highlights areas of confusion for both survey participants and experts in the proposed scheme, and we offer suggestions for addressing these issues.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Privacy, Notifications

Keywords

notifications, privacy, public policy, notice, security, user study, mobile, smartphones

1. INTRODUCTION

As smartphone and mobile device usage grows, so too grows the amount of data being collected by apps (such as those available on the Apple App Store and the Google Play Store). Concerns about data collection have led different platforms to provide notice about what permissions for data an app requests. This notice is not standardized across platforms; for example, Android shows a notice and dialog when an app is being installed, while Apple iOS shows a notice and dialog the first time an app transmits certain data.

In an effort to improve transparency and usability, the National Telecommunications and Information Administration (NTIA) has initiated a multi-stakeholder effort to develop a standardized short-form privacy disclosure on mobile devices. These standardized disclosures will show the user both what data is being shared and with which entities it is being shared. The NTIA multi-stakeholder process (NTIA MSHP) has been underway since June 2012, and is now near completion. The NTIA stakeholders – representing app developers, consumer groups, and government – have developed a Code of Conduct for app developers for a short-form privacy notice.[1] This Code outlines seven categories of data and eight categories of third-party entities that apps should include in short-form privacy notifications.

While the Code and its categories of data and entities have been in development for over a year, this paper presents the first human-subjects testing of its usability. We present the results of a 791-participant online study in which we investigate whether participants are able to categorize realistic data-sharing scenarios using the NTIA MSHP categories. We also present results from four experts who participate in the NTIA MSHP process. Of the 52 examples given in our scenarios, participants showed low common agreement for how to classify the data or entity in 23 cases. Overall, we found that many of the proposed categories and definitions were not consistently understood by our participants, including our expert participants. We discuss categories that need clarification, and offer suggests for improving the Code based on our findings.

We discuss background and related work in the Section 2. In Section 3, we describe our methodology. In Section 4, we present the results from our study. We explore the limitations of our study in 5. We discuss our findings and recommend steps for improvement in Section 6.

2. BACKGROUND AND RELATED WORK

This section describes the development of the code of conduct for a short-form mobile privacy policy, which includes the terms we test in this research. We also discuss existing research about user's conceptions of smartphone privacy and security.

2.1 Policy and Usability

In 2012, the White House issued a report on consumer data privacy, which included a Consumer Privacy Bill of Rights [3]. The second principle in the bill of rights is transparency, which is summarized as: "Consumers have a right to easily understandable and accessible information about privacy and security practices."

The report emphasizes the role of multi-stakeholder processes to develop and define privacy practices and technologies, and to

[1] http://www.ntia.doc.gov/files/ntia/publications/mobileappdraftmay16_2013.pdf

develop "enforceable codes of conduct." It calls upon the Department of Commerce's National Telecommunications and Information Administration (NTIA) to lead multi-stakeholder processes. The NTIA launched one such initiative on Mobile Application Transparency in 2012. The result, at the time of writing, includes a draft code of conduct on mobile short-form notices. This draft defines a standard short-form privacy notice for apps, which does not substitute for a longer complete privacy policy.

In a 2013 privacy report, the Federal Trade Commission encourages consumer testing of privacy notices "to ensure meaningful consumer comprehension" [12]. The NTIA multi-stakeholder group struggled with the role of usability testing in drafting the policy. While a usability subgroup was initiated and met several times, so far the larger group has been unable to reach consensus on what should be tested. The work reported in this paper was initiated and run independently of the usability group. Our goal was to examine one portion of the notice, in particular the understandability of the wording suggested for the short form notices. We believe our results can inform the process and offer insights into how the code of conduct can be implemented successfully.

Usability and consumer testing has previously played a role in developing standards and policy for privacy and technology. User tests conducted by members of the working group were used during the development of P3P to test the feasibility of user tools [13].

The Gramm-Leach-Bliley Act of 1999 (GLBA) required financial institutions to provide a privacy notice to it consumer customers. Kleimann Communication Group developed a privacy notice prototype using participant studies. Their qualitative research involved iterating over prototypes with several participant tests — including focus groups, pretests, and usability testing. [1]. This prototype was then tested against several others with quantitative testing [27], and the results were used to develop the final 'model' form, which presents information in a tabular format [2].

Kelley et al. used a similar approach to develop and test a "privacy nutrition label" for websites. They also found that a tabular format was liked by users and facilitated policy comparison [22].

The Future of Privacy Forum researched consumer's responses to notices about online behavioral advertising (OBA). They found that transparency and choice increased people's comfort with OBA. This study also compared the effectiveness of different icons in communication effectivensss [18]. Unfortunately, the icon revealed as the most effective was not selected by the ad industry. Further work found that the icon and the tagline selected by industry were not noticed by users, and that users did not understand the tagline and were afraid to click on the icon and text [24].

2.2 User Conception of Smartphone Privacy and Security

Several categories of smartphone data raise privacy concerns. Biometrics data can serve as a unique identifier for linking to a user's other activities [11]. These unique identifiers can cause particular privacy concern as they often cannot be revoked or changed even when stolen [32]. Users' concerns about the collection of their browsing history have been documented a number of times [26, 29, 39]. Additional privacy issues inherent in the collection of metadata, such as logs of browsing, phone calls, or text messages, have been publicized in the wake of revelations about the U.S. National Security Agency's PRISM program. Phone usage data and metadata can be used to infer hobbies, medical conditions, and beliefs [35]. A user's beliefs and activities can often be inferred from the people with whom they associate [35]. Users' privacy can be violated from simply learning their associations, as contained in their address book or social network connections. The collection of users' contacts has led to privacy outrage in the past, such as when Facebook's smart-phone app was discovered uploading the names and phone numbers from users' address books to Facebook's servers without providing notice [5]. The metadata from a user's emails alone can be used to infer their real-life social network and associations, as demonstrated by the art project Immersion [20,33]. Furthermore, the fact that data is collected can have a chilling effect on individuals' free speech [34], and most individuals would likely be unaware when their data and metadata could reveal them to be violating the law [28].

Sensitive information may exacerbate privacy concerns. Financial information can cause privacy issues both because an individual might be loath to disclose information about their earnings, as well as fear about the potential of price discrimination [41]. Similarly, privacy is fundamental to a doctor-patient relationship, and disclosure of health information could cause financial harm if used by a health-insurance company to deny coverage to a patient [4].

Location data can also arouse privacy concern, particularly when the location where a user is located is not a location visited by many people [37] or location information is highly granular [7]. Users also find their files, such as photos and videos, to be sensitive [30]. Furthermore, nearly all participants in a study by Felt et al. would have been upset if the text messages and emails stored on their phone were shared publicly [15].

In addition to the type of data, users are concerned about with whom the data is shared. Social networks, government, and advertisers may all be of particular concern. A PewResearch Study found that 63% of Americans would feel their privacy had been violated if they knew the government had collected information about their calls and online communication [36]. Social networks may be a concern both due to the accidental leakage of private information (willingly provided by the user) to unanticipated parties [17, 23]. Urban et al. found survey participants were unwilling to share contact information with advertisers [40]. Several studies have found that Americans are concerned about online behavioral advertising, and often do not understand the mechanisms (including data reselling and data aggregation) behind it [29, 38, 39]. Balebako et al. found that smartphone users often did not recognize the names of third-party advertisers or data aggregators with which smartphone games shared data [6].

