

# Donating Context Data to Science: The Effects of Social Signals and Perceptions on Action-Taking

YONG LIU<sup>1,\*</sup>, DENZIL FERREIRA<sup>2</sup>, JORGE GONCALVES<sup>2</sup>, SIMO HOSIO<sup>2</sup>,  
 PRATYUSH PANDAB<sup>2</sup> AND VASSILIS KOSTAKOS<sup>2</sup>

<sup>1</sup>*Department of Information and Service Economy, Aalto University School of Business, Helsinki, FI-00076, Finland*

<sup>2</sup>*Center for Ubiquitous Computing, University of Oulu, Oulu, FI-90014, Finland*

\*Corresponding author: yong.liu@aalto.fi

It is becoming increasingly easy for researchers to develop context-aware applications for smartphones. A perennial challenge, however, is to convince a large number of people to install them and donate contextual data for scientific purposes. Our empirical study seeks to address this challenge by investigating how people's perception and attitude affect their willingness to donate context data to researchers and quantifies the effects of social signals on donation action-taking. Our findings indicate that the *perceived need for donation* and *perceived organization reputation* are key determinants in deciding whether to donate: people with altruistic personality do not necessarily donate if they cannot see the need to take an action. Furthermore, we provide evidence that even if people indicate a willingness to donate, they are hesitant to take action towards donating data unless catalysts like social signals (hints about the actions of others) are present.

## RESEARCH HIGHLIGHTS

- Social signal affects people's actual human-computer interaction data donation behavior.
- Participants who saw social signal were five times more likely to become donors
- Perceived organization reputation and need to donate are key determinants.

*Keywords: empirical study; HCI; data collection; data donation; context-aware applications; privacy*

*Editorial Board Member: Dr. Javier Bargas-Avila*

*Received 21 July 2015; Revised 15 March 2016; Accepted 31 March 2016*

## 1. INTRODUCTION

An important challenge the human-computer interaction (HCI) research community needs to address is recruiting and motivating substantial number of participants to donate potentially personal data to advance scientific research. As early as 1998 (Hilbert and Redmiles, 1998), researchers elaborated on this challenge of large-scale data collection noting that '(data collection) can be difficult [in relation to the then-emerging World-Wide-Web] due to the distribution of users, the time and labor involved in collecting data, the lack of scalable tools for automatic data collection, and the lack of proper incentives to support high-quality voluntary data collection on

the part of users.' They go on to note that as a result, most studies 'are limited to small scale tests in the [...] lab, feedback from beta testing is typically reported manually by beta testers themselves, ...[and] the quality and quantity of data is limited.'

Although the above comments were published 15 years ago, researchers still face the same challenge. This is particularly true for researchers who develop systems that are deployed in 'living' settings and have the potential to collect very rich contextual data that volunteers generate. This challenge resonates with participatory sensing, and while some work has considered recruitment of volunteers based on their

geographical context (Reddy *et al.*, 2010), a more grounded analysis of why people may choose to donate contextual data is lacking. Many HCI studies published today are still conducted with a few volunteers and participants, over short period of time, raising extent and replication concerns in our field (Hombæk *et al.*, 2014).

It is still difficult to request personal and contextualized data from people despite the technological advances in recent years. This is mostly because we still do not fully understand how we sustain motivation or encourage initial participation, especially for longitudinal studies. Therefore, a clear understanding and an actionable strategy on data collection would substantially benefit our discipline in generating high-quality scientific contributions by relying on larger cohorts and richer data sets. We begin to establish such an actionable strategy in this article by drawing parallels between people giving contextual data to scientists versus people donating in general (e.g. blood, money and time). We focus on contextual data that people generate as a byproduct of their daily routine and not data generated through some type of focused labor (e.g. crowdsourcing). Our aim is to use the previous literature on donation as a lens through which we can inspect people's perspectives and actions in relation to 'donating' their contextual data.

We describe an online experiment that investigates the factors driving people to actually donate their contextual data to scientific research via a data collection mobile application. We define contextual data as any sensor data that can be captured by a person's smartphone, e.g. used applications, social network activities, location history, communication patterns, just to name a few. This data is therefore contextualized by the everyday behavior of this person.

We distributed an online survey to collect participants' demographic features, perceptions and willingness towards data donation. This enabled us to model their donation decision-making process. Immediately after participants completed the survey, an opportunity was given to them to actually donate contextual data by installing a mobile application on their own device. At this point, we manipulated our experiment by randomly assigning participants to different conditions to examine how participants would respond to 'social signals', i.e. hints about how others acted on this specific donation. We then analyzed how many participants actually donated data by installing the application on their smartphone.

We make two contributions:

- (1) Drawing on the literature, we have developed a theoretical framework for understanding the drivers behind data donation behavior and we have empirically assessed its rationality through the data collected from the survey;
- (2) Our discussion highlights the implications of our framework for researchers who wish to recruit volunteers to donate their data to science.

## 2. BACKGROUND

### 2.1. Prior studies on donation behavior

There is a long history of investigating the driving factors of charitable donation (e.g. money) as well as blood donation. Previous work has identified a long list of factors that act as motivations to charitable giving. These include social responsibility, altruistic desire, overcoming guilt, an anticipation for some future return, accountability, trust, empathy, donation of others, self- and outcome efficacy, all of which have been found to be important motivators (Cheung and Chan, 2000; Martin and Randal, 2008, 2009; Prendergast, 2013). Demographic features are found to associate with people's propensity to donate. For instance, donations tend to increase with age until the age of 65 years, after which they decrease (Schlegelmilch *et al.*, 1997). Interestingly, a substantial body of research suggests that economists are less generous than other professionals and those economics students are less generous than other students (Bauman and Rose, 2011).

In a similar manner, a number of prior studies have examined the motivators and deterrents of blood donation. In a meta-analysis of existing literature, Bednall and Bove (2011) summarized that the major motivators of blood donation include convenience, prosocial motivation, personal values, perceived need for donation, indirect reciprocity, marketing communications, incentives and social norms. However, the major deterrents include low self-efficacy, low involvement, inconvenience, poor marketing communication, ineffective incentives, lack of knowledge, negative service experience, fear, negative attitudes and personal values. Despite rich literature on blood donation and charitable donation behavior, limited prior work has considered how to motivate people to donate personal data to science.

### 2.2. Citizen science projects

While the majority of citizen science projects are focused on 'collecting observational data of such phenomena as weather and precipitation, air and water quality, and species abundance and distribution' (Sheppard *et al.*, 2014), there are still a few projects in which more personal data is donated. Examples of such projects include the *American Gut Project* (aimed at sharing data about microbes in the participants' gut) (American Gut Project, 2014), the *Big Sleep Survey* (aimed at gathering information about participants' sleeping habits) (The Big Sleep Survey, 2014), the *Kinsey Reporter* (aimed at sharing data anonymously about sexual behaviors) (The Kinsey Reporter, 2014), and the *Personal Genome Project* (aimed at sharing data about genes) (Personal Genome Project, 2014). A number of motivators have been used to a great effect in some of these projects, including altruism (Goncalves *et al.*, 2013; Rotman *et al.*, 2012), need for

donation (Nov *et al.*, 2014), psychological empowerment (Goncalves *et al.*, 2014), self-improvement (American Gut Project, 2014) and prizes (The Big Sleep Survey, 2014).

Apart from the above-mentioned citizen science projects, we noticed a number of citizen science projects with the presence of business entity. For instance, Google Map Maker allows users to add specific buildings and services onto the map, thus donating their knowledge and data to the service. The *myPersonality Project* collected both personality and Facebook usage data, which has led to a number of interesting scientific reports (Youyou *et al.*, 2015). Therefore, it would be interesting to investigate how the organization reputation and donation behavior of others would affect an individual's data donation behavior.

The methodology used in our study was adopted from citizen science research. Our work involves investigating what drives the quantity and quality of online citizen science participation. We use structural equation modeling (Nov *et al.*, 2014) to investigate the effects of social signals and perceptions on action-taking. One of our goals is to highlight the potential of these parameters for future citizen science projects that require the systematic collection of contextual data.

