



Utrecht University



# *Willingness and consent to data donation: Methodological and practical aspects*

Bella Struminskaya

Utrecht University

Data Donation Day at UU, March 11, 2022

Questions about COVID-19? VISIT [cdc.gov/coronavirus](https://www.cdc.gov/coronavirus) | [Español](#)



### Corona-Datenspende-App 2.0

Daten spenden und Umfragen beantworten - Ihr Beitrag gegen Corona

Corona-Datenspende - Infografik (Quelle: RKI) ... Resonanz und Teilnahme an der Corona-Datenspende-App. Bisher haben sich bereits mehr als 530.000 Personen (Stand: 30.09.2021) die App heruntergeladen, ihr Fitnessarmband oder ihre Smart Watch verknüpft und ihre Zustimmung zur wissenschaftlichen Datenauswertung gegeben. Herzlichen Dank dafür!



- 1. App herunterladen:** Installieren Sie die Datenspende-App einfach aus dem App Store oder dem Google Play Store auf Ihrem Smartphone.
- 2. Daten freigeben:** Stellen Sie die Datenschutzerklärung zu. In dieser wird erklärt, welche Daten genau gespeichert werden.
- 3. Fitnessarmband o. Smartwatch verbinden:** Unterstützt wird die Anbindung von Apple Health, Amazfit, Fitbit, Garmin, Google Fit, Huawei Health, Oura, Polar, Samsung Health und Withings.
- 4. An weiteren Studien teilnehmen:** Nehmen Sie an weiteren wissenschaftlichen Studien des Robert Koch-Instituts teil und helfen Sie bei der Betrachtung weiterer

## Center for Digital Health

Donate Your Data



why donate

a Penn Medicine initiative

### About Open Humans

Open Humans is dedicated to empowering individuals and communities around their personal data, to explore and share for the purposes of education, health, and research. We want to help people access and understand their personal data, and to help them do and share things that use that data.

For individuals, we have [community support for self-research](#), [tools for personal data access](#), and [data analysis notebooks](#) you can run in your browser.

For communities, we make it easy to choose to share your data with [group studies and activities](#). We provide the same features to researchers and citizen scientists alike: all members can adapt and share new [data analysis notebooks](#), as well as [create new group](#)

Coronavirus Digital health

# Data donation: better health and quality of life for all

Making a difference with data

Nora Tophof Maximilian Tischer

Mai 4, 2020

Estimated reading time 7 min.



Home

## The PHT concept

[Naar Nederlandstalige versie](#)



The Personal Health Train (PHT) is designed to enable health care innovators and researchers to work with health data from various

[Add data](#) [Explore & share](#) [Create](#) [Log in](#) [Join now!](#)

uring  
ual patients



o the datasource  
d to the research  
workflows' for  
of survival after  
data stations

# Data Donation (DD)

- Active consent to provide researchers with digitally collected personal data (digital platforms, apps and wearables, IoT, etc.)
- Often generated for purposes other than research
- Possible due to GDPR: individuals have access to a copy of their personal data

## Representativity concerns:

- coverage (not everyone has data to donate)
- willingness (not everyone wants to)
- participation (not everyone shares)

# Motivation to donate & reasons against DD

- 54% likely to donate for health research, 31% not likely (Skatova & Goulding 2019)
  - 40% of fitness tracker owners willing to donate (Toepoel et al. 2021)
  - 30% share Facebook data, 24-40% Twitter data, 60% Spotify (Silber et al. 2021)
- + Prosocial behavior (Skatova & Goulding 2019)
- + Insight into own results / quantified self (Bietz et al. 2019)
- Not gaining direct benefits from data donation (Skatova & Goulding 2019)
- Need to know the consequences of donation (Skatova & Goulding 2019)
- Similar to reasons to share/not share sensor & app data
  - Privacy concerns may play a role (e.g., for apps/sensors: Keusch et al. 2019; Struminskaya et al. 2020; Struminskaya et al. 2021)

Similarities to mechanisms of willingness to share and sharing of app and sensor data

### **Struminskaya et al. 2021**

- WTS & actual sharing
- Cross-section\* (NL)  
COOP2=54%
- GPS, photos, video; no app

### **Struminskaya et al. 2020**

- Willingness to share (WTS)
- Prob. LISS Panel (NL)  
2 waves, RR1 = 89%, 84%
- Share GPS, photos, video

### **Keusch et al. 2019**

- Willingness to share (WTS)
- Nonprob. panel (DE)  
2 waves
- Download tracking app

prev. exp., survey exp.

# Implementation (Struminskaya et al. 2021)



General consent



Framing, autonomy, & privacy explanation



GPS measurement



Photos & Video

### **Struminskaya et al. 2021**

- WTS & actual sharing
- Cross-section\* (NL)
- GPS, photos, video; no app
- Requests with rand. assig.:  
Autonomy over data collection  
Benefit framing  
Confidentiality assurance
- Fixed order of measurements
- Privacy concern, tech skills, prev. exp., survey exp.

