



From Surveys to Digital Traces

Insights from Data Donation Studies in the Dutch LISS Panel

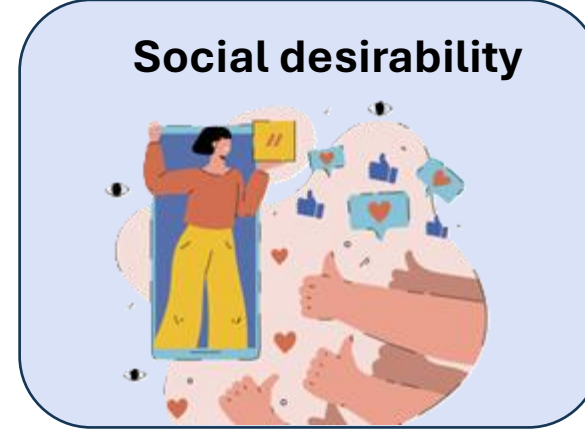


Joris Mulder
senior researcher
LISS quality & innovation lead

Why surveys are not always enough



- Surveys successfully used for collecting self-reports, perceptions and opinions
- But prone to:



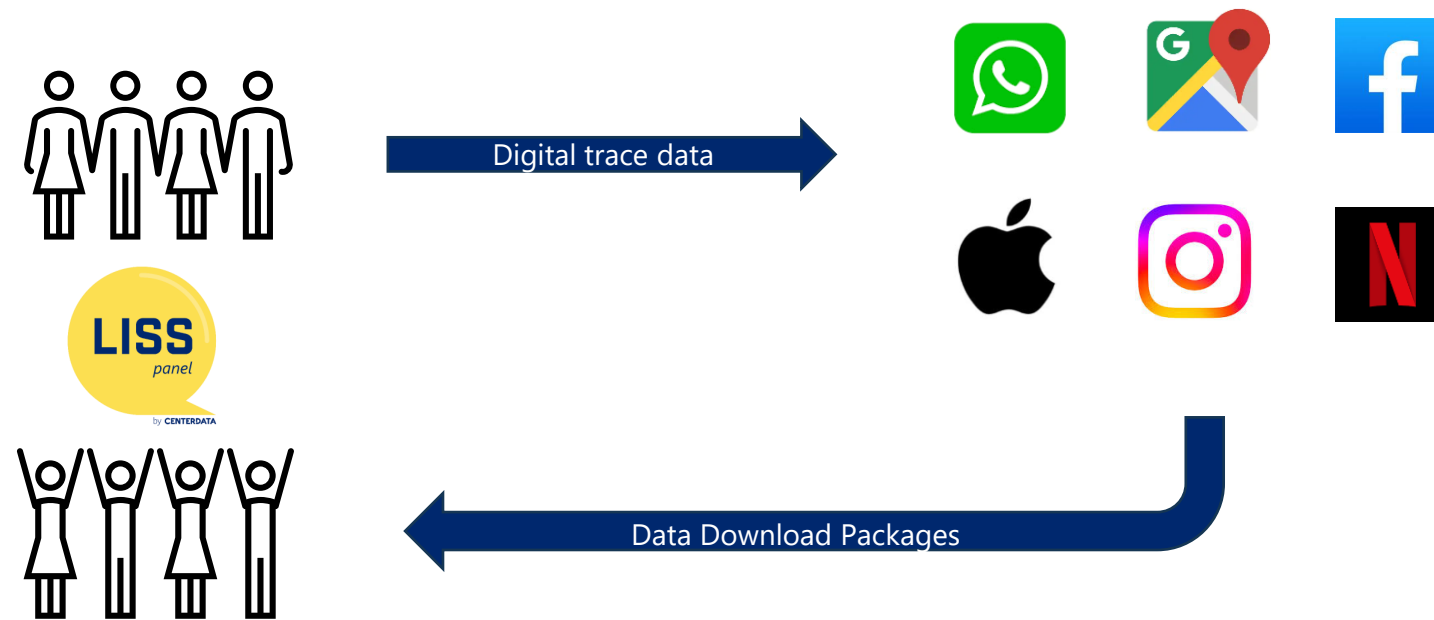
- Many behaviors **occur digitally** nowadays
- **Digital traces** can substantially **enrich survey data**