### 2.3. The economics of revealing information

When it comes to donating contextual information, the thorny issue of *privacy* becomes prominent. It seems that *protecting* personal information substantially impedes people sharing their personal information even for scientific purposes. In this regard, prior studies have shown that there is dichotomy between self-professed privacy attitudes and actual self-revelatory behavior (Reynolds *et al.*, 2011), and that occasionally people are willing to sell their information (Acquisti, 2013).

Tedeschi (2002) reported on a Jupiter Research study that the overwhelming majority of online shoppers surveyed would give personal information to new shopping sites in exchange for a chance to win \$100 in a sweepstakes. The studies of Cvrcek *et al.* (2006) and Danezis *et al.* (2005) reported that people value their mobile location information in very different ways, and there are doubts as to whether people can, or do, value their privacy correctly and appropriately.

A number of experiments in different settings reported that even if most individuals stated that privacy was important to them, they were willing to trade privacy for convenience and discounts (Acquisti, 2013; Acquisti and Gross, 2006; Acquisti and Grossklags, 2005; Spiekermann *et al.*, 2001). A recent study of IPG Mediabrand and Microsoft Corp. (Consumers Find Value in Sharing Digital Information with Brands, 2013) shows that 59% of global consumers are much more willing to buy a product or service from a brand that offers a reward in exchange for their digital data. Grossklags and Acquisti

(2007) showed that people are more willing to sell rather than protect their personal information. In a general sense, prior studies have suggested a potential privacy paradox: 'people want privacy, but do not want to pay for it, and in fact are willing to disclose sensitive information for even small rewards' (Acquisti, 2013). This is partly confirmed by a recent work of Staiano *et al.* (2014) on the collection and sale of personally identifiable information, showing that people care about their privacy information but are also willing to sell contextual data. However, it remains unclear whether privacy matters in the case of *donating* data for scientific purposes.

### 3. RESEARCH FRAMEWORK

Synthesizing our findings from the previous literature on charitable behavior, blood donation and data privacy, we identified a number of factors that we proceed to validate in a research framework. Our framework includes eight different factors interacting in seven hypothesized ways. The factors include five latent (i.e. non-observable) variables:

- (1) *Privacy*: the extent to which a participant is concerned about privacy.
- (2) *Organization reputation*: the extent to which people appreciate the reputation of the organization that collects the donation (Bednall and Bove, 2011; Bennett and Ali-Choudhury, 2009).
- (3) *Need for donation*: the extent to which a participant believes that there is an urgent and necessary need for the donated data.
- (4) *Altruism*: the extent to which a participant has exhibited altruistic behavior in the past.
- (5) *Attitude*: the degree to which donation data to scientific research is valued.

And also three directly observed variables:

- (1) *Willingness to install*: the stated willingness of the participant to donate data by installing a smartphone application.
- (2) *Social signal*: whether the participant was exposed to a social signal by our system. These signals provide hints about the donation behavior of other donors.
- (3) *Actual donation behavior*: whether the participant actually donated data.

We now synthesize the ways in which the literature suggests these variables may interact. Previous work suggests that people care about privacy, and privacy concerns may prevent people from using particular technologies (Featherman and Pavlou, 2003; Featherman *et al.*, 2010; Liu *et al.*, 2013). Regarding donating contextual data to science, it is possible that people who worry about a loss of privacy will adopt a negative attitude toward data donation to science. Protecting

privacy is also a major concern of citizen science projects as well (Bowser *et al.*, 2014). Therefore, we formulate Hypothesis 1:

**Hypothesis 1: Privacy concern negatively relates to Attitude.**

In a summary of prior studies on driving factors of blood donation, Bednall and Bove (2011) noted that reputation of the collection agency plays an important role in motivating donation behavior. Bendapudi *et al.* (1996) indicated that a charity's image could comprise the single most critical element of its promotional program, because a charitable organization's image frequently determined whether the (donation) decision-making process would be initiated. A positive reputation of the collection agency encourages people to donate their blood (Bednall and Bove, 2011). High organization reputation has been reported to enhance customers' confidence and reduce risk perceptions, increasing customers' expectations of the organization's capability and integrity in providing an excellent service (Bednall and Bove, 2011; Keh and Xie, 2009). Since a data donor cannot see how the data will be finally utilized, organization reputation would be an important clue for people to build confidence that their data will be properly used. Furthermore, it is expected that when the donation request is from a reputable organization, people will be more likely to perceive the related donation activity to be urgent and necessary. Based on the above literature, we establish Hypotheses 2a, 2b and 2c:

**Hypothesis 2a: Organization reputation negatively relates to Privacy concern.**

**Hypothesis 2b: Organization reputation positively relates to Need for donation.**

**Hypothesis 2c: Organization reputation positively relates to Attitude.**

A key component of charitable behavior and participation in citizen science projects is altruism (Rotman *et al.*, 2012). Altruism refers to people's altruistic personality (Rushton *et al.*, 1981), measuring the extent of an individual's willingness to act in the interests of others without the expectation of reward or positive reinforcement in return (Karra *et al.*, 2006). Altruism can explain why people donate money to unrelated individuals and organizations (Andreoni, 1990; Prendergast, 2013). Prior studies have interpreted altruism in the sense of emotions and empathy (Prendergast, 2013; Ray, 1998), humanitarianism (Cermak *et al.*, 1994) and simply a desire to help others (Harvey, 1990). Therefore, it is possible that altruistic individuals will be more likely to develop a positive attitude to donation action and will be easier to convince of the need to donate. In this regard, we introduce Hypotheses 3a and 3b:

**Hypothesis 3a: Altruism positively relates to Need for donation.**

**Hypothesis 3b: Altruism positively relates to Attitude.**

Need for donation reflects one's awareness that donation is necessary for scientific research (Cheung and Chan, 2000). The need for donation is a stimulus for triggering donation behavior according to social cognitive theory (Cheung and Chan, 2000) and a significant determinant in donating money (Cheung and Chan, 2000; Diamond and Kashyap, 1997) and blood (Bednall and Bove, 2011). In citizen science, a similar construct of collective motivations is utilized to interpret project participation intention (Nov *et al.*, 2014). Therefore, we introduce Hypothesize 4.

**Hypothesis 4: Need for donation positively relates to Attitude.**

Attitude towards an activity, a concept introduced by the theory of reasoned action, affects people's intention or willingness to adopt the activity and in turn brings about the actual behavior. Developed by Ajzen and Fishbein (1980) and Fishbein and Ajzen (1975), the theory of reasoned action has been one of the most applied social science theories. Consistent with the theory, we elicit Hypotheses 5 and 6:

**Hypothesis 5: Attitude positively relates to Willingness to install.**

**Hypothesis 6: Willingness of installation positively relates to Actual installation.**

Prior studies have repeatedly suggested that social signals or electronic word-of-mouth has a strong effect on people's behavior. Specifically, they suggest that social norms can have a strong impact on casual decisions to donate (Bennett and Ali-Choudhury, 2009; Gounaris and Stathakopoulos, 2004; Mitra and Gilbert, 2014). People may be drawn into making a donation as a result of the social influence of peers (Bennett and Ali-Choudhury, 2009). A multitude of social commerce and the electronic word-of-mouth literature suggest that people are more likely to adopt the products or the activities that more of the others adopted (Bond *et al.*, 2012). Martin and Randal (2008) conducted a field study investigating voluntary contributions to a public good and manipulated the social information available to patrons by altering what was visible in the donation box. Their experiment shows that the provided social information had a significant impact on donation composition, frequency and value. Similarly, in the context of data donation, we establish Hypothesis 7:

**Hypothesis 7: Social signal positively relates to Actual installation.**

## 4. RESEARCH METHODOLOGY

Our methodology consisted of a typical online survey with an experimental manipulation at the end to measure social signals' impact on contextual data donation. Once participants completed the survey and were redirected to the 'Thank You' page, which contained social signals. Participants were then given the opportunity to download a smartphone application that would enable them to donate context data to science. At this point, we manipulated the social signals shown to participants. We could track how many people actually downloaded and installed the application successfully and subsequently donated context data to us. Furthermore, every participant who installed the application could be linked back to the survey and their responses. Thus, our methodology combines *subjective attitude* data (questionnaire) with *explicit attitude* data (installing the application and donating data).