### **Struminskaya et al. 2020**

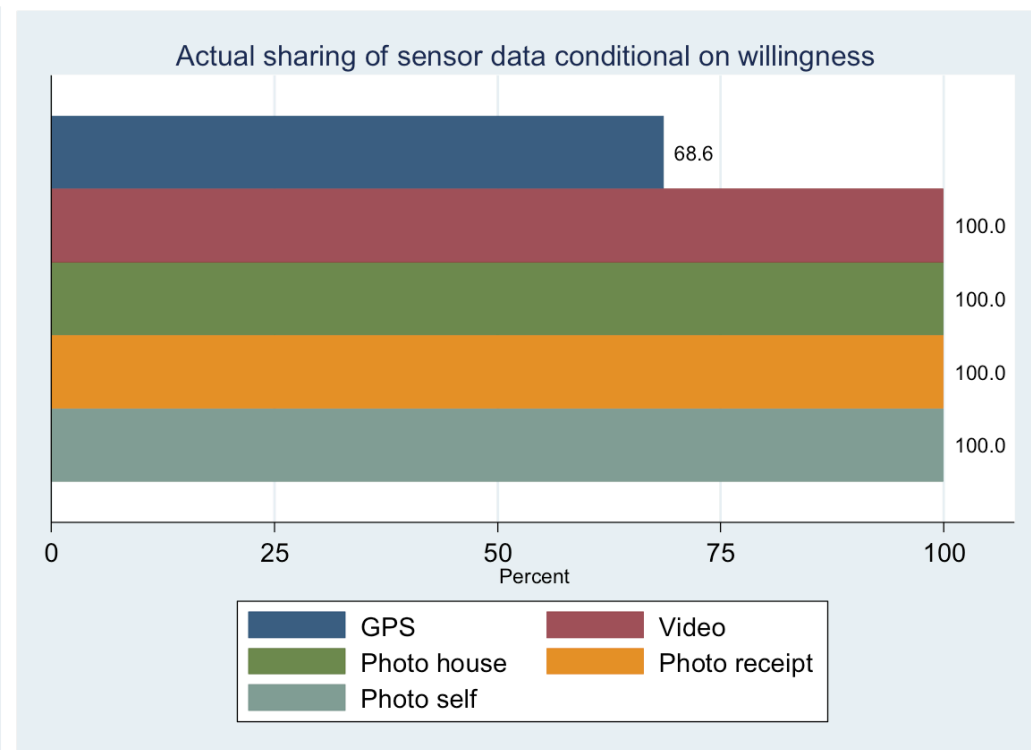
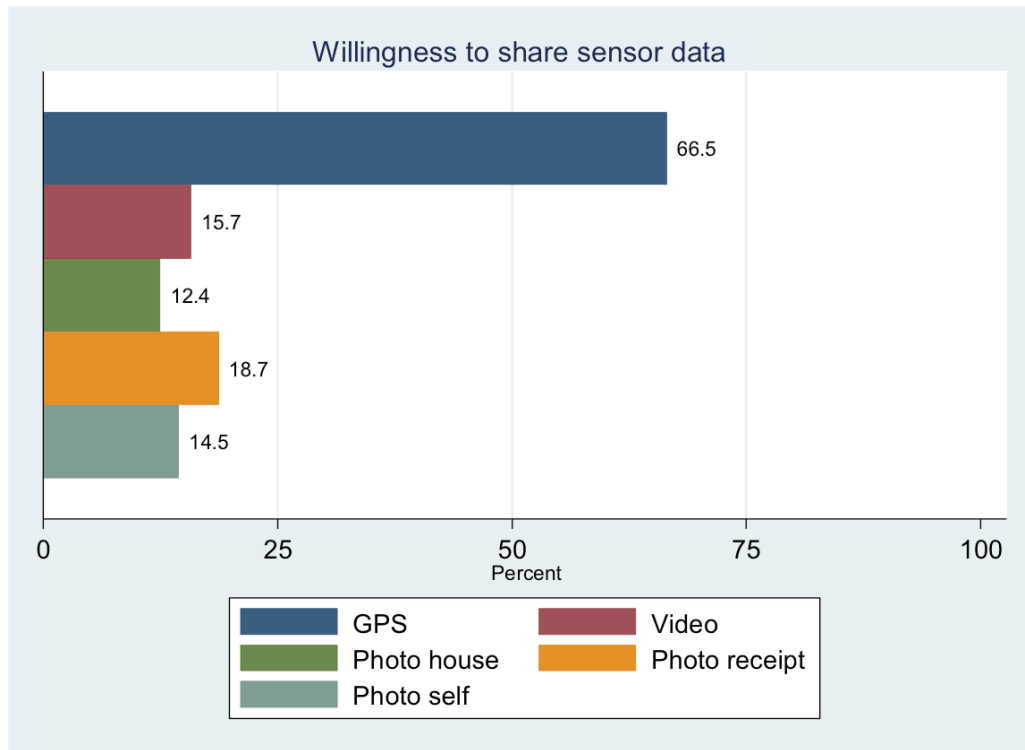
- Willingness to share (WTS)
- Prob. LISS Panel (NL)  
2 waves
- Share GPS, photos, video
- Vignettes w rand. assig.:  
Sponsor  
Autonomy over data collection  
Benefit framing  
Confidentiality assurance
- Randomized order of tasks
- Privacy concern, tech skills, prev. exp., survey exp.

### **Keusch et al. 2019**

- Willingness to share (WTS)
- Nonprob. panel (DE)  
2 waves
- Download tracking app
- Vignettes w rand. assig.:  
Sponsor  
Autonomy over data collection  
Duration  
Topic  
Incentive  
Questions in-app
- Randomized order of vignettes
- Privacy concern, tech skills, prev. exp., survey exp.



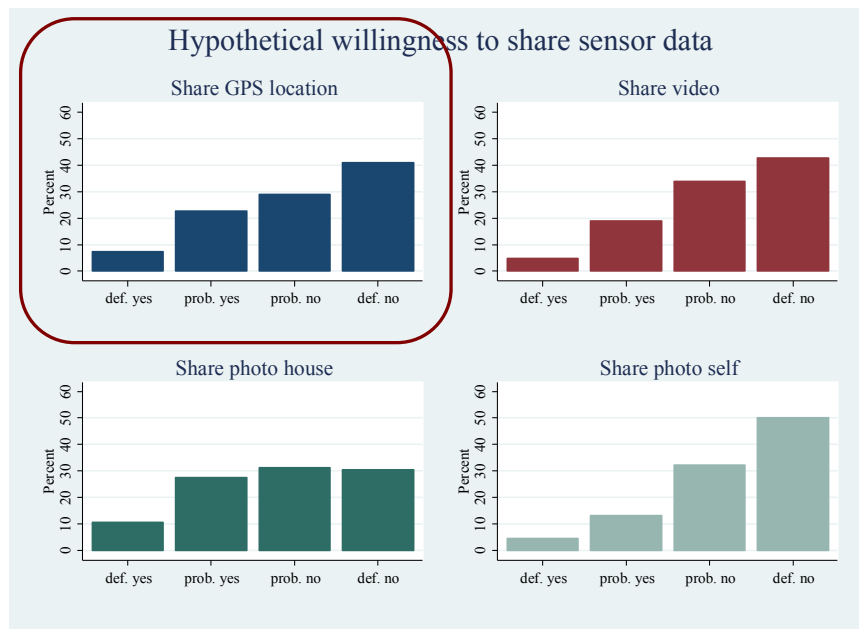
# Willingness and actual sharing (Dutch cross-section)