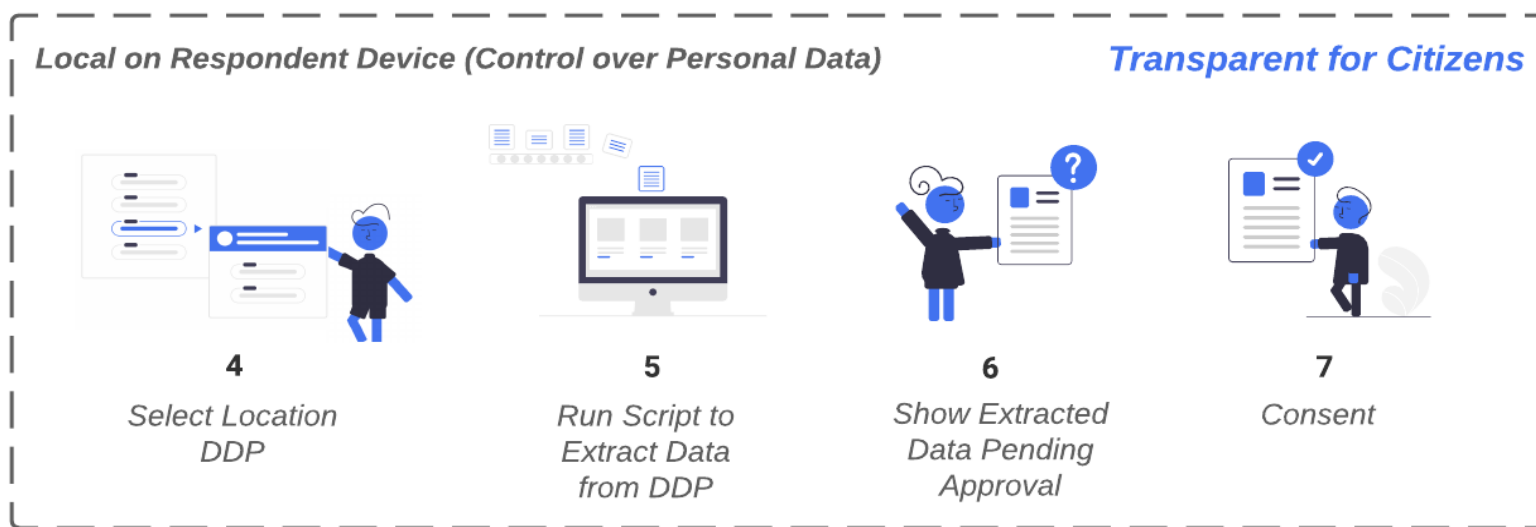
Data Donation as a method

Takes advantage of GDPR (EU)

- *Right of access by the data subject & data portability (articles 15 & 20)*
 - Receive personal data in structured **Data Download Package** (DDP)
 - In the US? Most major platforms now provide DDPs globally
 - Ask panel members to request data package and donate for research



Port: a privacy preserving donation tool



<https://github.com/eyra/feldspar>

Boeschoten, L., De Schipper, N.C., Mendrik, A.M., Van der Veen, E., Struminskaya, B., Janssen, H. and Araujo, T. (2023). [Port: A software tool for digital data donation](#). Journal of Open Source Software, 8(90), 5596.