We offered a lucky draw to all respondents of the survey, providing an Apple iPad Mini or a new Samsung Galaxy Tab 2 7.0 as the prize. We distributed the questionnaires by posting the link of the survey to the university email list, university's Facebook page, as well as publicly available mailing lists.

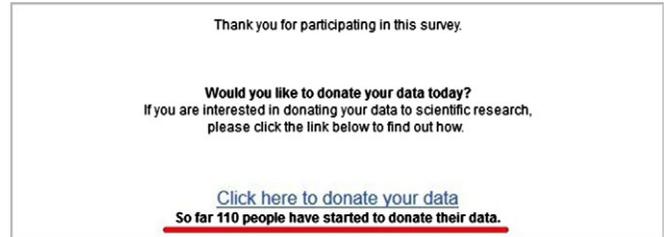
### 4.1. Questionnaire and samples

The online questionnaire consisted of two sections. The first measured participants' demographic features. The second section collected data on participants' perception regarding donation behavior and their willingness to donate contextual data. In our research framework, privacy, organization reputation, need for donation, altruism and attitude are latent variables measured by questionnaire items.

A single five-point Likert-scale ranging from disagree (1) to agree (5) was used to measure every reflective latent variable included in the research framework. We reworded questionnaire measurements from prior studies to fit the purpose of our experiment. Specifically, the items for measuring altruism, organization reputation, privacy and attitude are adjusted based on the measurement developed by Rushton *et al.* (1981), Cheung and Chan (2000), Featherman and Pavlou (2003) and Kim *et al.* (2009), respectively. The measurement of need for donation is based on the work of Bednall and Bove (2011) and Cheung and Chan (2000).

At the beginning of the second section, we gave a description of the data donation software, in which information with regard to organization reputation (University of XXXX) and privacy statement (will not log personal information) are provided to facilitate a better understanding of our research context:

A mobile application that helps scientists of University of XXXX to non-intrusively collect phone usage information from you, and the application will not log personal information, such as phone numbers, contacts information and any identifier that could be



**Figure 1.** The 'Thank You' page including a link to download an application for donating data and a social signal (underlined in red for annotation purposes).

breaching your privacy. For instance, the application may record that a SMS is sent at a specific time, but not the content of SMS itself. You can choose which sorts of data to donate, as we will specify later.

Willingness to install is measured by asking respondents to indicate for how long they would be willing to donate contextual data, ranging from no willingness to donate (1), 1 week (2), 1 month (3), 6 months (4), 1 year (5) and as long as the experiment requires (6). We stress that at this point in the survey, participants did not know that they would be given an opportunity to actually do so at the end.

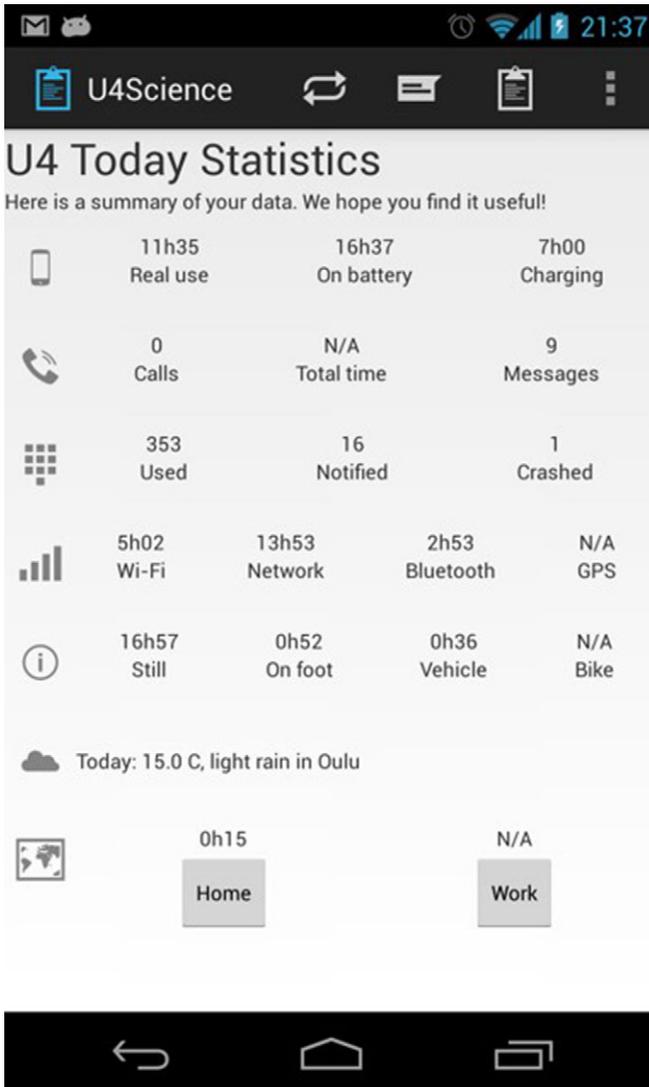
At the end of the survey, we offered participants an opportunity to actually donate contextual data to science by installing our software (i.e. a mobile application). Specifically, after respondents clicked to submit the questionnaire, a page appeared to express our thanks for their participation and in addition provide a link to the software so that participants can actually start to donate. All participants were informed that the survey was complete at this point, and so they had no obligation to continue with the installation. We manipulated the social signal shown on this page, which was included in the research framework as a priming factor. We manipulated the interface by adding a social signal message, as shown in Fig. 1 (underlined in red). Participants were randomly assigned to one of the three social signal conditions:

- No social signal is shown.
- So far 110 people have started to donate their data.
- So far 510 people have started to donate their data.

### 4.2. Software for donating data

The respondents who followed the 'Click here to donate your data' link were directed to another page with step-by-step instructions on how to donate data using our smartphone application (Fig. A1). We built our application (U4Science) using AWARE's (Ferreira, 2013) mobile context framework.

The instruction page gave the same information to all visitors, including screenshots of the application. A short URL was provided to the user to install the software. Users were instructed to visit this URL using their smartphone's browser.



**Figure 2.** Example statistics of the data donated by a user.

This minimized the amount of typing on the mobile phone. U4Science required a custom PIN to unlock it, thus avoiding accidental data donations and allowing us to link software installations to survey responses. Our most important objective for the mobile software design was to make sure that the donors would be in control of their donation at all times. The interface allowed donors to send feedback to us (Fig. A2, left), stop donating data (Fig. A2, center) and choose what contextual data to upload and when to share it with us (Fig. A2, right).

### 4.3. Donated data

The donors were given granular controls to choose which contextual data to send. U4Science collected a combination of the following data (Fig. 2):

**Table 1.** Demographics of participants and donors.

Demographic	Categories	Participants	Donors
Gender	Male	247 (74%)	12
	Female	87 (26%)	1
	Total	334	13
Age	<18	1 (0.3%)	0
	18–25	114 (34.1%)	5
	26–35	151 (45.2%)	7
	36–45	51 (15.3%)	1
	46–55	12 (3.6%)	0
	>55	5 (1.5%)	0
	Total	334	13
Education	High school or lower	73 (21.9%)	3
	Bachelor degree	94 (28.1%)	3
	Master degree	125 (37.4%)	6
	Doctor degree or higher	42 (12.6%)	1
	Total	334	13
Income (€)	<2000	178 (53.3%)	8
	2000–2999	78 (23.4%)	3
	3000–3999	43 (12.9)	1
	4000–4999	11 (3.3%)	0
	>5000	24 (7.2%)	1
	Total	334	13
Field of study	IT-related	169 (50.6%)	8
	Economics-related	20 (6%)	0
	Science	39 (11.7%)	2
	Engineering	84 (25.1%)	3
	Arts	22 (6.6%)	0
	Total	334	13

- Phone: battery and display usage;
- Communications: calls and message statistics (metadata only);
- Applications: application usage statistics;
- Activity: activity recognition (e.g. idle, walking, biking, in vehicle)
- Location: time spent in specific locations (e.g. home and work), with low location accuracy (i.e. using only network triangulation);
- Weather: weather conditions at the user's location.