Several studies have examined smartphone privacy notifications. An Internet survey of 308 Android users and a laboratory study of 25 Android users found that only 17% paid attention to the permissions when installing an application. They also found that only 3% of the Internet survey respondents demonstrated full comprehension of the permissions screen [16]. Kelley et al. found that when Android users were presented with privacy information, they chose apps with fewer permission requests [21]. Balebako et al. examined users' reactions to a user interface displaying information about data collected, and found users were surprised by the amount and destinations of data [6]. Felt et al. propose a framework for smartphone platforms to request permission for data from the user [14].

2.3 NTIA MSHP draft wording

We tested the wording used in the NTIA MSHP draft code published on April 29, 2013.[2] We deliberately did not change, add, or in any way modify the wording or punctuation. The draft includes seven categories of information to include in app privacy disclosures. It also includes eight categories of entities with which data

[2] http://www.ntia.doc.gov/files/ntia/publications/mobileappdraftapril29_2013_draft1b_fs.pdf

might be shared. The draft includes short definitions for all information types and entities — referred to throughout the paper as the "parenthetical" text — shown in parentheses below.

The categories for data types are:

- Biometrics (information about your body, including fingerprints, facial recognition, signatures and/or voice print.)
- Browser History and Phone or Text Log (A list of websites visited, or the calls or texts made or received.)
- Contacts (including list of contacts, social networking connections or their phone numbers, postal, email and text addresses.)
- Financial Information (Includes credit, bank and consumer-specific financial information such as transaction data.)
- Health, Medical or Therapy Information (including health claims and information used to measure health or wellness.)
- Location (precise past or current location and history of where a user has gone.)
- User Files (files stored on the device that contain your content, such as calendar, photos, text, or video.)

The categories for entities with which data was shared are:

- Ad Networks (Companies that display ads to you through apps.)
- Carriers (Companies that provide mobile connections.)
- Consumer Data Resellers (Companies that sell consumer information to other companies for multiple purposes including offering products and services that may interest you.)
- Data Analytics Providers (Companies that collect and analyze your data.)
- Government Entities (Any sharing with the government except where required or expressly permitted by law.)
- Operating Systems and Platforms (Software companies that power your device, app stores, and companies that provide common tools and information for apps about app consumers.)
- Other Apps (Other apps of companies that the consumer may not have a relationship with)
- Social Networks (Companies that connect individuals around common interests and facilitate sharing.)

3. METHODOLOGY

We conducted an online survey using Amazon's Mechanical Turk crowdsourcing service (MTurk)[3] over a two-week period in May 2013. Participants were recruited with the text, "Give us your opinion about information about smartphone apps. This should take 15-25 minutes," and paid $1 for completing the survey.

Previous research has shown that MTurk studies provide fairly representative samples of the US population, and demonstrated that offline experimental results can be successfully replicated using MTurk [9, 31]. Furthermore, while MTurk workers are younger and more technically savvy than the general US population, MTurk has been shown to provide a more diverse sample than a university lab survey [10, 19]. Using MTurk has allowed us to conduct our study with a larger and more diverse sample than would otherwise have been possible.

We also invited NTIA MSHP members to participate in the same study. We advertised the study to MSHP members through announcements by email and a brief presentation at one of their meetings. MSHP members answered two additional questions about their role in the process. MSHP participants were not compensated. The process for participating in the NTIA is open, but requires a time commitment and dedication to attend and participate in the meetings. These participants are considered experts, since they are familiar with objectives of the NTIA and have worked to shape the draft Code.

Results from both the MTurk participants and NTIA MSHP experts are described in Section 4. In this section, we describe how our survey was designed and tested, and who participated.

3.1 Survey Design

Our survey presented participants with a sequence of 10 smartphone-app scenarios. In each scenario, we described the app's purpose, what data it collects, and with which entities it shares that data. Some scenarios also included an explanation about why the data is collected. We then asked participants to categorize both the data being collected and the entities with which it is shared, according to the NTIA categories. An example scenario is below; all ten scenarios are provided in the Appendix.

> The Fitness app integrates with your FitMonitor (FitMonitor is a special pedometer and activity monitor, purchased separately) to allow you to track and improve your fitness activities and level.
>
> Fitness app will collect information on how many steps you have taken, how long you've slept, and allow you to enter you weight and body fat.
>
> Fitness app will notify sports and health companies if you achieve certain goals, and these companies will send you valuable coupons as awards.

We attempted to represent every data category and every entity category from the NTIA draft in our scenarios.

Our scenarios were designed to be realistic. Many scenarios were based on real apps or websites, though we changed the names and adjusted the wording in order to avoid confusion if the participant was already familiar with the real app. In three cases, we used the names of real companies — Apple, Facebook, and Google — in order to investigate whether participants considered them to be social networks or operating systems.

We included several scenarios that may be considered privacy sensitive. Two scenarios described collecting financial information and another described collecting the user's weight. The "FindMyKid" app allowed a user to set up tracking on someone's phone without that person being aware; such an app could be used by stalkers or abusive partners with physical access to a victim's phone.

3.2 Data and Entity Categories

After participants read the scenario, they were asked to categorize each type of data and third-party entity with which the data would be shared, based on the NTIA MSHP short-form terms. We presented the categories using the exact same wording, in the same order, as used in the NTIA MSHP draft, published April 29, 2013. We also added "None of the Above" and "Not Sure" options.

The NTIA provides both names and explanatory text for each category. In order to gain a better understanding of the utility of including this explanatory text, we conducted our study as a between-subjects survey. Participants in the *terms only* condition were shown only the category names in each scenario; participants in the *parentheticals* condition were also shown the NTIA's explanatory text for each category.

3.3 Pilot Studies

We designed our online survey after conducting several in-person pilots, in which the survey-taker walked through the survey with the researcher and thought out loud. These pilots allowed us to refine

[3] https://www.mturk.com/

our study design. For example, in these pilot surveys, we found that participants were skeptical about the scenarios giving them complete information about what data would be shared, and were apt to make inferences about additional types of data that might be shared. Therefore, we designed the survey so that participants would select a data or entity option only for elements mentioned explicitly in the scenario. Furthermore, we added a notice on every page stating, "The scenarios describe the data collection and sharing completely, so you do not need to guess anything outside of what is described."

3.4 Data Analysis

Each of our participants was shown a sequence of ten scenarios; each scenario had at least one data item and at least one third-party entity with which data is shared. Participants were asked to classify each data item and each entity according to the NTIA categories, or as "None of the Above" or "Not Sure." In all, participants were asked to make 52 categorizations. The data type items we asked participants to categorize are shown in the second column of Table 1, and the third-party entities are shown in Table 2.

We cannot determine how many of our participants were "correct" in each scenario, because we have no ground-truth on which to base that assessment. Instead, our analysis focuses on how consistently our participants categorized the data items and entities. For each data item and entity, we considered the most-commonly selected category to be the *winner*. We then looked at the percentage of participants who selected the *winning* category for each data item and entity, and we call this percentage the *common understanding* for that data item or entity.

We classify each data item and each entity as being either *low common understanding* or *high common understanding*. A data item or entity in which more than 60% of our participants agreed on its categorization is considered to be *high common understanding* (that is, more than 60% of participants categorized it as its *winning* categorization). A data item or entity with 60% or lower categorization agreement is considered to be *low common understanding*. This split is based on what appears to be a bi-modal distribution of common understanding, as shown in Figure 1.