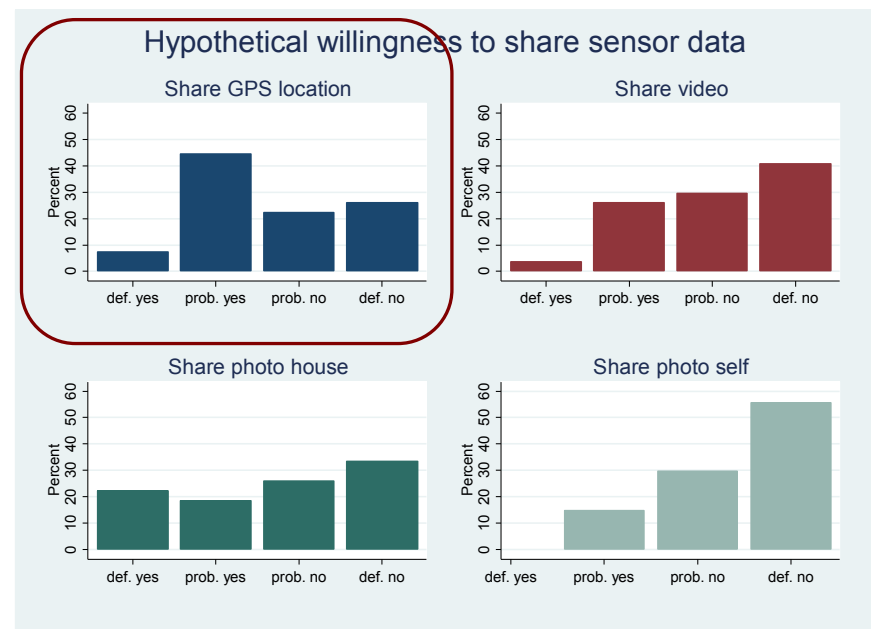
Participation rate GPS: 45.6%; n=1883 Dutch smartphone and tablet users

# Hypothetical willingness & Order effects

Overall, randomized order



Order: GPS, Video, Photo house, Photo self



Order effect: Average marginal effect +5.6 p.p  
(Struminskaya et al. 2020)

% Willing to share GPS:

- **If asked first: 41%**
- **If asked last: 26%**

# Willingness mechanisms

Predictors	WTS GPS	Share GPS	Share video	Share photo house	Share photo receipt	Share photo self
Benefit framing	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Autonomy over data collection	.11***	-.06*	n.s.	n.s.	.04*	n.s.
Privacy	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

n=1,853; Average marginal effects; covariates not shown

Predictors	Sharing
Order (asked first)	0.02 **
Sponsor University	0.09***
Sponsor Market Research	n.s.
Benefit framing	-0.02*
Autonomy over data collect.	n.s.
Privacy	n.s.

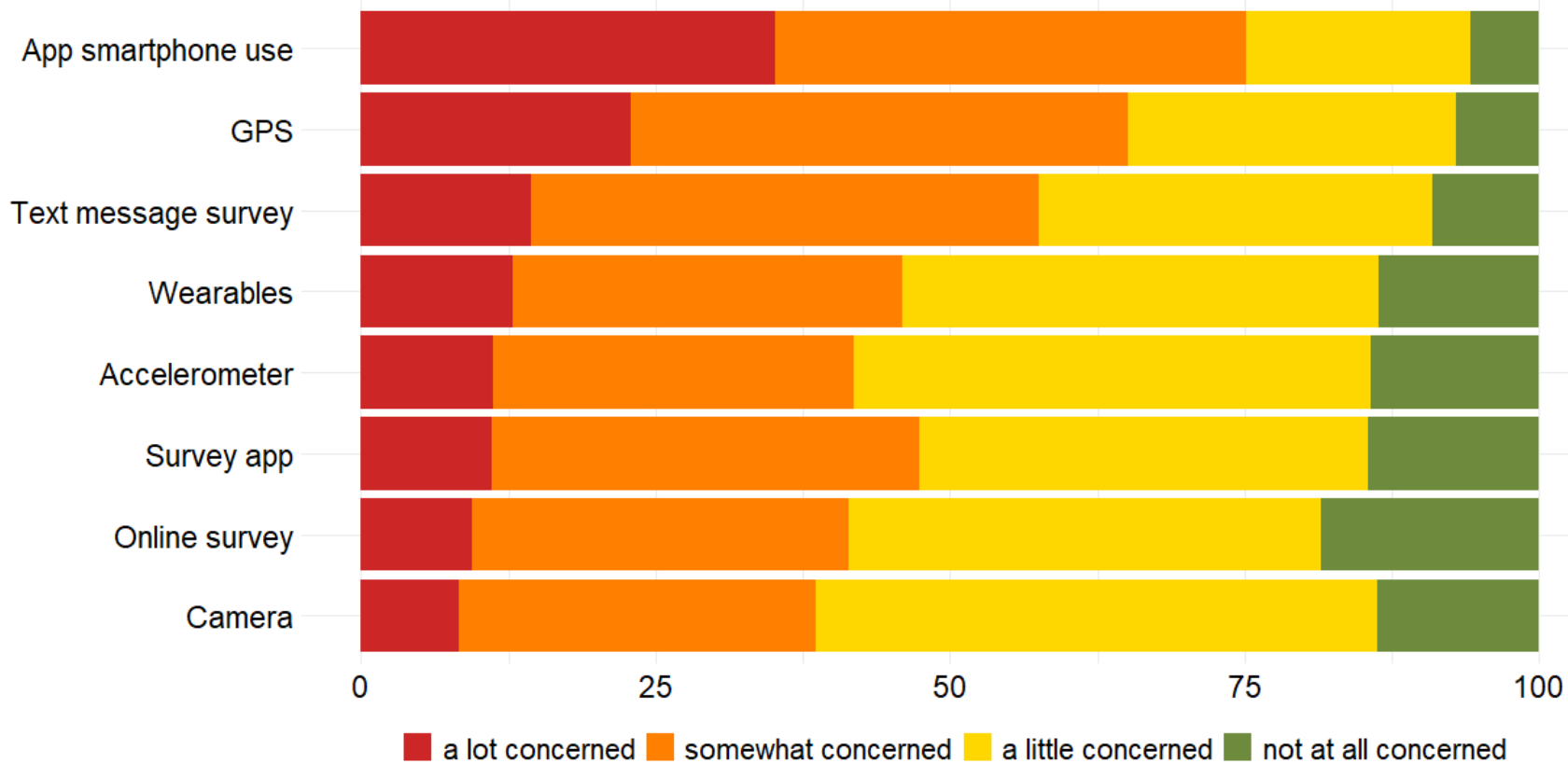
n=2,669; Average marginal effects; covariates not shown