## 5. Findings

The online survey was deployed for 4 months with all the responses collected between 8 July and 4 November 2013. In total, we collected 360 completed questionnaires. The demographics of participants and donors are shown in Table 1.

However, 26 participants claimed not to use a smartphone or being unable to install the application. Their responses

were therefore excluded from the analysis. Participants were from 21 different countries, based on the IP address when filling in the survey, with the majority (79%) from Finland, while the rest are primarily from western countries including USA, UK, Portugal, Switzerland, Canada and Italy. We provide a breakdown of how many participants progressed to each stage of our study:

- Survey responses: 360 people;
- Responses retained for further analysis: 334 people;
- Responses with intention to donate data: 287 people;
- Clicked the donation link at the end: 175 people;
- Installed the mobile application: 13 people.

### 5.1. Who donated data

In the survey, 287 participants expressed a willingness to donate data to various extents. Despite such a large number of ‘willing’ donors, only 13 people actually donated their data, implying a substantial gap between intention and behavior. Of these 13 people, most were from Oulu ( $N = 6$ ) or Helsinki ( $N = 3$ ) and one from each of Canada (Toronto), Portugal (Porto), Mexico (Obregon) and Sweden (Umeå). This suggests that 3.4% of respondents from Finland installed the application, compared to 4.2% of respondents outside Finland. All donors have responded on the survey with the intention to donate data.

Of the 13 participants who actually installed our application, 12 are men, between 18 and 35 years old and with an income of <3000 Euros per month and a master’s degree or lower. Regarding their majors of education, eight donors studied IT-related subjects, three studied engineering and two studied science. All donors are experienced mobile phone users and have installed several applications in their phones already.

Therefore, the demographic features of donors suggest that young (18–35 years) male IT or engineering students with rich mobile application usage experience are more likely to donate contextual data. Regarding the effect of our social signal manipulation, only one participant in the group with no social signal actually installed the application, as shown in Table 2. A chi-square test shows that the participants who saw a social signal were significantly more likely to actually install the application ( $\chi^2 = 3.977$ ,  $df = 1$ ,  $P < 0.05$ ). No significant difference was found between the number of respondents who clicked the link at non-social signal versus social signal groups ( $\chi^2 = 1.268$ ,  $df = 1$ ,  $P = 0.296$ ).

### 5.2. Donated data

Regarding the actual data that were donated, most donors provided battery data ( $N = 8$ ), screen usage data ( $N = 7$ ), calls statistics ( $N = 5$ ), messaging statistics ( $N = 5$ ) and data on running application ( $N = 4$ ). Only two participants donated activity recognition data, while only one participant is found

**Table 2.** Summary of donators and exposed social signal.

Social signal	Respondents	Clicked the link	Donors
No signal	111	63	1
110 ... donate	108	54	6
510 ... donate	115	58	6

**Table 3.** Reliability and convergent validity statistics.

Construct (number of items)	$\alpha$	CR	Minimum FL
Organization reputation (3)	0.81	0.81	0.639
Privacy (2)	0.90	0.90	0.852
Need for donation (3)	0.85	0.85	0.739
Altruism (3)	0.62	0.63	0.538
Attitude (3)	0.83	0.82	0.635

**Table 4.** Discriminant validity.

Construct	OR	PA	ND	AL	AT
OR	<b>0.77</b>				
PA	-0.10	<b>0.90</b>			
ND	0.17	-0.23	<b>0.81</b>		
AL	0.02	0.05	0.27	<b>0.61</b>	
AT	0.23	-0.21	0.69	0.25	<b>0.78</b>

The bold diagonals are the square roots of the average variance extracted of the individual constructs. Off-diagonal values are the correlations between constructs.

OR, organization reputation; PA, privacy; ND, Need for donation; AL, Altruism; AT, Attitude

to donate location or weather information. This variation in the donated data is enabled by the granular control mechanisms we provided to the donors (Fig. A2).

Our results also indicate a large discrepancy regarding the length of data donation. Within 2 days of installing our application, six participants stopped donating data. However, one donor gave data for 123 days, another for 32 days, while three donors stopped their donation after 20 days. The difference in donation activities was enabled by our easy opt-out mechanism and suggests that donors were highly conscious of their data donation activities.

### 5.3. Questionnaire measurement validity and reliability

Confirmatory factor analysis was utilized to test the adequacy of the measurement model using *Amos 19*. One item for measuring privacy was found to have a low loading value. After removing this item, we repeated the confirmatory factor analysis, the results of which demonstrated a satisfactory fit (Table 3). The values of Cronbach’s alpha ( $\alpha$ ) and composite reliability (CR) and minimal factor loading (FL) of the constructs are all over the thresholds of 0.7, 0.6 and 0.5, respectively (Tseng, 2006), as shown in Table 3. The squared roots of average variance extracted are higher than their correlations

**Table 5.** Fit Indices for Measurement and Structural Models.

Fit statistics	Measurement model	Structural model	Recommended value
$\chi^2/df$	1.943	2.035	<3
GFI	0.940	0.929	>0.90
AGFI	0.903	0.900	>0.80
RMSEA	0.053	0.056	<0.08
CFI	0.956	0.944	>0.90

with other constructs, as shown in Table 4. The results show that all items fit their respective factors quite well. The results suggest unidimensionality, convergent and discriminant validity of the measures. Harmon's one-factor test is applied to test common method bias. No factor is found to account for the majority of the covariance in the variables, which suggests that common method bias is an unlikely concern in the data.

Multiple common model-fit indices were used to estimate the measurement model and structural model fit (Hair *et al.*, 2006): Chi-square/degrees of freedom; the Goodness-of-Fit Index (GFI); the Adjusted Goodness-of-Fit Index (AGFI); Root Mean Square Error of Approximation (RMSEA); and the Comparative Fit Index (CFI) as seen in Table 5. All the indices indicate a good model fit.

#### 5.4. Model evaluation and hypotheses testing

Structural model testing indicated a good fit between the model and data. Except for Hypothesis H3b, all hypotheses are supported.

The perceived organization reputation has significant effects on both privacy ( $\beta = -0.109$ ,  $P < 0.05$ ) and need for donation ( $\beta = 0.175$ ,  $P < 0.01$ ). Altruism has significant influence on need for donation ( $\beta = 0.267$ ,  $P < 0.001$ ) but has no significant effect on attitude. Privacy negatively relates to attitude ( $\beta = -0.098$ ,  $P < 0.05$ ). In addition, organization reputation ( $\beta = 0.114$ ,  $P < 0.05$ ) and need for donation ( $\beta = 0.642$ ,  $P < 0.001$ ) are direct determinants of attitude. Consistent with our expectation, attitude ( $\beta = 0.358$ ,  $P < 0.001$ ) significantly relates to willingness to install. Willingness to install ( $\beta = 0.177$ ,  $P < 0.001$ ), together with social signal ( $\beta = 0.1$ ,  $P < 0.05$ ), significantly influences the likelihood of actual installation.

Table 6 summarizes the results of the hypotheses tests. Overall, the model interpreted 1.2% of variance of privacy, 10.4% of need for donation, 50.9% of attitude, 12.8% of willingness of installation and 4.1% of actual installation, as shown in Fig. 3. The explained variance of actual installation is quite low, which is partly caused by the use of binary measurement.