Whitelist approach

- **Extract only** what you currently need
- Selected datapoints
- Privacy-by-design

For instance, **only**:

- Number of contacts in WhatsApp group
- Number of steps per day

Blacklist approach

- **Extract everything** in the DDP for future use
- Minus Personal Identifiable Information
- Extra cleaning may be necessary

For instance, **all**:

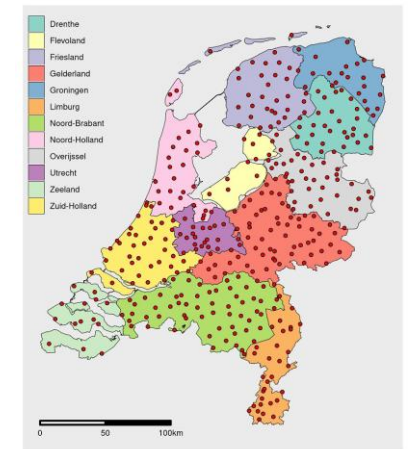
- Likes, comments & posts on Insta & TikTok
- Visited URLs

The methodological choice determines privacy risk, data richness, and cleaning complexity



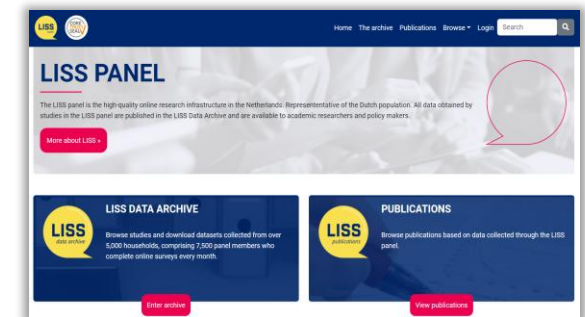
Integration into the LISS panel

- Probability-based sample Statistics Netherlands (CBS)
- Collecting longitudinal data since 2007
- ~ 7,500 panel members
- 70 - 80% monthly survey response
- Open data in LISS Data Archive: <https://lissdata.nl>
- Data linkable to CBS register data



Integration of Port for Data Donation allows us to:

- Test it within a high-quality panel infrastructure
- Enrich longitudinal survey data with digital trace data



Three Data Donation projects



Studying Social Networks in WhatsApp

- Communication patterns
- Group dynamics

Corten et al. (2026)

<https://doi.org/10.1007/s13278-025-01550-8>



Physical Activity among 50+ adults

- Apple Health
- Samsung Health
- Google Location History

Keusch et al. (2025)

<https://www.jmir.org/2025/1/e69799>



Public perceptions and experiences in society with.

- Artificial Intelligence
- Automated Decision Making systems

AlgoSoc.org

<https://doi.org/10.1080/13645579.2025.2568497>



Study 1: WhatsApp group chat donation



Raw DDP

```

_chat - Notepad
File Edit Format View Help
[21/01/2022, 17:55:56] You created group "UU WhatsApp p
[21/01/2022, 17:55:56] UU WhatsApp project [x]: Messages
[21/01/2022, 18:02:32] Research Colleague 1: Hi!
[21/01/2022, 18:02:53] Laura: 🙋
[21/01/2022, 18:03:18] Laura: Javier heeft een speciale
[21/01/2022, 18:03:34] Research Colleague 1: Ow OK
[21/01/2022, 18:03:46] Laura: Dus die komt later
[21/01/2022, 18:15:12] Research Colleague 1: 👍
[21/01/2022, 19:56:11] Research Colleague 2 joined usin
[21/01/2022, 21:47:27] Research Colleague 1: Welcome Ja
[22/01/2022, 13:10:31] Research Colleague 2: Thanks! Is
[22/01/2022, 13:24:17] Research Colleague 1: 🙋
[22/01/2022, 13:24:29] Research Colleague 1: You can of
[22/01/2022, 13:51:43] Laura: You are also allowed to t
[22/01/2022, 13:52:09] Laura: image omitted
[22/01/2022, 19:38:02] Research Colleague 2: Het ziet e
[17/02/2022, 13:25:15] Research Colleague 2 added Resea
[16/03/2022, 15:16:30] Research Engineer 2 : Hi Laura,
[16/03/2022, 15:17:11] Research Engineer 2 : I will sh
[16/03/2022, 15:17:32] Laura: I made you admin
[16/03/2022, 15:17:45] Research Engineer 2 : Great, th
[16/03/2022, 15:18:27] Research Engineer 2 added Rese