In all 3 studies: sig. effects of smartphone use behaviors, mixed findings about the effect of privacy concerns, attitudes toward surveys, prior app download

(Struminskaya et al. 2020, 2021)

# Concern by Type of Collected Data

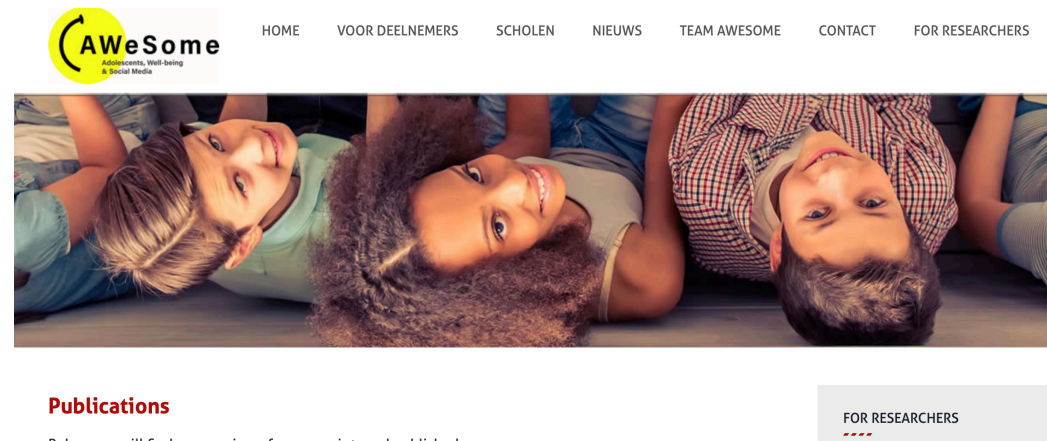
LISS Panel



# Selectivity in data donation in the AWeSome Study

# Is there selectivity in DD of social media data?

- [Project AWeSome](#) (Adolescents, Well-being, and Social Media) by University of Amsterdam
- Topics: social media use, well-being, social relationships, self-regulation
- Teenagers 13-15 yo in the Netherlands, recruited f2f at school, parental consent provided (N = 388)
- Asked to participate in:
  - surveys (16 biweekly, 6 months)
  - ESM (6 per day 7 days)
  - donate Instagram data (raw, deidentified pre-analysis)



# (1) Who donates data?

- 312 (80%) have at least one Instagram account
- 100 (32%) donated Instagram data
- 4 participants donated who said they don't have Instagram (n = 316)

% (n)	Donated?		
	Yes	No	Total
M (SD)			
<i>Social characteristics</i>			
Social comparison	3.5 (1.4)	3.3 (1.4)	3.3 (1.4)
How many good friends	15.8 (78.8)	9.2 (6.5)	11.4 (45.5)
Friendship quality	4.2 (0.6)	4.2 (0.5)	4.2 (0.5)
Parental phone restrictions	2.4 (1.5)	2.1 (1.5)	2.2 (1.5)
Parental knowledge	3.5 (0.7)	3.3 (0.6)	3.4 (0.6)
Adolescent disclosure and secrecy	2.1 (1.1)	1.8 (0.8)	1.9 (1.0)

Disclosure & secrecy  
(teens tell parents what  
they do on SM)  
predictive of DD  
( $p=.005$ )

## (2) Who donates data?

% (n)	Donated?			
	Yes	No	Total	
<i>Socio-demographic characteristics</i>				
Sex	Female	21.8 (69)	37.3 (118)	59.2 (187)
	Male	10.8 (34)	29.8 (94)	40.5 (128)
	Other	0.3 (1)	0.0 (0)	0.3 (1)
Grade	8	17.7 (56)	37.0 (117)	54.8 (173)
	9	15.2 (48)	30.1 (95)	45.3 (143)
Level of education	prevocational	10.8 (34)	29.8 (94)	40.5 (128)
	general	11.1 (35)	19.9 (61)	31.0 (96)
	academic	11.1 (35)	17.4 (54)	28.5 (89)

37% of females  
donate data vs.  
27% of males ( $p=.04$ )

Self-esteem somewhat  
lower for those who  
donate ( $p=.027$ )

<i>Psychological characteristics</i>			
Affective well-being	5.6 (1.2)	5.7 (1.2)	5.7 (1.2)
Cognitive well-being	5.5 (2.9)	5.4 (3.1)	5.5 (3.0)
Mobile phone use self-monitoring	2.5 (0.8)	2.6 (0.7)	2.5 (0.8)
Positive affect	3.6 (0.6)	3.7 (0.6)	3.7 (0.6)
Negative affect	2.3 (0.7)	2.2 (0.6)	2.2 (0.7)
Self-esteem	3.8 (1.0)	4.0 (0.7)	4.0 (0.8)
Loneliness	2.0 (1.0)	1.8 (0.9)	1.9 (0.9)
Self-regulation	3.5 (0.7)	3.3 (0.6)	3.4 (0.6)