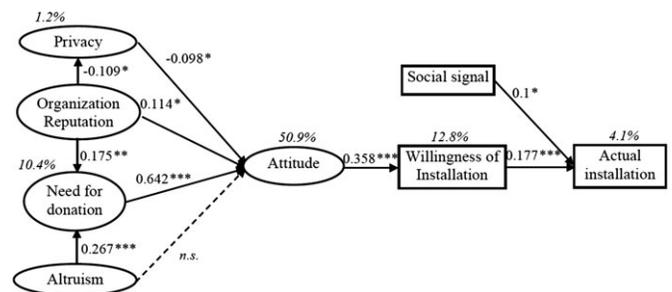
## 6. DISCUSSION

Despite the increasing awareness on the importance of contextual data collection in HCI, actionable strategies remain

**Table 6.** Results of hypotheses test.

Hypotheses	Coefficients	Finding
H1: Privacy $\rightarrow$ Attitude	-0.098*	Supported
H2a: OR $\rightarrow$ Privacy	-0.109*	Supported
H2b: OR $\rightarrow$ Need for donation	0.175**	Supported
H2c: OR $\rightarrow$ Attitude	0.114*	Supported
H3a: Altruism $\rightarrow$ Need for donation	0.267***	Supported
H3b: Altruism $\rightarrow$ Attitude	NS	Not supported
H4: Need for donation $\rightarrow$ Attitude	0.642***	Supported
H5: Attitude $\rightarrow$ Willingness	0.358***	Supported
H6: Willingness $\rightarrow$ Installation	0.177***	Supported
H7: Social signal $\rightarrow$ Installation	0.1*	Supported

\*  $< 0.05$ , \*\*  $< 0.01$ , \*\*\*  $< 0.001$ .

**Figure 3.** Result of testing the research model.

scarce (Reddy *et al.*, 2010). Our study aims to investigate the underlying issues relating to the donation of contextual data for scientific research. Particularly, we sought to model the decision process of individuals when donating their contextual data for scientific research.

#### 6.1. Why do people donate contextual data?

Overall, we found that people are *hesitant* to donate their contextual data. More crucially, our results suggest that *altruism* is not a directly significant determinant, unlike in the broader scope of charitable donation. Furthermore, we found a significant *disparity* between respondents' *claims* and actual *actions* towards *donating contextual data* to science. While 287 participants expressed some degree of willingness to donate, only 13 of them actually donated when it came to fulfilling their promise. However, this finding does not necessarily mean that respondents lied, particularly if we consider the positive and significant path coefficient between willingness of installation and actual installation. Rather, it suggests that people might require additional stimuli or a catalyst to 'take the plunge' because their attitudes are not likely to be enough to turn into action.

An example of such a catalyst, which we controlled in this study, is social signals. We found that only 1 from 111 participants (0.9%) in the no social signal group actually became a donor, in comparison to 12 donors from 223 participants (5.38%) in the social signal groups. Participants who saw the social signal were about five times more likely to become donors. This finding has clear implications for approaching potential data donors: *use social signals during recruitment*.

While social signals are crucial in ultimately turning willingness into action, our study shows that willingness is nevertheless an important prerequisite. Therefore, it is important to understand how people formulate a willingness to donate data in the first place. Having a *positive perception of the organization's reputation* is an important driver of people's willingness to donate. It helps reduce the anxiety regarding privacy, it increases the perception of the need for donation and it has a direct influence on attitude. This finding has two important implications. First, *reputable organizations have an important advantage in attracting participants for their research*. Therefore, in the long run, raising the reputation of one's organization (e.g. through university promotion campaign) will make it easier to approach potential data donors in the future. Second, individuals who respect the organization are good candidates to become data donors, i.e. university alumni.

The perceived need for donation is also an influential determinant. It mediates the effects of altruism on attitude. In other words, our results show that altruistic people are not necessarily willing to help scientific research unless they realize a clear need for donation. A possible interpretation for this is that because the output of HCI research (e.g. as a result of data donation) is probably not as immediate, visible or critical as organ, blood and charitable donation, altruistic individuals may hesitate to form a positive attitude toward data donation. Thus, altruism may only take an effect when the need for donation is perceived. Therefore, it is important to clearly elucidate to potential data donors the research purpose as well as the significance of conducting research. Accordingly, producing a well-motivated 'advertisement' may contribute to an effective strategy to recruit donors, particularly people who have altruistic personality (e.g. blood donation or charitable giving experience) are good candidates to respond to the advertisement.

Furthermore, we found that *privacy concerns have a significant and negative influence on attitude, albeit relatively marginal* ( $\beta = -0.098$ ,  $P < 0.05$ ). Therefore, a clear statement on the privacy protection of using personal data can ease the concerns of potential donors. However, considering its marginal effects, the privacy statement will not substantially increase people's willingness to donate data. Interestingly, our study was conducted during the immediate aftermath of Edward Snowden's disclosure of the NSA's PRISM surveillance program. Based on our findings, it seems that privacy concerns do not cause a strong reaction to effectively prevent people from donating contextual data.

It is worth noting that all donors are experienced mobile phone users with multiple applications already installed in their phones before the experiment. This fact may partly explain the gap between the large number of participants who expressed a willingness to donate data and the small amount of actual donors. According to a study of [Kaiser and Schultz \(2009\)](#), the attitude-behavior relationship deteriorates in the context of very difficult behaviors. Therefore, when respondents lack of confidence in installing and configuring the donation software, they may not turn their willingness to actual behavior. In this light, the presence of social signal may deliver respondents a hint that configuring the donation software is not difficult, since many other people have already donate their data successfully, thus motivating respondents to take the action.

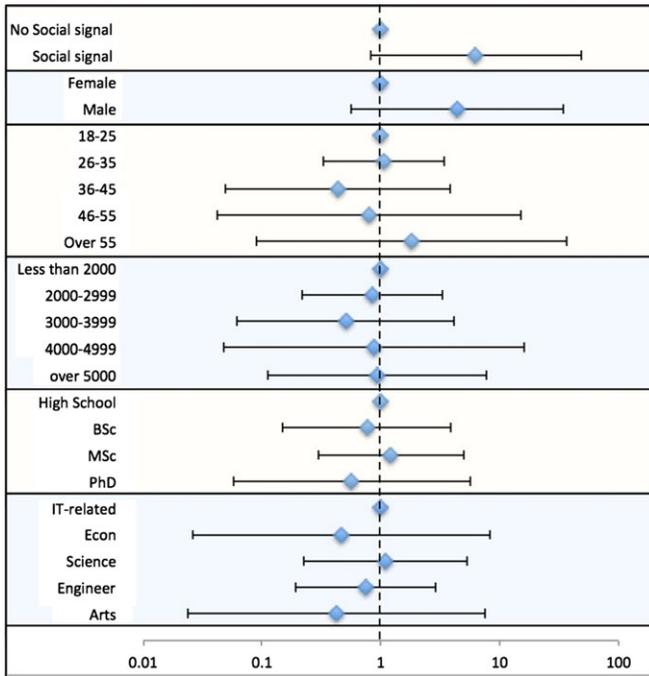
## 6.2. Donating contextual data is not the same as other types of donation

A key insight of our findings is that *scientific data donation differs in some important ways from other charitable donations*. Most prior studies on charitable donation have suggested that in general, people's propensity to donate increases with age and education, and that females donate more often than males ([Schlegelmilch et al., 1997](#); [Simmons and Emanuele, 2007](#); [Srnska et al., 2003](#)). However, our findings suggest that this is not the case for scientific data donation. Specifically, the results suggest that *as age increases, contextual data donors decrease*. In addition, *as income or education increases, contextual data donors decrease*. Furthermore, males are substantially more likely to donate data than females.

If we draw an analogy with money, previous research suggests that people of higher education and age are more likely to donate money, and we argue that these are demographic groups that actually have more money. What we found in our study is that *younger and (and by implication perhaps slightly less educated) individuals are more likely to donate data*. We argue that in fact these are the demographics of people that have more data and is the generation known as the *Millennials* or *Generation Y*.

This younger demographic may be more active in online social sharing, keeping in touch electronically, generating lots of media content and trying out new applications and services. In this sense, therefore, we argue that contextual data donors should be sought within data-rich communities (sophisticated phone users). In [Fig. 4](#), we visualize the 'odds ratio' of various demographic groups. These indicate how much more likely (or less likely) they are to donate data given the baseline group (categories without whiskers).

[Figure 4](#) is a standard forest plot that visualizes the odds ratio at 95% confidence intervals for variables. These values are obtained from a pairwise binary logistic regression between the outcome variable (whether individuals donated



**Figure 4.** Odd ratios of comparing different demographic and social signal features. Odds ratios (at the 95% confidence interval) indicate how much more/less likely is a category to donate compared to a baseline category (shown without whiskers).

data or not) and each of the variable in the diagram. Note that the X-axis scale is logarithmic. We have to arbitrarily set the baseline for each variable to be one of the levels of the variable and estimate the odds ratio for the remaining levels of the variable according to their 95% confidence interval. These values are calculated automatically by our statistics software (SPSS) in order to visualize the relative effect, other than using  $P$  values. Because these intervals are obtained by considering each variable independently, this analysis is not as robust as the SEM analyses, which considers all values simultaneously. Therefore, we include this figure to visualize the magnitude of the effect, not as a statistical test.