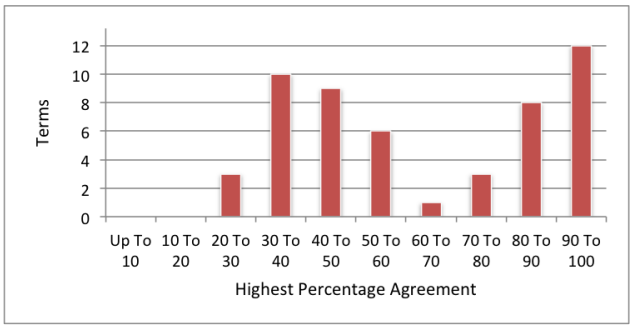


Figure 1: Histogram of percentage responses correct for MTurk Respondents

4. RESULTS

We discuss our participants in Section 4.1. In Section 4.2, we give an overview of our expert results, focusing on categorizations on which the experts did not agree amongst themselves. Then, in Section 4.3, we discuss results from our MTurk participants. We highlight differences between our two study conditions in Section 4.4, and show that while participants in the *parenthetical* condition were more likely to agree on the most-selected category, this was not the case in every scenario. We then discuss differences between expert participants and MTurk participants.

4.1 Participants

The four NTIA MSHP participants in our study, whom we call our *expert participants*, were a diverse group. They each held different professions and represented different stakeholders in the NTIA process; we do not report their demographics to preserve their anonymity. Expert participants were evenly split between our two conditions; because we had only two expert participants in both conditions, we do not report differences based on these conditions for expert participants.

For our MTurk participants, we analyzed data only for participants in the United States who had completed the survey, and we excluded participants who entered gibberish answers for open-text fields. This left us with 791 MTurk participants (375 parenthetical and 416 term-only). The data was collected in two batches, one of 503 responses and one of 288 responses. The second batch included three data entities accidentally omitted from the first. The data entities were: Sports and Health Companies in the Fitness scenario, and AdMeMetric in the Salsa scenario, and are indicated in Table 2. We combine the results from these two batches, except when discussing the three questions that had only 288 responses.

51% of the MTurk respondents were female. Participants ranged in age from 18 to 73 years, with a mean of 33 and a standard deviation of 11 years. Participants took an average of 17 minutes to complete the survey. Every US State was represented. Participants were generally educated: 38% have a Bachelors degree, and another 30% have some college. 82% own a smartphone.

4.2 Expert Participant Results

Our four expert participants showed unanimous agreement on how to classify only half (26) of the data items and entities. There was disagreement on how to classify the other data items and entities. For example, for data being shared, expert participants were divided on whether *Inseam*, *Waist Size*, and *Steps Taken* should be categorized as *health* or *biometrics* data. Further, they were divided on whether *Home Address*, when used for shipping, should be categorized as *None*, *Location*, or *User Files*.

As an example of expert disagreement over how third-party entities should be classified, two expert participants said that a "Traffic Data Company" was none of the categories, one said "Data Analytics," and the fourth said "Consumer Data Resellers." Experts were also evenly split over whether Apple iCloud is "Operating Systems and Platforms" or "None." GreatReading, "an app that organizes local book clubs," yielded a split between "Other Apps" and "Social Networks."

Further, when there was a majority agreement between expert participants, it did not always match the most common categorization by MTurk participants. Table 1 and 2 show in italics when the majority of experts disagreed with the most common categorization by MTurk participants. For example, for some data items, the majority of experts selected *Health*, but the most common MTurk participant categorization was *Biometrics*. In another example, the majority of experts categorized an entity as not belonging to any category, while a majority of MTurk participants selected *Consumer Data Reseller*.

4.3 MTurk Participant Results

This subsection describes results for the 791 participants whose answers were collected through MTurk (our *MTurk participants*). Our results are summarized in Tables 1 and 2. The figures in the Appendix provide more detailed information about how par-

Scenario	Data	Expert Response	Winning Participant Response	Parenthet-icals[1]	Term only[1]	p-value
HipClothes	Inseam	Biometrics (2)	Biometrics	69.1	45.9	<.001*
	Waist Size	Biometrics (2)	Biometrics	69.6	46.4	<.001*
	Clothing Preference	None (3)	None	48	38	<.001*
	Location	Location (4)	Location	91.7	89.9	.494
Salsa	Call History	Browser History (4)	Browser History	88.5	87.5	.463
	Text History	Browser History (4)	Browser History	89.3	90.1	.184
	Video History	Browser History (4)	Browser History	51.5	70	<.001*
	Games Played	Browser History (3)	Browser History	45.9	50.5	.021*
	Photos	User files (3)	User Files	77.6	69.2	.005 *
SuperTax	Photo of W2	Financial Information (3)	User Files	59.2	75.5	.001*
	Salary	Financial Information (4)	Financial Information	92.3	93.3	.502
	Interest Income	Financial Information (4)	Financial Information	92.5	91.8	.066
Fitness	Steps Taken	Health (2)	Biometrics	40.3	46.2	.225
	How Long Slept	*Health (4)*	Biometrics	39.7	44.2	.148
	Weight	*Health (4)*	Biometrics/Health	54.1	50.2	<.001*
	Body Fat	*Health (4)*	Biometrics/Health	53.3	49.5	.005 *
EasyApply	Work History	*None (3)*	None/Financial Information	33.3	34.4	<.001*
	Medical Insurance	Health (3)	Health	85.9	81	.161
	Medical Payments	Health (4)	Health	59.7	52.2	.127
	Number of Children	None (3)	None	41.1	35.1	<.001*
	Marital Status	None (3)	None	43.5	35.1	<.001*
	Income	Financial Information (4)	Financial Information	88.5	91.6	.063
CallCalendar	Call Time	Browser History (4)	Browser History	91.2	86.8	.222
	Call Duration	Browser History (4)	Browser History	90.1	86.3	.189
	Name from Contact List	Contacts (3)	Contacts	71.2	82.5	<.001*
GoodDriver	GPS Location	Location (4)	Location	94.1	94.7	.788
	Gyroscope Bumps	None (3)	None	33.6	33.9	.252
FindMyKid	Location	Location (4)	Location	94.1	94.7	.176
iTunes	Credit Car Info	Financial Information(3)	Financial Information	96	92.3	.304
	Song and Artist Names	*None (3)*	User Files	57.1	53.1	.443
Bookstore	Book Title	None (4)	None	34.4	36.1	.502
	Home Address	None (2)	Location	49.1	58.7	.008*
	Credit Card	Financial Information(4)	Financial Information	94.1	91.1	.092

[1] Participant level of common understanding for winning term by condition
* Difference between conditions is significant at p<.05 with χ^2 test Benjamini and Hochberg FDR correction.

Table 1: Data Type categories selected for each term by NTIA experts and MTurk participants. For this table, the categories "Health, Medical or Therapy Information" has been abbreviated to "Health" and "Browser History and Phone or Text Log" to "Browser History." In the expert column, we show all categories selected by two or more experts, with the number of experts that selected each category in parenthesis. The terms in which the majority of experts and participants differed are in italics. If the conditions in the participant study had different winners, both are shown in the participant column.

ticipants categorized data items.

As described in Section 3, we divided each data item and entity into *high common understanding* and *low common understanding*, with the former having greater than 60% agreement on its categorization. Figure 1 shows a histogram of the data items and entities with the highest percentages of categorization agreement among our MTurk participants. For 12 total data items and entities, at least 90% of participants agreed on the categorization. Conversely, for 28 total data items and entities (over half), 60% or fewer of our MTurk participants agreed on their categorization (they had low common understanding).