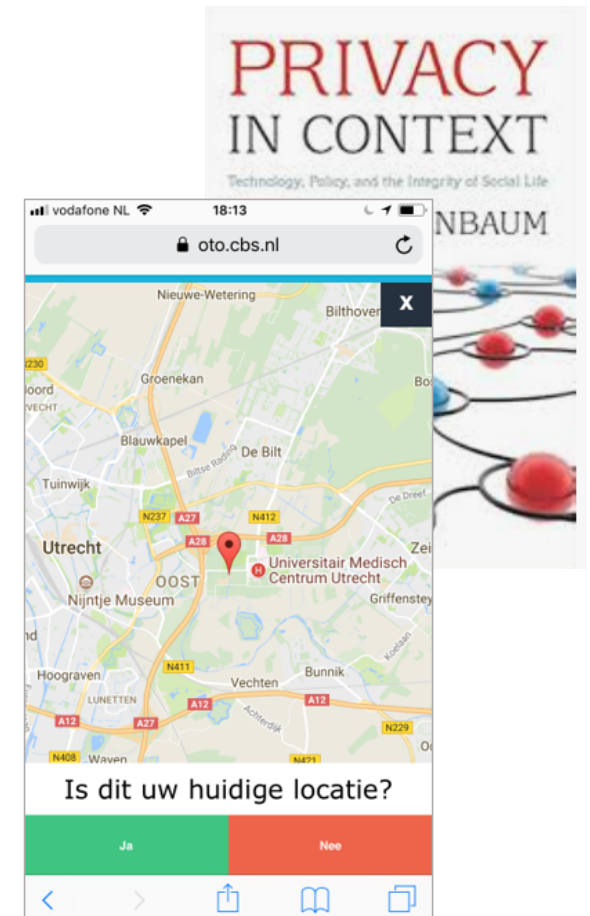
### (3) Who donates data?

% (n)	Donated?			
	Yes	No	Total	
<i>Device / Study participation</i>				
Mobile phone type	Android	19.0 (56)	37.3 (110)	56.3 (166)
	iPhone	15.3 (45)	28.5 (84)	43.7 (129)
M (SD)				
	Completed ESM 1	97.9 (19.2)	93.6 (21.2)	95.1 (20.6)
	Completed ESM 2	83.5 (30.6)	71.9 (34.7)	76.0 (33.7)
	Completed surveys	15.1 (1.9)	14.4 (2.6)	14.7 (2.4)

Prior participation is predictive of donation  
( $p=.004$  &  $p=.007$ )

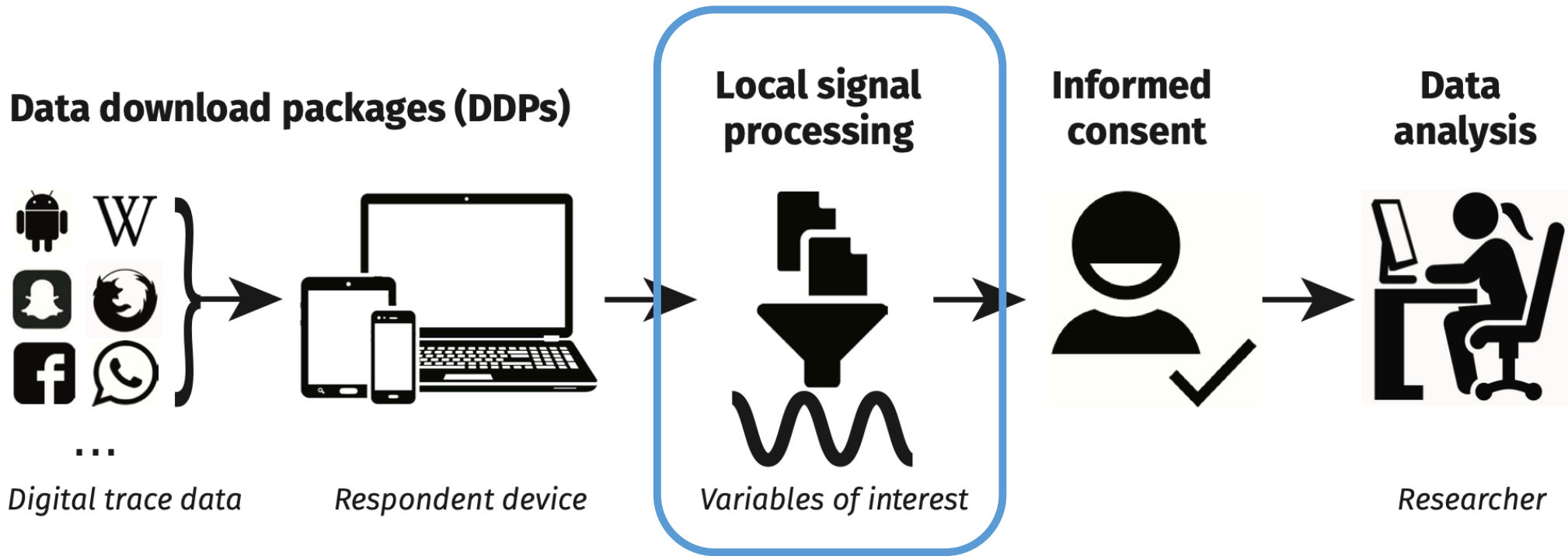
# Summary so far

- Decisions about sharing seem to be situation-specific, nuanced
- Hypothetical behavior (willingness) differs from actual participation behavior
- The nature of the task is relevant
- Clear communication of who asks to share & for what purpose
- Balance between maximizing sharing and providing detailed information about the data (“backfire effects”)
- Ceiling effects possible due to loyalty, trust in sponsor



“Is this your current location? Yes/No”

# Privacy-preserving Data Donation Workflow



# Google Location History data donation pilot

# DD locally processed Google Location History

- Pilot study w/friends, colleagues, family (n about 50)
- Goal: test the workflow including ERB, performance of the local processing
- Extracted number of hours/km travelled in different modes
- Intake survey
- Data Donation
- Evaluation survey

Dear friends, family and colleagues,

We would like to ask you to participate in our research project on trying out new technologies to study people's daily movement and travel behavior. Recently, we developed software that allows us to investigate Google Location History data from research participants while preserving the privacy of the participants. We call this software 'PORT'. Hereby, we would like to invite you to try out PORT for the first time!