```

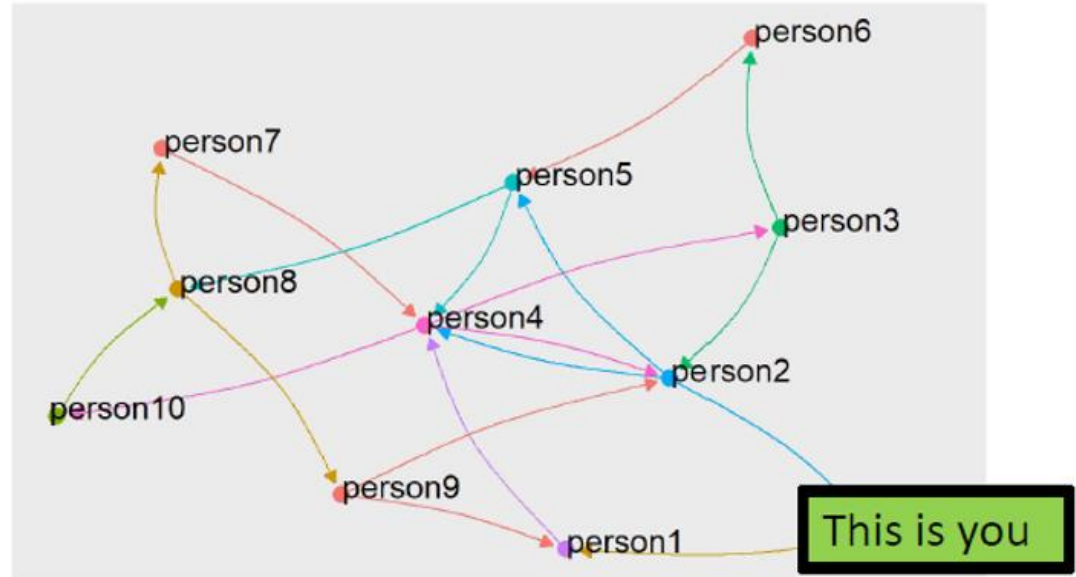
This is you

Description	Value
0 Total number of words	3177
1 Number of messages	330
2 Date first message	2020-03-13 14:08:27
3 Date last message	2020-08-20 09:46:39
4 Number of URLs	1
5 Who replies to you the most often?	person1
6 Who do you most often reply to?	person1

person1

Description	Value
0 Total number of words	3505
1 Number of messages	447
2 Date first message	2020-03-13 14:08:04
3 Date last message	2020-08-20 09:03:01
4 Number of URLs	12
5 Number of shared files	3
6 Who replies to you the most often?	person3
7 Who do you most often reply to?	person8