4.3.1 High Common Understanding

All of the data items in which the name of the data item closely aligned with an NTIA category had high common understanding. This included all examples of GPS Location, found in the HipClothes, GoodDriver, and FindMyKid scenarios. In addition, most data related to financial information had high common understanding, including income, salary, and credit card information. On the other hand, *work history* and *photo of W-2* had low common understanding. Further, the data category *Browser History and Phone or Text Log* was generally understood to include call history, text history, call time, and call duration. However, participants did not agree on whether this category included video history or games played.

Most entities had low common understanding, though two had high common understanding: government-related entities and Facebook as a *Social Network*.

Scenario	Data	Expert Response	Winning Participant Response	Parenthe-ticals[1]	Term only[1]	p-value
HipClothes	OtherClothingStores	*None (3)*	Consumer Data Reseller/None	31.5	33.3	<.001*
Salsa	Advertising Companies	Ad Networks (4)	Ad Networks	80.5	79.2	.520
	AdmeMetric[2]	*Consumer Data Reseller (3)*	Consumer Data Reseller	43.8	38	.086
SuperTax	State Agency	Government Entity (4)	Government Entity	93.9	96.2	.465
	Federal Agency	Government Entity (4)	Government Entity	94.7	95.4	.518
Fitness	Sports Companies[2]	*None (3)*	Consumer Data Reseller	38.4	26.8	.027
	Health Companies[2]	*None (3)*	Consumer Data Reseller	31.5	24.6	.022*
EasyApply	State Agency	Government Entity (4)	Government Entity	92	93.3	.208
CallCalendar	Carrier	Carrier (4)	Carrier	90	88.2	.173
	Google Calendar	Other Apps (3)	Other Apps	47.1	51	.066
GoodDriver	Traffic Data Company	None (2)	Data Analytics	59.7	58.4	.770
	Car Insurance	*None (4)*	Consumer Data Reseller	35.7	26	<.001*
	Car Rental	*None (4)*	Consumer Data Reseller	36.3	25.7	<.001*
FindMyKid	Parents Phone	None (3)	None	34.4	46.6	.034
	Local Police	Government Entity (4)	Government Entity	80	85.3	.333
iTunes	Facebook	Social Network (3)	Social Network	89.6	92.1	.714
	Apple iCloud	OS and Platforms (2), None (2)	OS and Platforms	37.9	34.9	.799
Bookstore	Facebook	Social Network (3)	Social Networks	88.8	90.6	.566
	GreatReading	Social Network (2), Other Apps (2)	Other Apps	37.6	40.1	.410

[1] Participant level of common understanding for winning term by condition.
[2] 288 Responses Only
* Difference between conditions is significant at p<.05 with χ^2 test and Benjamini and Hochberg FDR correction.

Table 2: Third-Party Entities categories selected for each term by NTIA experts and MTurk participants. In the expert column, we show all categories selected by two or more experts, with the number of experts that selected each category in parenthesis. The terms in which the majority of experts and participants differed are in italics. If the conditions in the participant study had different winners, both are shown in the participant column.

4.3.2 Low Common Understanding

MTurk participants generally had low understanding (60% or lower agreement) when the most common categorization for a given data item was *Biometrics* or *Health, Medical or Therapy Information* (*Health Information*). For example, over 60% of participants did agree that *Medical Insurance* fell under the latter category; however, while medical payments were most commonly classified as *Health Information*, less than 60% of MTurk participants made this categorization. Furthermore, all of the data items pertaining to body measurements — inseam, waist size, steps taken, amount of sleep, weight, and body fat — were typical divided between *Biometrics* and *Health Information*.

Categorization of third-party entities was less consistent than for data items. While government entities tended to have high common understanding, entities such as consumer data resellers, data analytics, operating systems, and other apps did not reach 60% categorization agreement among our MTurk participants. Over 60% of our MTurk participants did categorize Facebook as a *Social Network*; however, less-known social networking GreatReading was not met with high common understanding.

Participants often thought that third-party entities that purchase data from the app were *Consumer Data Resellers*, even if there was no indication that those companies resold the data. In scenarios in which the third-party offered discounts (for example, a health company or clothing store), participants were particularly likely to consider them resellers.

The description of *Data Analytics Providers* is sufficiently vague ("Companies that collect and analyze your data") that several of the example entities could have fallen under this category. However, there was only one entity that many participants described as a *Data Analytics Provider*: The *Traffic Data Company* that "specializes in traffic data so that congestion and problems can be predicted and analyzed." However, even this category had low common understanding, falling just below 60%.

When any data element did not fall into one of the given categories — where *None* would have been the most appropriate response — we also see low common understanding. Examples of this include number of children, clothing preference, and gyroscope bumps. While it may be impossible to come up with a complete taxonomy of all types of data, our results indicate that participants expected all the data to fall into at least one of the categories.

4.3.3 Ambiguous Data Types

Many of our scenarios were based on real-world apps. We did not intend to confuse participants, but the scenarios did include several data items and entities that were ambiguous, leading to low common understanding. In this section, we describe some of these ambiguous terms and how they might reasonably be considered to belong to multiple categories.

The *SuperTax* scenario collects a photo of a W-2 file, a yearly earning statement in the US, typically used for tax reporting. This could reasonably be classified as either *Financial Information* or

User Files: The W-2 and its information are financial, but the photo itself is a user file.

The *Bookstore* app, which allows users to purchase books and share that information on social networks, collects "your home address where the book will be shipped." Since the home address information is entered by the users, and not by GPS sensors, it may or may not fall under the *Location* category.

Both the Apple iCloud and the GoogleCalendar entities are ambiguous in that they are both apps created by major smartphone platform and OS vendors. Google creates the Android operating system, so sharing information with the Google Calendar could be interpreted as sharing with Google, the OS creator. The same is true for Apple iTunes, though in this scenario we described iTunes as being available on Android (not the Apple OS, so it was not being provided by the user's OS provider).

The *GreatReading* app is described as "an app that organizes local book clubs." This could reasonably be described as either a *Social Network* since it assists in creating social groups, or an *Other App*, as it is could be described by the parenthetical text for *Other Apps*: an app from a, "company that the consumer may not have a relationship with."

Weight, body fat, inseam, and waist size are "information about your body" and could therefore be considered *Biometrics*. However, they can also indicate health status, such as obesity, and therefore can reasonably be considered *Health, Medical or Therapy Information*. Arguably, in the context of a clothes shopping app, inseam and waist size would not be considered either *Biometrics* or *Health*.

4.4 Differences between conditions

In this section, we look at differences between our two conditions for MTurk participants. Recall that participants in the *parentheticals* condition were shown a brief description for each category, taken directly from the draft Code, while participants in the *terms only* condition were not given a definition for the categories. We tested whether participants in our two conditions classified data items and entities significantly differently. We used a χ^2 test, which examines the distribution of responses across all categories. Tables 1 and 2 show the p-values for difference between conditions for each data item and entity. Overall, 14 of the 33 data items had significant differences between conditions, and 4 of the 19 entities had significant differences (with $p<.05$ and Benjamini and Hochberg False Discovery Rate correction [8]).

Below, we discuss some of the more salient differences between conditions. Section 4.4.1 discusses data items and entities in which the category with the most agreement differed between conditions. Section 4.4.2 discusses data items and entities in which the *parenthetical* condition had less agreement on most common category than the *terms only* condition.