This project consists of two questionnaires and trying out PORT, software developed by Utrecht University and Eyra. To use PORT, you first request your Google Location History at Google. Next, you download this and save it on your computer. Next, you load it into PORT, and PORT will only extract data on various activity types that you have undertaken, such as the number of hours you spent biking, walking or taking the train between 2016 and 2021. An example of how this looks like is shown below.

## Travelled by bike

Year	Month	Nr. of hours
2017	JANUARY	8,85
2017	FEBRUARY	13,40
2017	MARCH	12,43
2017	APRIL	16,47
2017	MAY	30,01
2017	JUNE	7,12
2017	JULY	7,37
2017	AUGUST	10,47
2017	SEPTEMBER	5,34
2017	OCTOBER	14,70
2017	NOVEMBER	7,91
2017	DECEMBER	10,38

Although Google Location History contains detailed information on all the locations you visited, this will not be shared with us. Only the information similar to what is shown in the table (number of

# Intake survey

[https://  
survey.uu.nl/  
jfe/form/  
SV\\_afxB0W3X  
0xZ0kDA](https://survey.uu.nl/jfe/form/SV_afxB0W3X0xZ0kDA)

- Smartphone type
- Smartphone use behaviors (idle/leaving at home, etc.)
- Hours biking/vehicle travel/public transport/on foot/still during the week
- Hours biking/vehicle travel/public transport/on foot/still during the weekend
- How worried about privacy
- Frequency of using location-aware apps
- Self-rated smartphone skills
- Basic demographics

Thank you for participating in our methodological research project on trying out new technologies to study people's daily movement and travel behavior!

The first step in the process is to fill out the questionnaire. It will take you about 5-7 minutes. There are no right or wrong answers to the questions, and you can stop participating at any time. The data that you provide in this questionnaire will be only used in aggregated form and only the researchers from this project will have access to the information you provide. This survey is anonymous and personal information can not be inferred from your answers.

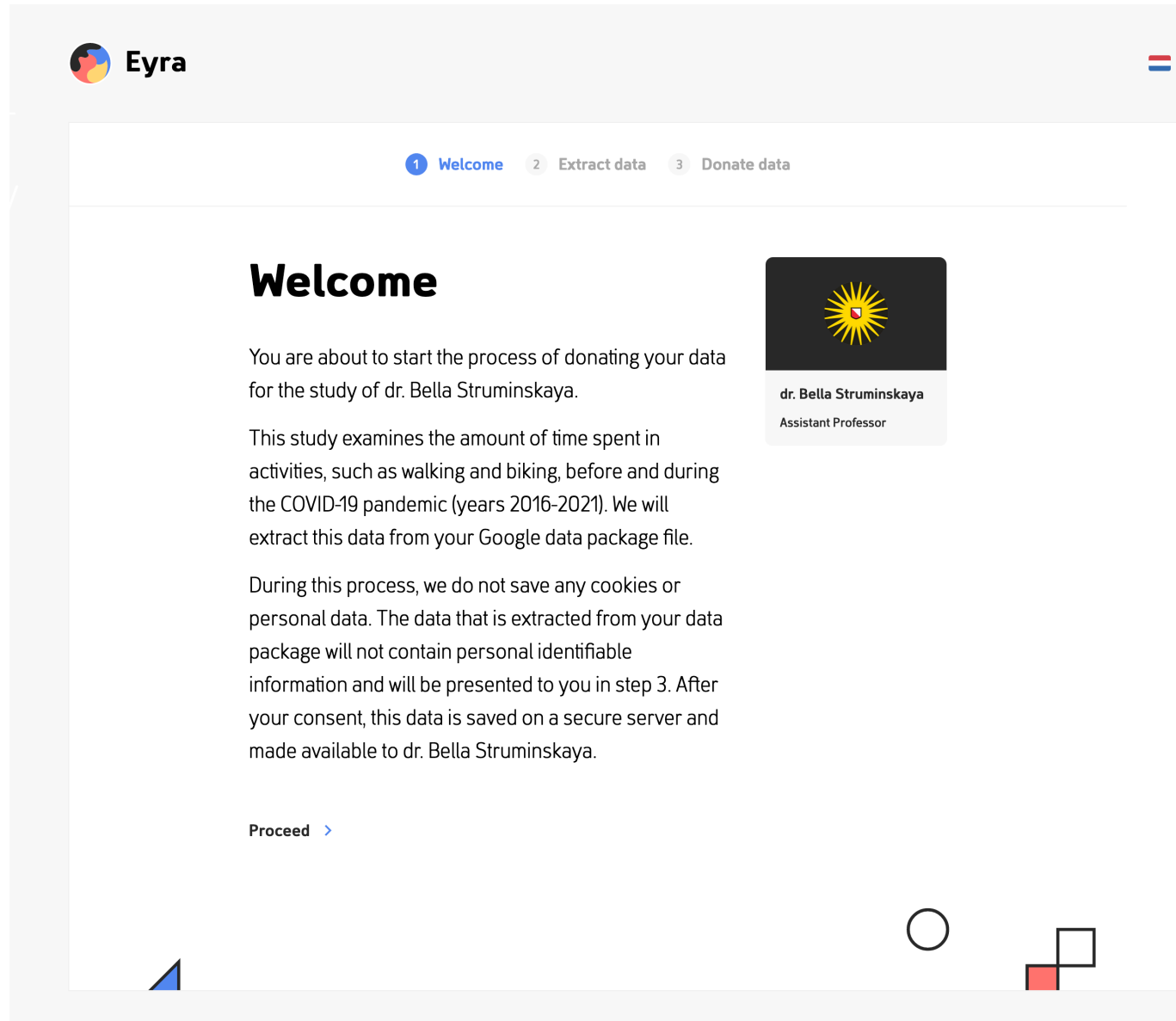
If you have any questions about this project, please contact Bella Struminskaya at [b.struminskaya@uu.nl](mailto:b.struminskaya@uu.nl) or Laura Boeschoten at [l.boeschoten@uu.nl](mailto:l.boeschoten@uu.nl).