## 7. RECOMMENDATIONS

Based on our results, we list actionable suggestions for researchers who wish to approach contextual data donors.

- Build on reputation. Highlight the organization (e.g. using its logo) if it is well respected with a strong and positive reputation. If it does not have a strong reputation, then consider approaching individuals with ties to the organization (e.g. alumni of a university) since they are more likely to perceive it positively.
- Highlight the need for data donation. In your communication and messages, clearly motivate the need for the research and the potential benefits of conducting the

research. This is key in convincing altruistic people to donate data.

- Identify key donors. These are more likely to be younger (18–35 years old) men with master degree or lower in Science or IT-related fields. It can be challenging to recruit females, seniors and people with other majors. While such participants may also be data rich, our results show they are more hesitant to donate their phone usage data under the presence of social signal and may require additional incentives to contribute. Hence, researchers who want to recruit individuals with diverse demographic backgrounds may have to place additional effort to recruit donors who fall within these demographics.
- Give strong social signals. Inform potential donors that they are not alone and that others have decided to donate. For example, highlight how many people have donated or possibly how much data have been donated.

### 7.1. Limitations

The scope of our study is limited to donating contextual data from mobile phones. This context differs from previous studies that either involve direct rewards for participants or consider different technologies such as online social networks or desktop computers. Therefore, there exists a diversity of donation contexts, which may somewhat affect participants' donation intention and behavior. In addition, the sample size of our study is relatively small, in particular, considering the number of people who actually installed the mobile software, which may affect the generalizability of our results. For instance, the results suggest that social signal makes it five times more effective to attract actual data donation. However, this reported strength of the effect may partly be affected by our relatively small sample size ( $N = 334$ ). More precise value can be achieved via the use of a bigger sample.

Also most respondents major in IT-related subjects or engineering, and therefore they are over-represented in our sample. This may be partly caused by the possibility that respondents with IT-related subjects or engineering background are more interested in participating in an IT-related project aimed at understanding data donation using mobile phones. Therefore, we assume that data donation projects may tend to attract respondents with a background related to IT-related subjects or engineering. Nonetheless, our moderate sample size may limit the generalizability of our research findings and audience should be aware of this in applying our findings.

Despite our efforts to minimize the effort to donate data using smartphones, our method (questionnaire–instructions–activation) might have impacted the donation adoption from our respondents. An alternative approach could have been to use an application store as the recruiting mechanism. However, this would entail additional research challenges (Ferreira *et al.*, 2012) that go beyond the scope of this work.

Furthermore, the questionnaire described the need of donation in a generic manner, which is to support scientific research. In this light, including a more specific purpose of data usage, such as developing better applications, may enable more concrete perceptions on the need of donation among participants.

Prior studies show that people are willing to disclose sensitive information for even small rewards (Staiano *et al.*, 2014). In our study, we did not provide any incentives to donors. Therefore, it is unclear to what degree monetary rewards would help raise the actual donation rate, resulting in possible avenues for future research.

## 8. CONCLUSION

This article investigates why it is challenging to recruit donors of contextual data. Drawing on previous work on charitable behavior, citizen science and donation, we developed a research framework that we hypothesize is underlying the decision-making process of potential donors. In addition to deploying a questionnaire to validate our model, we also provided participants an opportunity to actually take action and donate contextual data using a smartphone application we developed.

Our findings highlight that the key demographics for donating data are not necessarily the same as those for donating money. We summarize our findings as a set of four recommendations for recruiting data donors: build on reputation, highlight the need for data donation, identify data-rich donors and provide strong social signals.

Our work, findings and recommendations focus on contextual data donation, and therefore do not consider the use of monetary or other extrinsic rewards in this process. Clearly, there is potential for individuals to sell their data, but our study indicates that about 5% of respondents are likely to donate data given the presence of strong social signals, and this ratio is likely to increase when following our recommendations.

## FUNDING

Academy of Finland (276786-AWARE, 285062-iCYCLE, 286386-CPDSS, 285459-iSCIENCE); European Commission (PCIG11-GA-2012-322138 and 645706-GRAGE).

## REFERENCES

- Acquisti, A. (2013). The economics of privacy: Theoretical and empirical aspects. The economics of privacy: Theoretical and empirical aspects [Web page]. Retrieved from Google Scholar: <http://cusp.nyu.edu/wp-content/uploads/2013/09/C03-acquisti-chapter.pdf>.
- Acquisti, A. and Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy enhancing technologies*, pp. 36–58. Springer, Berlin Heidelberg. doi:10.1007/11957454\_3.
- Acquisti, A. and Grossklags, J. (2005). Privacy and rationality in individual decision making. *Security & Privacy*, 3, 26–33. doi:10.1109/MSP.2005.22.
- Ajzen, I. and Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- American Gut Project. (2014). American Gut project. [Web page]. Retrieved from <http://humanfoodproject.com/americangut/>.
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *Econ. J.*, 100, 464–477. Retrieved from <http://www.jstor.org/stable/10.2307/2234133>.
- Bauman, Y. and Rose, E. (2011). Selection or indoctrination: Why do economics students donate less than the rest? *J. Econ. Behav. Organ.*, 79, 318–327. doi:10.1016/j.jebo.2011.02.010.
- Bednall, T.C. and Bove, L.L. (2011). Donating blood: A meta-analytic review of self-reported motivators and deterrents. *Transfus. Med. Rev.*, 25, 317–34. doi:10.1016/j.tmr.2011.04.005.
- Bendapudi, N., Singh, S.N. and Bendapudi, V. (1996). Enhancing helping behavior: An integrative framework for promotion planning. *J. Mark.*, 60, 33–54. Retrieved from <http://www.jstor.org/stable/10.2307/1251840>.
- Bennett, R. and Ali-Choudhury, R. (2009). Second-gift behaviour of first-time donors to charity: An empirical study. *IJNVSM*, 180, 161–180. doi:10.1002/nvsm.
- The Big Sleep Survey. (2014). The big sleep survey. [Web page]. Retrieved from <http://www.sleepsurvey.net.au/index.php>.
- Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D.I., Marlow, C., Settle, J.E. and Fowler, J.H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489, 295–8. doi:10.1038/nature11421.
- Bowser, A., Wiggins, A., Shanley, L., Preece, J. and Henderson, S. (2014). Sharing data while protecting privacy in citizen science. *Interactions*, 21, 70–73. Retrieved from Google Scholar.
- Cermak, D.S.P., File, K.M. and Prince, R.A. (1994). A benefit segmentation of the major donor market. *J. Bus. Res.*, 29, 121–130. Retrieved from <http://www.sciencedirect.com/science/article/pii/0148296394900167>.
- Cheung, C. and Chan, C. (2000). Social-cognitive factors of donating money to charity, with special attention to an international relief organization. *Eval. Program. Plann.*, 23, 241–253. doi:10.1016/S0149-7189(00)00003-3.
- Consumers Find Value in Sharing Digital Information with Brands*. (2013). IPG Mediabrands and Microsoft Corp. (Original work published 2013) Retrieved from <http://news.ipgmediabrands.com/ipgmediabrands/press-releases/microsoft-digital-data-mediabrands.htm>.
- Cvrcek, D., Kumpost, M., Matyás, V. and Danezis, G. (2006). The value of location information: A European-wide study. In *Security protocols-14th international workshop*, Vol. 5087, pp. 112–121. LNCS, (Original work published 2006) Retrieved from Google Scholar.
- Danezis, G., Lewis, S. and Anderson, R.J. (2005). How much is location privacy worth? *WEIS*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.236&rep=rep1&type=pdf>.