4.4.1 Conditions yield different winners

In four cases, participants in different conditions selected different categories most frequently: *Weight* and *Body Fat* in the Fitness scenario, *Work History* in the EasyApply scenario, and *Other Clothing Stores* in the HipClothes scenario. The difference between the conditions was significant in all cases ($p<.05$). These terms also had low common understanding overall, meaning that in no case did 60% of participants agree on a category. We discuss each term below and propose explanations for why the results were different.

The Fitness scenario describes collecting both weight and body fat. In the parenthetical condition, more participants felt that weight and body fat were *Biometrics*. Those in the *terms only* condition selected *Health, Medical or Therapy Information* more frequently. The parenthetical text for biometrics includes the explanation "information about your body," which is vague enough to encompass weight and body fat, as well as other types of medical information. Interestingly, the NTIA experts all selected the *Health* category, suggesting that the Biometrics text may be misleading.

Participants in the *parentheticals* condition mostly categorized *Work History* as belonging to *None* of the categories, while those in the *terms only* condition selected *Financial Information*. We hypothesize that parenthetical text for Financial Information — "Includes credit, bank and consumer-specific financial information such as transaction data" — is so specific that it causes participants to rule out the financial category. This may be the intended result; most NTIA experts categorized *Work History* as *None* of the categories.

Participants in the *parentheticals* condition perceived *Other Clothing Stores* to be *Data Resellers*, while the *terms only* participants selected *None* most frequently. This was also one of the entities with the least common understanding. The parenthetical describes the purpose of Consumer Data Resellers collecting the information as "offering products and services that may interest you," which the other clothing stores would do in the given scenario. We note that this is the only category whose parenthetical includes a description of the purpose of the data. Most NTIA experts classified this as *None*.

4.4.2 Parentheticals that reduced agreement

For the majority of data items and entities in which participants in both conditions agreed on the classification, participants in the *parentheticals* condition were more likely to agree with the most-common categorization. When the two conditions did not agree on classification, *parentheticals* participants were more likely to agree with their most-common categorization than *terms only* with theirs — 7 out of 11 for data items and 4 out of 5 for entities. In other words, the presence of parentheticals to define the terms generally lead to more agreement. In this section, we describe those data items and entities for which having parentheticals appears to have reduced agreement with the most-popular categorization.

Contacts. In the *CallCalendar* scenario, a name from the contact list is collected from the apps. In both conditions, the majority of participants agreed that this fell under the *Contacts* category, though with more agreement from participants who did not see the parentheticals (*parentheticals*=71.2%, *terms only*=82.5%). Further, in the *parentheticals* condition, 22.9% thought this would qualify as *Browser History and Phone or Text Log*, while only 8.4% in the terms only condition opted for the that category. The parenthetical text for this second choice is, "(A list of websites visited, or the calls or texts made or received.)" Participants were presumably uncertain whether information about the calls made or received includes the name. It may be unclear to users where and how the name of the contact is stored in their phone's architecture, and what exactly is included in the log of calls and texts.

Games Played and Video History. The *Salsa* scenario describes an app that allows users to make video calls and share games, and shares information about video history and games played. Our results show participants were divided about whether this should be categorized as *Browser History and Phone or Text Log* or as a *User File*. Participants may have debated whether video history and game-playing belong in the log category because information about the file, and not the file itself, was shared. The parenthetical text steered more users to selecting *User File*: In the *parentheticals* condition, 39.7% of participants considered games played and video history to be *User files*, compared to 22% in the term condition. Since the parenthetical text for *User Files* includes the word

"video," it may have led users who saw that text to select *User File*, while users who did not see the parenthetical text may have focused on the distinction between sharing logs and files.

Home Address. The *Bookstore* app collects the user's home address in order to ship a book. The parenthetical text for location describes it as "(precise past or current location and history of where a user has gone.)," which may indicate to participants that the automatic collection of where they are counts as location. The users in the parenthetical condition who did not select *Location* did not have clear agreement on how to categorize it, and were more likely to say *Not Sure* or *none* than a category. This indicates the *Location* category should be more clear about whether it includes user-entered data, or only that collected by location sensors.

4.5 Demographic Factors

We looked at whether any of the demographic factors significantly affected participants' responses. We looked at whether owning a cellphone, education, or knowing a programming language had any affect on choices for each item (χ^2 test with the conservative $p < 0.001$ due to multiple tests). Only one term showed a significant difference among these factors: Education made a difference in how participants categorized *local police*. Participants with lower education (some high school or high school), more frequently selected *none* as the appropriate category for local police, while higher-education participants were more likely to recognize that they are *Government Entities*.

4.6 What Categories are the most sensitive

We asked participants, "Which of the following types of data would you want to know about an app collecting?" and "Which of the following entities would you want to know if an app shared data with?" The response options were, "Want To Know," "Don't Care," and "It Depends." The response options were randomized between participants to avoid bias. In this analysis, we look at responses from participants who provided exactly one response to the question.

Among the types of data about which we asked participants, participants most wanted to know when "Financial Information" (89.5%) and "Health, Medical, or Therapy Information" (86.1%) was disclosed. The results for each datum are shown in Table 3.

Table 3: The percentage of participants who responded with "Want To Know" to the question "Which of the following types of data would you want to know about an app collecting?" for each entity.

Entity	% want to know
Financial Information	89.51%
Health, Medical or Therapy.Information	86.09%
Browser History and Phone or Text Log	82.55%
User Files	80.03 %
Contacts	79.77%
Location	71.05%
Biometrics	68.65%

Among the entities about which we asked, participants most wanted to know about data sharing with government entities (79.7%), followed by consumer data resellers (77.4%). However, for each of the entities about which we asked, over half of participants wanted to know when data would be shared with that entity. This is shown in Table 4.

Table 4: The percentage of participants who responded with "Want To Know" to the question "Which of the following entities would you want to know if an app shared data with?" for each entity.

Entity	% want to know
Government Entities	79.65%
Consumer Data Resellers	77.37%
Social Networks	74.97%
Ad Networks	72.31%
Data Analytics Providers	69.03%
Carriers	65.61%
Other Apps	63.34%
Operating Systems and Platforms	58.15%

It is worth noting that, while some entities and information appear more sensitive than others, over half of participants want to know about disclosure in each of the cases asked about. We also examined whether responses differed by condition or demographics. We examined the responses of our participants across three factors (omitting participants who did not indicate one): participant gender, the condition to which the participant was assigned, and whether the participant indicated using a mobile device. Separately for each of these three factors, we compared whether there was a significant difference in the proportion of participants who responded "Want To Know" to each entity and datum using a χ^2 test. All p-values, separately for each factor, were corrected using Holm-Bonferroni correction.

We found no significant difference in response to entities or data when we look at participants by gender or by condition ($p > .05$). However, even with correction, if we compare participants who do and do not have a mobile device, we see a significant difference ($p < .05$) in wanting to know about disclosure to operating system and platform (p=.002), ad networks (p=.008), carriers (p=.015), consumer data resellers (p=.012), other apps (p<.001), and social networks (p=.022). In each case, participants who did not use mobile devices were significantly more likely to want to know about disclosure than those who do use mobile devices. We cannot determine, based on this data, whether users who are more privacy-sensitive are less likely to use a mobile device, or whether using a mobile device makes users less privacy-sensitive.

5. LIMITATIONS

This survey is designed to measure whether participants understand the NTIA categories by giving them an explanation of an app, and an explanation of the data shared, including such details as with whom the data is shared and the purpose of sharing the data. Participants may see more information than they would in practice. Our results for understanding, therefore, may be an overestimate of true understanding in practice. Further, as stated above, while we can measure the extent to which participants agree on how to categorize a given data item or entity, it is impossible to determine whether that categorization is "correct."