If you have complaints about this study, please email to this address: [klachtenfunctionaris-fetcsocwet@uu.nl](mailto:klachtenfunctionaris-fetcsocwet@uu.nl).

Please enter your personal code from the invitation email to start the survey.

Personal code (6 digits)

# Donation



The screenshot shows a web interface for 'Eyra'. At the top left is the 'Eyra' logo, which consists of a colorful globe icon followed by the text 'Eyra'. In the top right corner, there is a small flag icon. Below the header is a progress bar with three steps: '1 Welcome' (highlighted in blue), '2 Extract data', and '3 Donate data'. The main content area has a large 'Welcome' heading. Below it, there are three paragraphs of text explaining the data donation process. To the right of the text is a profile card for 'dr. Bella Struminskaya', Assistant Professor, featuring a yellow sunburst icon on a dark background. At the bottom of the page, there is a 'Proceed >' button and a decorative footer with a blue triangle, a white circle, and a red square.

**Eyra**

1 Welcome 2 Extract data 3 Donate data

## Welcome

You are about to start the process of donating your data for the study of dr. Bella Struminskaya.

This study examines the amount of time spent in activities, such as walking and biking, before and during the COVID-19 pandemic (years 2016-2021). We will extract this data from your Google data package file.

During this process, we do not save any cookies or personal data. The data that is extracted from your data package will not contain personal identifiable information and will be presented to you in step 3. After your consent, this data is saved on a secure server and made available to dr. Bella Struminskaya.

[Proceed >](#)

dr. Bella Struminskaya  
Assistant Professor

# Evaluation survey

[https://survey.uu.nl/jfe/form/SV\\_aeFK7wZkc](https://survey.uu.nl/jfe/form/SV_aeFK7wZkc)



- Difficulty to understand instructions
- Difficulty of upload /
- Reasons for not uploading
- Technical difficulties
- Android OS version

Thank you for participating in this short evaluation questionnaire. We would like to know how you experienced the data donation process. It will take you about 3 minutes to answer the questions.

Please enter your personal code from the invitation email to start the survey.

personal code (6 digits)

Next

Pov

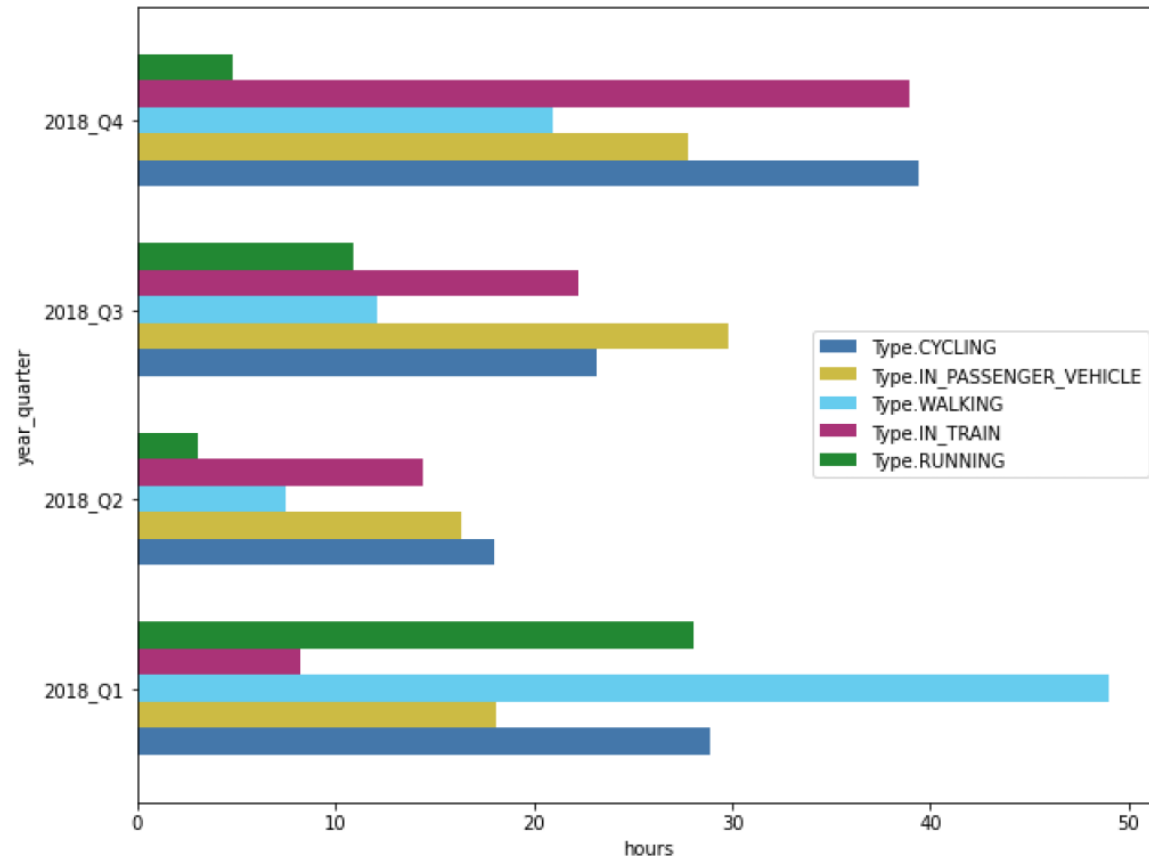


# Ethical Review Board Clearance (for the pilot)

- Iterative process
- Took about 7 weeks to address
- Examples of requests for clarification:
  - Whether the data is anonymized/pseudoanonymized
  - Only to state that the data is anonymous during publication/archiving, not for any of the surveys or the donation
  - As there is more data in the package that the subject uploads to the data donation environment → Who is responsible for the data in that step (UU/Eyra)? How are the risks minimized during that step? (DPIA)