- Diamond, W.D. and Kashyap, R.K. (1997). Extending models of prosocial behavior to explain university alumni contributions. *J. Appl. Soc. Psychol.*, 27, 915–928. doi:10.1111/j.1559-1816.1997.tb00278.x.
- Featherman, M.S., Miyazaki, A.D. and Sprott, D.E. (2010). Reducing online privacy risk to facilitate e-service adoption: The influence of perceived ease of use and corporate credibility. *J. Serv. Market.*, 24, 219–229. doi:10.1108/08876041011040622.
- Featherman, M.S. and Pavlou, P.A. (2003). Predicting e-services adoption: A perceived risk facets perspective. *Int. J. Hum. Comput. Stud.*, 59, 451–474. doi:10.1016/S1071-5819(03)00111-3.
- Ferreira, D. (2013). *Aware: A mobile context instrumentation middleware to collaboratively understand human behavior*. PhD Thesis. Retrieved from <http://herkules.oulu.fi/isbn9789526201900/isbn9789526201900.pdf>.
- Ferreira, D., Kostakos, V. and Dey, A.K. (2012). Lessons learned from large-scale user studies: Using android market as a source of data. *IJMHCI*, 4, 28–43. doi:10.4018/jmhci.2012070102.
- Fishbein, M. and Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Goncalves, J., Ferreira, D., Hosio, S., Liu, Y., Rogstadius, J., Kukka, H. and Kostakos, V. (2013). Crowdsourcing on the spot: Altruistic use of public displays, feasibility, performance, and behaviours. *UbiComp'13: Proc. 2013 ACM. Int. Joint Conf. Pervasive and Ubiquitous Computing*, pp. 753–762. ACM. doi:10.1145/2493432.2493481.
- Goncalves, J., Kostakos, V., Karapanos, E., Barreto, M., Camacho, T., Tomasic, A. and Zimmerman, J. (2014). Citizen motivation on the go: The role of psychological empowerment. *Interact. Comput.*, 26, 196–207. doi:10.1093/iwc/iwt035.
- Gounaris, S. and Stathakopoulos, V. (2004). Antecedents and consequences of brand loyalty: An empirical study. *J. Brand Manage.*, 11, 283–306.
- Grossklags, J. and Acquisti, A. (2007). When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *WEIS*. Retrieved from Google Scholar.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2006). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall.
- Harvey, J.W. (1990). Benefit segmentation for fund raisers. *J. Acad. Market. Sci.*, 18, 77–86.
- Hilbert, D.M. and Redmiles, D.F. (1998). An approach to large-scale collection of application usage data over the internet. In *Proc. 20th Int. Conf. Software Engineering*, pp. 136–145. Retrieved from Google Scholar.
- Hornbæk, K., Sander, S.S., Bargas-Avila, J.A. and Grue Simonsen, J. (2014). Is once enough?: On the extent and content of replications in human-computer interaction. In *CHI '14: Proc. SIGCHI Conf. Human Factors in Computing Systems*, pp. 3523–3532. ACM. doi:10.1145/2556288.2557004.
- Kaiser, F.G. and Schultz, P.W. (2009). The Attitude–behavior relationship: A test of three models of the moderating role of behavioral difficulty. *J. Appl. Soc. Psychol.*, 39, 186–207.
- Karra, N., Tracey, P. and Phillips, N. (2006). Altruism and agency in the family firm: Exploring the role of family, kinship, and ethnicity. *Entrep. Theory. Pract.*, 30, 861–877. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6520.2006.00157.x/full>.
- Keh, H.T. and Xie, Y. (2009). Corporate reputation and customer behavioral intentions: The roles of trust, identification and commitment. *Ind. Market. Manage.*, 38, 732–742. doi:10.1016/j.indmarman.2008.02.005.
- Kim, Y.J., Chun, J.U. and Song, J. (2009). Investigating the role of attitude in technology acceptance from an attitude strength perspective. *Int. J. Inf. Manage.*, 29, 67–77.
- The Kinsey Reporter. (2014). *The Kinsey Reporter*. [Web page]. Retrieved from <http://kinseyreporter.org/>.
- Liu, Y., Kostakos, V. and Deng, S. (2013). Risks of using NFC mobile payment: Investigating the moderating effect of demographic attributes. *Int. Conf. Electronic Commerce*, pp. 125–134. Turku, Finland. Retrieved from Google Scholar: <http://www.ee.oulu.fi/~vassilis/files/papers/icec13.pdf>.
- Martin, R. and Randal, J. (2008). How is donation behaviour affected by the donations of others? *J. Econ. Behav. Organ.*, 67, 228–238. doi:10.1016/j.jebo.2007.08.001.
- Martin, R. and Randal, J. (2009). How sunday, price, and social norms influence donation behaviour. *J. Socio. Econ.*, 38, 722–727. doi:10.1016/j.socec.2009.03.010.
- Mitra, T. and Gilbert, E. (2014). The language that gets people to give: Phrases that predict success on kickstarter. *Proc. 2014 ACM. Conf. Computer Supported Cooperative Work*. Retrieved from Google Scholar.
- Nov, O., Arazy, O. and Anderson, D. (2014). Scientists@ home: What drives the quantity and quality of online citizen science participation? *PLoS. ONE.*, 9, e90375. Retrieved from Google Scholar.
- Personal Genome Project. (2014). *Personal genome project*. [Web page]. Retrieved from <http://www.personalgenomes.org/>.
- Prendergast, G.P. (2013). Donors' experience of sustained charitable giving: A phenomenological study. *J. Consum. Mark.*, 30, 130–139. doi:10.1108/07363761311304942.
- Ray, L. (1998). Why we give: Testing economic and social psychological accounts of altruism. *Polity*, 30, 383–415.
- Reddy, S., Estrin, D. and Srivastava, M. (2010). Recruitment framework for participatory sensing data collections. In *Pervasive'10: Proc. 8th Int. Conf. Pervasive Computing*, pp. 138–155. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-12654-3\_9.
- Reynolds, B., Venkatanathan, J., Goncalves, J. and Kostakos, V. (2011). Sharing ephemeral information in online social networks: Privacy perceptions and behaviours. In *INTERACT*, pp. 204–215. Springer: Lisbon, Portugal. doi:10.1007/978-3-642-23765-2\_14.
- Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. and Jacobs, D. (2012). Dynamic changes in motivation in collaborative citizen-science projects. In *CSCW '12: Proc. ACM. 2012 Conf. Computer Supported Cooperative Work*, pp. 217–226. doi:10.1145/2145204.2145238.

- Rushton, J.P., Chrisjohn, R.D. and Fekken, G.C. (1981). The altruistic personality and the self-report altruism scale. *Pers. Individ. Dif.*, 2, 293–302. Retrieved from <http://www.prenhall.com/divisions/hss/app/social/addchap4.html>.
- Schlegelmilch, B.B., Love, A. and Diamantopoulos, A. (1997). Responses to different charity appeals: The impact of donor characteristics on the amount of donations. *Eur. J. Mark.*, 31, 548–560. doi:10.1108/03090569710176574.
- Sheppard, S.A., Wiggins, A. and Terveen, L. (2014). Capturing quality: Retaining provenance for curated volunteer monitoring data. In *Proc. 17th ACM. Conf. Computer Supported Cooperative Work & Social Computing*, pp. 1234–1245. Retrieved from Google Scholar.
- Simmons, W.O. and Emanuele, R. (2007). Male-female giving differentials: Are women more altruistic? *J. Econ. Stud.*, 34, 534–550. doi:10.1108/01443580710830989.
- Spiekermann, S., Grossklags, J. and Berendt, B. (2001). E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *EC '01: Proc. 3rd ACM. Conf. Electronic Commerce*, pp. 38–47. ACM. doi:10.1145/501158.501163.
- Srnka, K.J., Grohs, R. and Eckler, I. (2003). Increasing fundraising efficiency by segmenting donors. *AMJ*, 11, 70–86. doi:10.1016/S1441-3582(03)70119-0.
- Staiano, J., Oliver, N., Lepri, B., de Oliveira, R., Caraviello, M. and Sebe, N. (2014). Money walks: A human-centric study on the economics of personal mobile data. In *UbiComp. ACM: Seattle, Washington, US*. doi:10.1145/2632048.2632074.
- Tedeschi, B.B. (2002). E-Commerce report; everybody talks about online privacy, but few do anything about it. *The New York Times*. Retrieved from NYTimes.
- Tseng, W. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Appl. Linguist.*, 27, 78–102. doi:10.1093/applin/ami046.
- Youyou, W., Kosinski, M. and Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans, *PNAS*, 112, 1036–1040.