The task presented to survey participants more closely resembles a realistic task for an app developer than a user. A more realistic user task might be to provide a notice that uses the terms from the Code and to ask users what data they think an app is collecting and with what entities they believe it is shared. However, this is actually an even harder task because each data category could potentially cover many types of data, and it is not necessarily possible to infer

what data is collected from a very brief description of an app.

This survey is limited to testing the particular terminology defined by the NTIA code. While the results indicates some categories are poorly understood, we do not test alternate wordings. Therefore, we are unable to offer better terminology; that may be an area for future work.

Furthermore, while we tried to present a broad swath of scenarios, we could not create a study that would present all possible scenarios to participants. There may be many more types of data that are ambiguous to users, or examples that are more clear than those in this survey.

6. DISCUSSION

The NTIA MSHP has selected several categories of data sharing about which mobile users should be informed on short-form privacy notices. Our investigation looked at user and expert understanding of these categories. Our survey found that the categories were not well understood by our participants. Of the 52 examples of data sharing given in our scenarios, participants showed low (less than 60%) common agreement for 23 of them. Furthermore, our expert participants also disagreed among themselves on how to categorize some of the examples, and had different majority responses from the study participants for 13 examples. We find that the *Biometrics* and *Health, Medical or Therapy Information* categories were especially prone to disagreement. Further, participants struggled to categorize many of the third-party entities. In particular, participants expected more entities to be categorized as "Consumer Data Resellers" than the experts expected.

Our main finding is that the current set of NTIA categories does not appear to offer a high level of transparency for users. The lack of common understanding, even among experts, also suggests that app developers may have trouble generating accurate notices using these terms and definitions. Next, we will discuss our main findings and offer our recommendations.

Parentheticals Help (Sometimes). In most cases, the difference between the parenthetical condition and the term-only condition was not significant. When it was significant, the parenthetical usually resulted in greater agreement with the most-popular category. However, this was not always the case; some parentheticals appeared to confuse our participants. For example, the parenthetical text for *Browser History and Phone or Text Log*, *User File*, and *Location* appear to need some improvement to make them more useful to users.

Better Definitions Are Needed. Some categories were not well understood, both by participants and by NTIA experts. Therefore, we recommend that the Code provide further guidance on how to interpret the categories. This may include definitions and examples, including edge cases. In particular, guidance is needed for *Biometrics*, *Health, Medical or Therapy Information*, and all of the third-party entities except *Government Entities*. Further, experts should clarify whether location includes only information from sensors (such as GPS) or user-entered information (such as home address).

Ambiguous Data Items Need Clarification. Several types of data items were confusing to participants. Some data items could reasonably be classified in two categories (e.g., a photo of a W-2 is both a user file and financial information). Some data items require an understanding of the platform architecture in order to classify them correctly (e.g., whether a contact name is stored in a call log or in a user file). In several of these cases, participants who saw the parenthetical text had less agreement than those who saw only the terms, indicating that the short phrases created confusion instead of clarification.

For improved transparency on ambiguous or poorly understood data types, we recommend that implementors of the short-form specify the data being collected. For example, a short form notice with the text "Health, Medical or Therapy Info: how many steps you have taken, how long you've slept, weight, and body fat" may be more clear to users than "Health, Medical or Therapy Info." Future research should investigate whether specific information is better understood, and whether implementors of a short-form notice should specifically say what is being collected instead of, or in addition to, the parenthetical text.

Third-Party Entities Are Poorly Understood. Many of the third-party entity categories were confusing to participants. For example, participants typically categorized any entity that purchased information as a *Consumer Data Reseller*. Our results show that participants struggled with many of the third-part entities, except *Government* and *Carriers*. In these cases, parenthetical text did not add much clarification.

On the other hand, specificity about third-party entities will only be helpful if users recognize the name of the entity. Previous research suggests that users are not familiar with the names of advertisers, data resellers, or analytics companies [6, 25]. Further research is needed on describing third-party entities in a transparent way.

Uncategorized Data and Entities. There are some privacy-sensitive data that do not fit into any of the existing categories (and therefore need not be indicated in a short-form notice). These include identifying information such as user name, phone id, or SSN. Since not all data sharing falls into a category covered by the short-form notice requirements, the app may be sharing data without notifying the user through the short form. Our results show that participants did not often categorize data and entities as *None*, and preferred to place data in one of the categories. This suggests participants believe the categories encompass all possibilities. Therefore, information about the smartphone notices should emphasize that the short form does not notify users about all types of data sharing.

Further User Testing is Needed. By providing realistic scenarios and asking survey participants to categorize data items shared and entities with which data is shared, our work highlights that the categories are not well understood. However, this is not typical task flow for users, and we did not test actual short-form notices. This work is a first step and indicates that more work is needed to develop a well-understood notice with categories and definitions that will be generally understood by American smartphone users.

Acknowledgements

We thank all the NTIA stakeholders who participated in the survey. This research was funded in part by NSF grant DGE0903659.

7. REFERENCES

[1] *Evolution of a Prototype Financial Privacy Notice: A Report on the Form Development Project*. FTC and Kleimann Communication Group, Inc, 2006.

[2] Final model privacy form under the Gramm-Leach-Bliley act; final rule. *Federal Register*, 74(229):62890–62994, Dec. 2009.

[3] *Consumer Data Privacy In A Networked World: A Framework For Protecting Privacy And Promoting Innovation In The Global Digital Economy*. White House, 2012.

[4] G. J. Annas. Hipaa regulations âĂŤ a new era of medical-record privacy? *New England Journal of Medicine*, 348(15), April 2003.

[5] C. Arthur. Is your private phone number on facebook? probably. and so are your friends'. The Guardian, http://www.guardian.co.uk/technology/blog/2010/oct/06/facebook-privacy-phone-numbers-upload, October 6, 2010.

[6] R. Balebako, J. Jung, W. Lu, L. Cranor, and C. Nguyen. Little brothers watching you: Raising awareness of data leaks on smartphones. In *Proc. SOUPS*, 2013.

[7] M. Benisch, P. G. Kelley, N. Sadeh, and L. F. Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal Ubiquitous Comput.*, 15(7):679–694, Oct. 2011.

[8] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 289–300, 1995.

[9] A. J. Berinsky, G. A. Huber, and G. S. Lenz. Using mechanical turk as a subject recruitment tool for experimental research. 2011.

[10] M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Persp. Psych. Sci.*, 6(1):3–5, 2011.

[11] A. Cavoukian. Privacy and biometrics. Information and Privacy Commissioner, Ontario, September 1999.

[12] F. T. Commission. *Mobile Privacy Disclosures, Building Trust Through Transparency*. Federal Trade Commission, 2013.

[13] L. Cranor. *Web privacy with P3P*. O'Reilly Media, Inc., 2002.

[14] A. Felt, S. Egelman, M. Finifter, D. Akhawe, and D. Wagner. How to ask for permission. *HOTSEC 2012*, 2012.

[15] A. Felt, S. Egelman, and D. Wagner. I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns. In *Proc. SPSM*, 2012.

[16] A. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: User attention, comprehension, and behavior. *Proc. of SOUPS*, 2012.

[17] R. Gross and A. Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.

[18] C. M. Hastak, M. Online behavioral advertising "icon" study. *Future Of Privacy Forum*.