# Future steps

- CentERpanel methodological study (Google Location History)
- Alternative visualization (transparency / usability)
- Randomized experiments on visualization & amount of info presented /consent



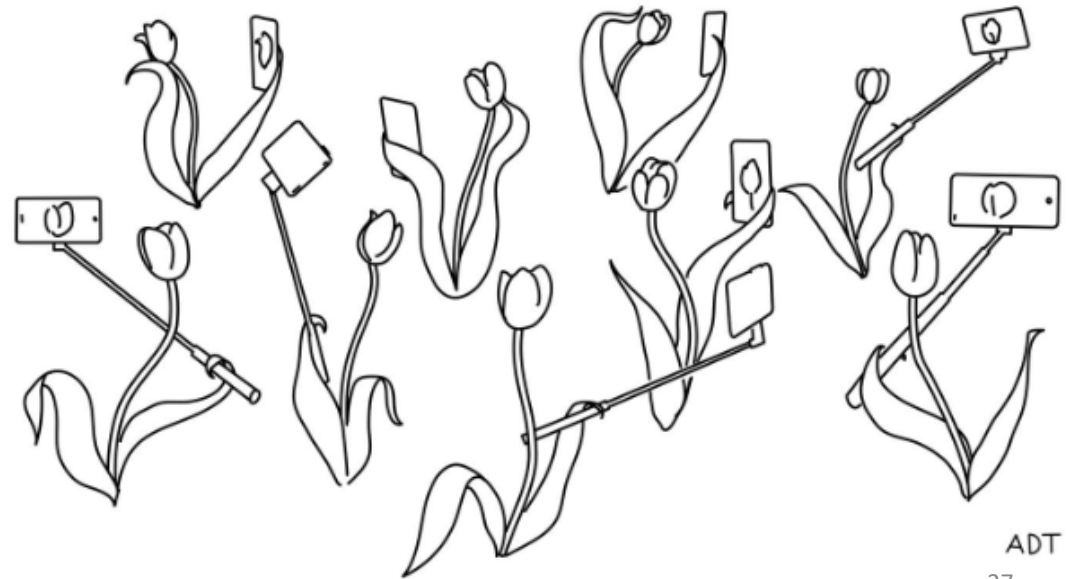
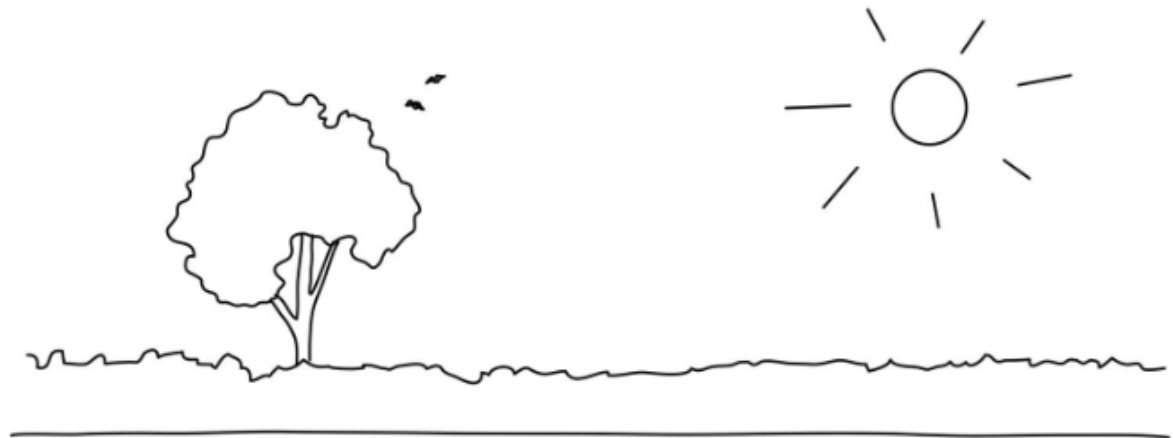
# Thank you!

Contact:

[b.struminskaya@uu.nl](mailto:b.struminskaya@uu.nl)

<https://bellastrum.com>

[@bellastrum](#)



ADT

27

[A.D. Thompson, The New Yorker](#)

# References:

- Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., Dolmans, R. (2021). Sharing data collected with smartphone sensors: Willingness, participation, and non-participation bias. *Public Opinion Quarterly*. <https://doi.org/10.1093/poq/nfab025>
- Struminskaya, B., Toepoel, V., Lugtig, P., Haan, M., Luiten, A., Schouten, B. (2020). Understanding willingness to share smartphone-sensor data. *Public Opinion Quarterly* 84(3), 725-759. <https://doi.org/10.1093/poq/nfaa044>
- Keusch, F., Struminskaya, B., Antoun, C., Couper, M., Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly* 83(S1), 210-235. <https://doi.org/10.1093/poq/nfz007>
- Bietz et al. (2019): <https://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.178/>
- Skatova & Goulding (2019): <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0224240>
- Boeschoten et al. (2020): <https://arxiv.org/pdf/2011.09851.pdf>
- Silber et al. (2021): [https://madoc.bib.uni-mannheim.de/58906/1/Linking%20surveys%20and%20digital%20trace%20data\\_Silber\\_et\\_al\\_2021.pdf](https://madoc.bib.uni-mannheim.de/58906/1/Linking%20surveys%20and%20digital%20trace%20data_Silber_et_al_2021.pdf)
- Toepoel et al. (2021): <https://www.surveypractice.org/article/22247-response-willingness-and-data-donation-in-a-study-on-accelerometer-possession-in-the-general-population>