## APPENDIX: MEASUREMENTS OF PERCEPTUAL VARIABLES

### *Organization reputation*

- (1) The image of the University of <city> is good.
- (2) I endorse the University of <city>'s notion of scientific research on human behavior.
- (3) The University of <city> has a good reputation in doing scientific research.

### *Privacy*

- (1) Participating in this research may cause a loss of my privacy.
- (2) My privacy information is in risk if I participate in this research.

### *Need for donation*

- (1) Donating some personal data for scientific research is important for future advance of science.
- (2) Donating some personal data for scientific research is critical for scientific advance.
- (3) It is necessary for people to donate some of their personal data to advance science.

### *Altruism*

- (1) I have given someone direction.
- (2) I have given money to someone who needed it (or asked for it).
- (3) I have done volunteer work for charity.

### *Attitude*

- (1) Donating personal data to scientific research is a good thing.
- (2) Conducting scientific research on a basis of real user data is a good idea.
- (3) It is a good thing to facilitate scientific research by donating personal data.

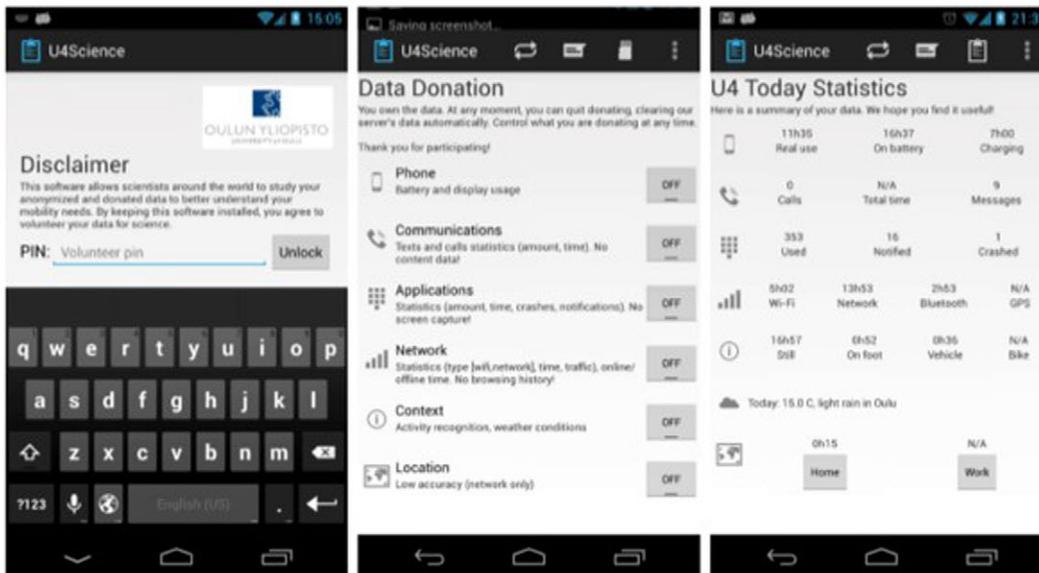
## Thank you for volunteering your data for science!

We are excited to hear that you want to help us change the world.

**DO NOT CLOSE THIS WINDOW YET!**

### Please follow these instructions:

- 1 Open link in your mobile phone browser: <http://tinyurl.com/U4SCI>
- 2 Install the software on your phone and open it (note: compatible with Android 2.3.3 or higher)
- 3 Insert this pin to unlock the software: **FYQP**
- 4 Select the data you wish to donate to science. **You can change this anytime you wish!**



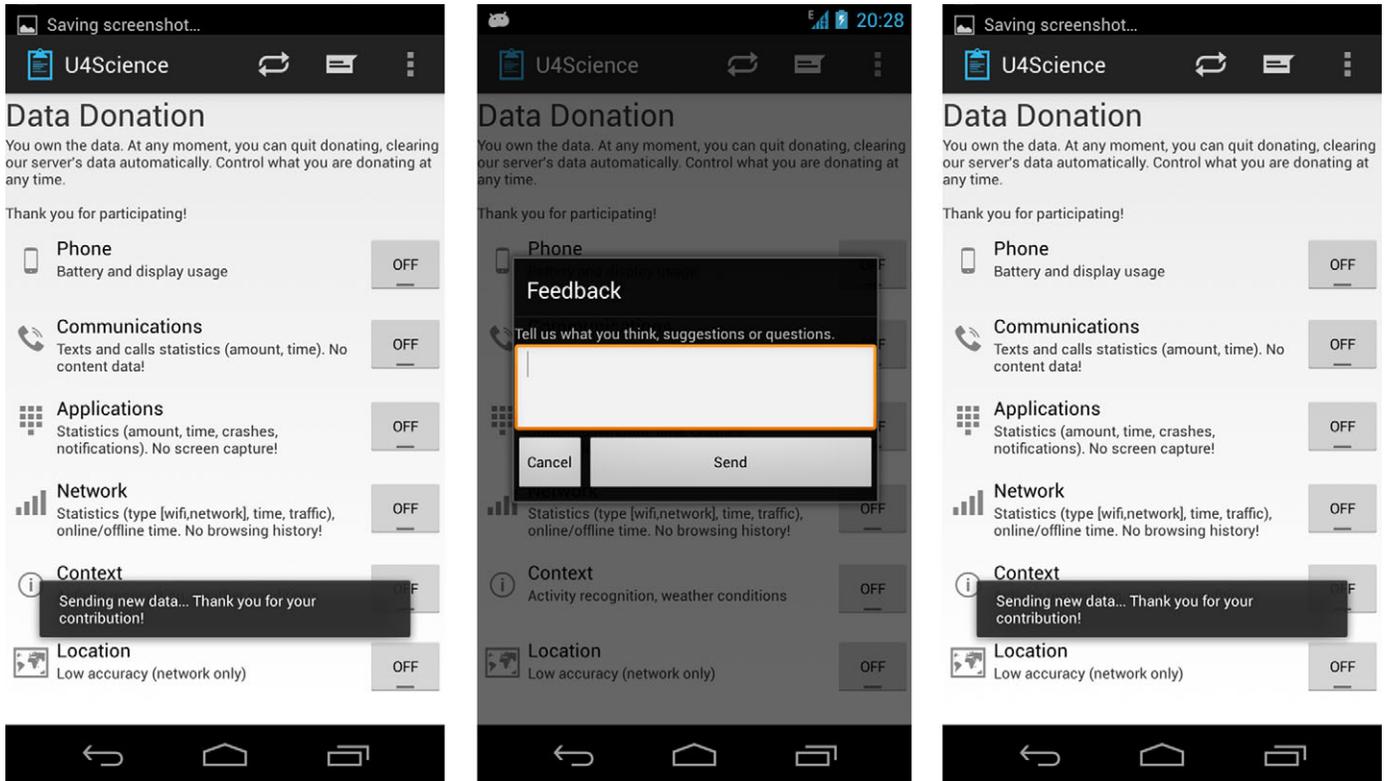
Depending on the data donations, you might require additional components:

- **Google Activity recognition:** you'll be prompted to install an energy efficient activity recognition algorithm provided by Google.
- **Weather:** you'll be prompted to install Weather, where you get current local weather reports from OpenWeather.
- **Google Fused Location:** you'll be prompted to instal an energy efficient location provider. We use it to help you find your work/home balance.

### Disclaimer

This software allows scientists around the world to use your **anonymous** donated data to better understand your mobility needs. By installing this software, you are agreeing to volunteer and donate your data for science. You are **free to remove** the application at **any** point in time or **modify** what data is donated to us.

**Figure A1.** The web page shown to participants who wished to donate data. The page provides step-by-step instructions on how to install our prototype.



**Figure A2.** U4Science's donor empowerment strategies. Left: provide feedback; Middle: quit study; Right: choose when to upload data.