[19] P. G. Ipeirotis. Demographics of Mechanical Turk. Technical Report CeDER-10-01, New York University, 2010.

[20] B. Isaacson. Immersion, an mit media lab creation, uses email metadata to map your connections. Huffington Post, http://www.huffingtonpost.com/2013/07/10/immersion-email-metadata_n_3567984.html, July 10, 2013.

[21] P. Kelley, L. F. Cranor, and N. Sadeh. Privacy as part of the app decision-making process. In *Proc. of CHI 2013*, 2013.

[22] P. G. Kelley, L. Cesca, J. Bresee, and L. F. Cranor. Standardizing privacy notices: an online study of the nutrition label approach. Proc. of CHI 2010, pages 1573–1582. ACM, 2010.

[23] B. Krishnamurthy and C. E. Wills. Characterizing privacy in online social networks. In *Proceedings of the first workshop on Online social networks*, WOSN '08, pages 37–42, New York, NY, USA, 2008. ACM.

[24] P. G. Leon, J. Cranshaw, L. F. Cranor, J. Graves, M. Hastak, B. Ur, and G. Xu. What do online behavioral advertising privacy disclosures communicate to users? In *Proc. WPES*, pages 19–30, 2012.

[25] P. G. Leon, B. Ur, R. Balebako, L. F. Cranor, R. Shay, and Y. Wang. Why Johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising. In *Proc. CHI*, pages 589–598, 2012.

[26] P. G. Leon, B. Ur, Y. Wang, M. Sleeper, R. Balebako, R. Shay, L. Bauer, M. Christodorescu, and L. F. Cranor. What matters to users? factors that affect users' willingness to share information with online advertisers. In *Proc. SOUPS*, 2013.

[27] A. Levy and M. Hastak. Consumer comprehension of financial privacy notices: A report on the results of the quantitative testing. *Federal Trade Commission*, pages 62890–62994, Dec. 2008.

[28] M. Marlinspike. Why 'I have nothing to hide' is the wrong way to think about surveillance. Wired, June 13, 2013.

[29] A. M. McDonald and L. F. Cranor. Americans' attitudes about internet behavioral advertising practices. In *Proc. WPES*, 2010.

[30] I. Muslukhov, Y. Boshmaf, C. Kuo, J. Lester, and K. Beznosov. Understanding Users' Requirements for Data Protection in Smartphones. *ICDE 2012*, pages 228–235, Apr. 2012.

[31] G. Paolacci, J. Chandler, and P. Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419, 2010.

[32] S. Prabhakar, S. Pankanti, and A. Jain. Biometric recognition: security and privacy concerns. *IEEE Security & Privacy*, 1(2):33–42, 2003.

[33] D. Smilkov, D. Jagdish, and C. Hidalgo. Immersion. https://immersion.media.mit.edu/, 2013.

[34] D. Solove. 'I've got nothing to hide'and other misunderstandings of privacy. *San Diego law review*, 44, 2007.

[35] D. J. Solove. Five myths about privacy. Washington Post, June 13, 2013.

[36] A. P. R. C. T. Survey. Public split over impact of nsa leak, but most want snowden prosecuted. June 17, 2013.

[37] E. Toch, J. Cranshaw, P. H. Drielsma, J. Y. Tsai, P. G. Kelley, J. Springfield, L. Cranor, J. Hong, and N. Sadeh. Empirical models of privacy in location sharing. In *Proc. Ubicomp*, pages 129–138, 2010.

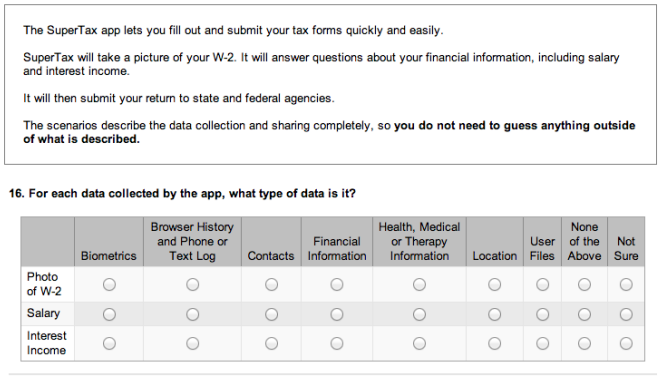
[38] J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessey. Americans reject tailored advertising and three activities to enable it, September 2009. http://ssrn.com/abstract=1478214.

[39] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang. Smart, useful, scary, creepy: perceptions of online behavioral advertising. In *Proc. SOUPS*, 2012.

[40] J. Urban, C. Hoofnagle, and S. Li. Mobile phones and privacy. *UC Berkeley Public Law Research Paper*, 2012.

[41] J. Valentino-Devries, J. Singer-Vine, and A. Soltani. Websites vary prices, deals based on users' information. Wall Street Journal, http://online.wsj.com/article/SB10001424127887323777204578189391813881534.html, December 24, 2012.

APPENDIX



The SuperTax app lets you fill out and submit your tax forms quickly and easily.

SuperTax will take a picture of your W-2. It will answer questions about your financial information, including salary and interest income.

It will then submit your return to state and federal agencies.

The scenarios describe the data collection and sharing completely, so **you do not need to guess anything outside of what is described.**

16. For each data collected by the app, what type of data is it?

	Biometrics	Browser History and Phone or Text Log	Contacts	Financial Information	Health, Medical or Therapy Information	Location	User Files	None of the Above	Not Sure
Photo of W-2	○	○	○	○	○	○	○	○	○
Salary	○	○	○	○	○	○	○	○	○
Interest Income	○	○	○	○	○	○	○	○	○

Figure 2: Screenshot of one scenario in the terms-only condition, showing how participants were asked to categorize the data types.

The text used to describe each scenario is presented here. A complete copy of the survey is available at `http://alturl.com/vbmki`. **HipClothes** The HipClothes app recommends clothing to you, and also shows you the stores closest to your location where you can find the clothes in your size.

The HipClothes app requests your inseam, waist size, and clothing preferences.

It also will share your information with two other clothing store chains that are owned by the same company.

Salsa The Salsa app allows you to make video calls, phone calls, text messages and include games, and picture sharing. Salsa stores your history in your online Salsa account.

Salsa collects your call, video, and text history, including copies of which pictures were shared, and information about which games were played.

Salsa will shows ads, and does so by sharing your information with advertising companies. Salsa will also share your information with AdMeMetric, which will resell information to companies that will provide you with coupons.

SuperTax The SuperTax app lets you fill out and submit your tax forms quickly and easily.

SuperTax will take a picture of your W-2. It will answer questions about your financial information, including salary and interest income.

It will then submit your return to state and federal agencies.

Fitness app The Fitness app integrates with your FitMonitor (FitMonitor is a special pedometer and activity monitor, purchased separately) to allow you to track and improve your fitness activities and level.

Fitness app will collect information on how many steps you have taken, how long you've slept, and allow you to enter you weight and body fat.

Fitness app will notify sports and health companies if you achieve certain goals, and these companies will send you valuable coupons as awards.

EasyApply This EasyApply app can be used to apply for government benefits such as Child Health Plus, Family Health Plus, Medicaid, and the Family Planning Benefit Program.

You will enter your income, work history, and whether you have any existing medical insurance and medical payments. You will also supply information about how many children you have, and your marital status.

EasyApply will save this information, and will submit your application to the state agency who will determine what benefits you and/or your children are eligible for.

CallCalendar The CallCalendar is an app that logs your phone activity and adds it to your Google Calendar.

You can select the type of calls to log (incoming, outgoing, and missed) and the calendar to log them in. CallCalendar will save your call log, including time, duration, and name of the person from the contact list.

CallCalendar will share your phone call information with your cellphone carrier so your cellphone carrier can improve its services. It will also share this information with Google Calendar.

GoodDriver The GoodDriver app is an application for your smartphone that will keep you and others safe on the roads.

It will use your GPS to detect your speed and location. It will use your gyroscope to detect road conditions (such as bumps). It will use your speed to tell you about traffic congestion and problems.

It shares information with that a company that specializes in traffic data so that congestion and problems can be predicted and analyzed. Your driving information will be sold to car insurance companies and car rental companies, who will offer you better rates for good driving.

FindMyKid The FindMyKid app can be installed on your childÕs phone to track its location and show you his or her whereabouts.

Without interrupting your child, you can see where he or she is at any time from your phone or on-line. FindMyKid app collects your child's location from his or her phone.

This app shares your child's location information with you (your phone). It will also share with local police, in case of emergency, with a simple button interface.

iTunes The popular iTunes app for playing music, and developed by Apple, is now available on Google Android phones.

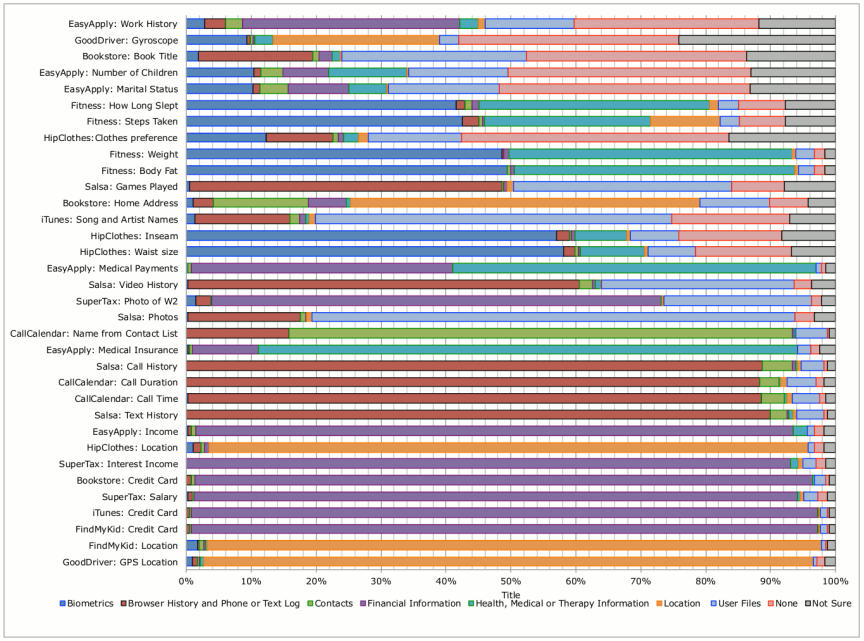
You can enter song and artists names, which is stored by Apple. You can make purchases by entering your credit card information, which is saved by Apple for further purchases.

iTunes will share information about what you are playing with Facebook. Your songs are stored on the Apple iCloud service.

Bookstore The Bookstore app allows you to purchase books from your cell phone.

You will pay using a credit card and enter your home address where the book will be shipped. Bookstore app will save this information in your online account so that you can use Bookstore online or from any device.

It also shares information about your purchase with Facebook and GreatReading (an app that organizes local book clubs).



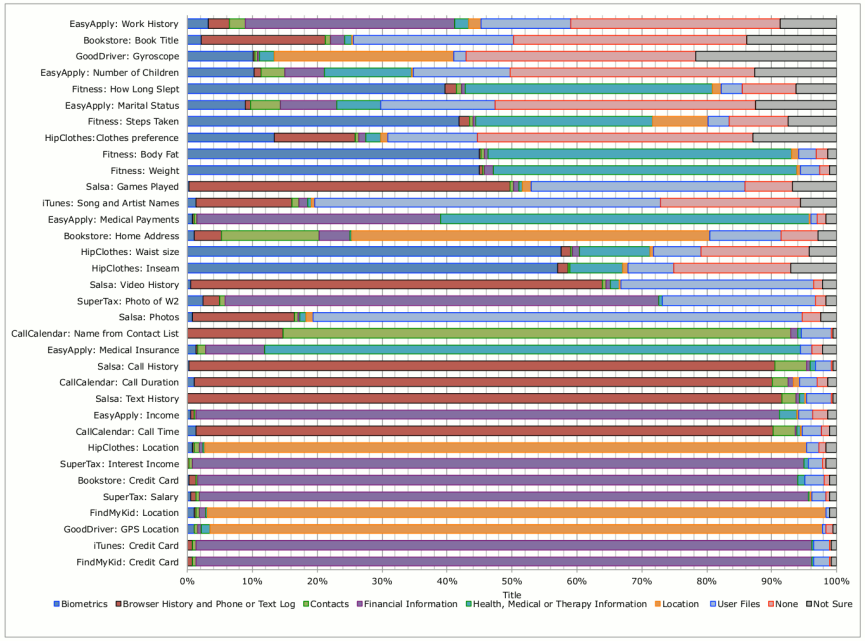
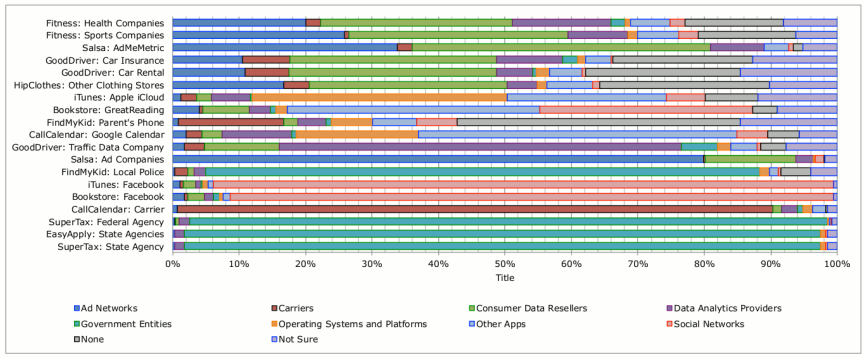


Figure 3: Participants' categorization of data types in parenthetical condition (left) and term-only condition (right).



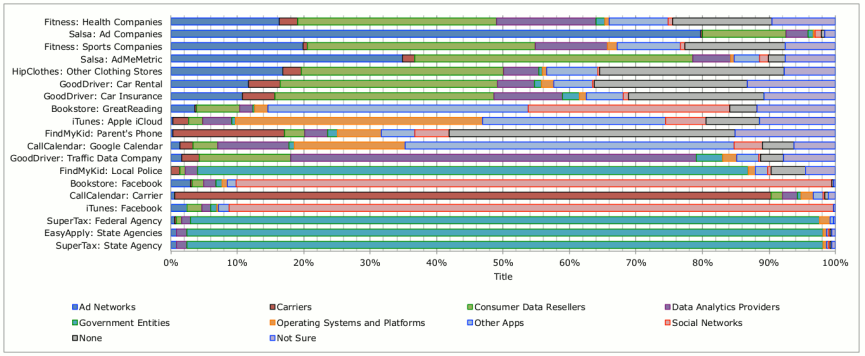


Figure 4: Participants' categorization of third-party entities in parenthetical condition (left) and term-only condition (right).