Processed data





Study 1: WhatsApp donation response and bias



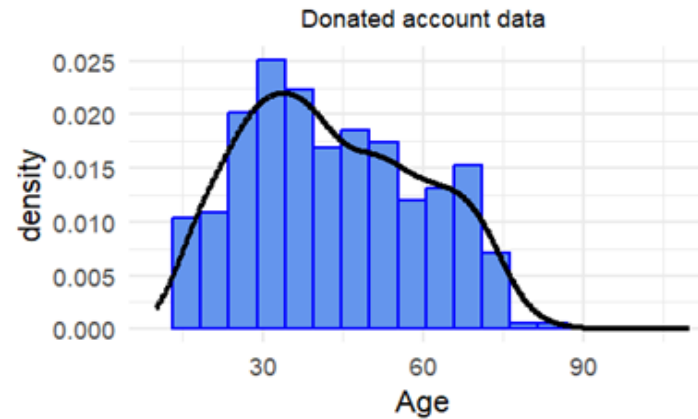
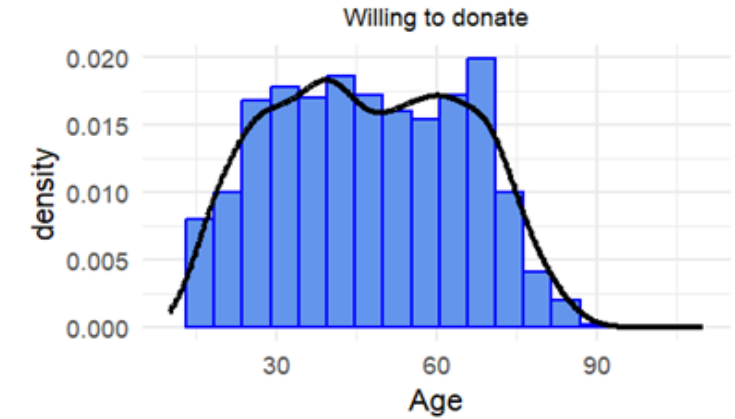
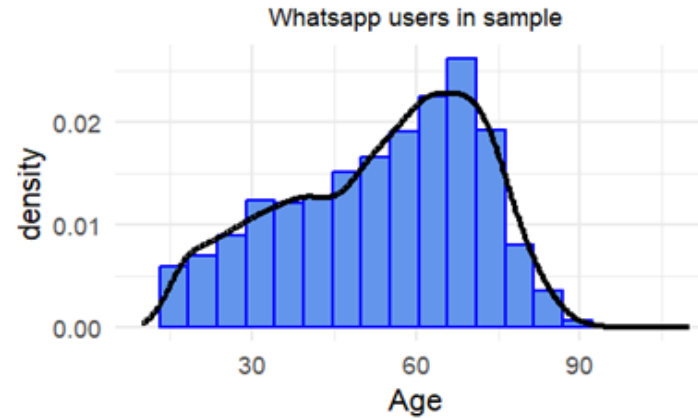
N = 4800 gross sample (100%)

N = 3598 participated in survey (75%)

N = 929 consented to donate (19%)

N = 492 donated at least 1 DDP (10%)

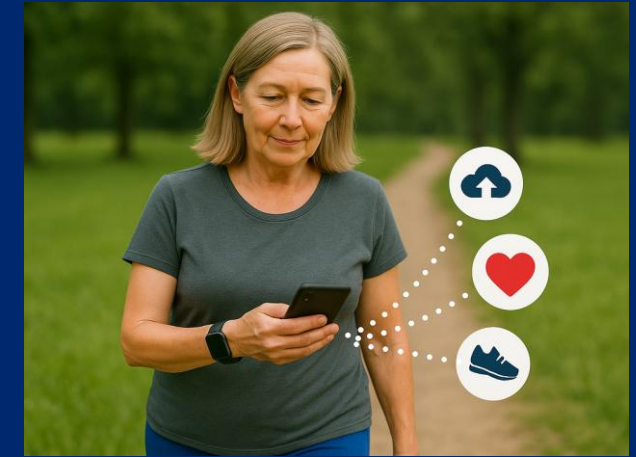
809 DDPs from 492 donors



Whatsapp users in sample
Willing to donate
Donated account data

	Mean age	Sd
Whatsapp users in sample	53.634	17.72
Willing to donate	47.364	17.61
Donated account data	43.266	16.42

Study 2: Measuring Physical Activity in Older Adults



Source: Image created by the Authors using DALL-E;
Copyright: N/A (AI-generated image); License: Public Domain (CC0)

Aantal stappen

< 1 2 3 4 5 6 7 > 323 pagina's Zoeken

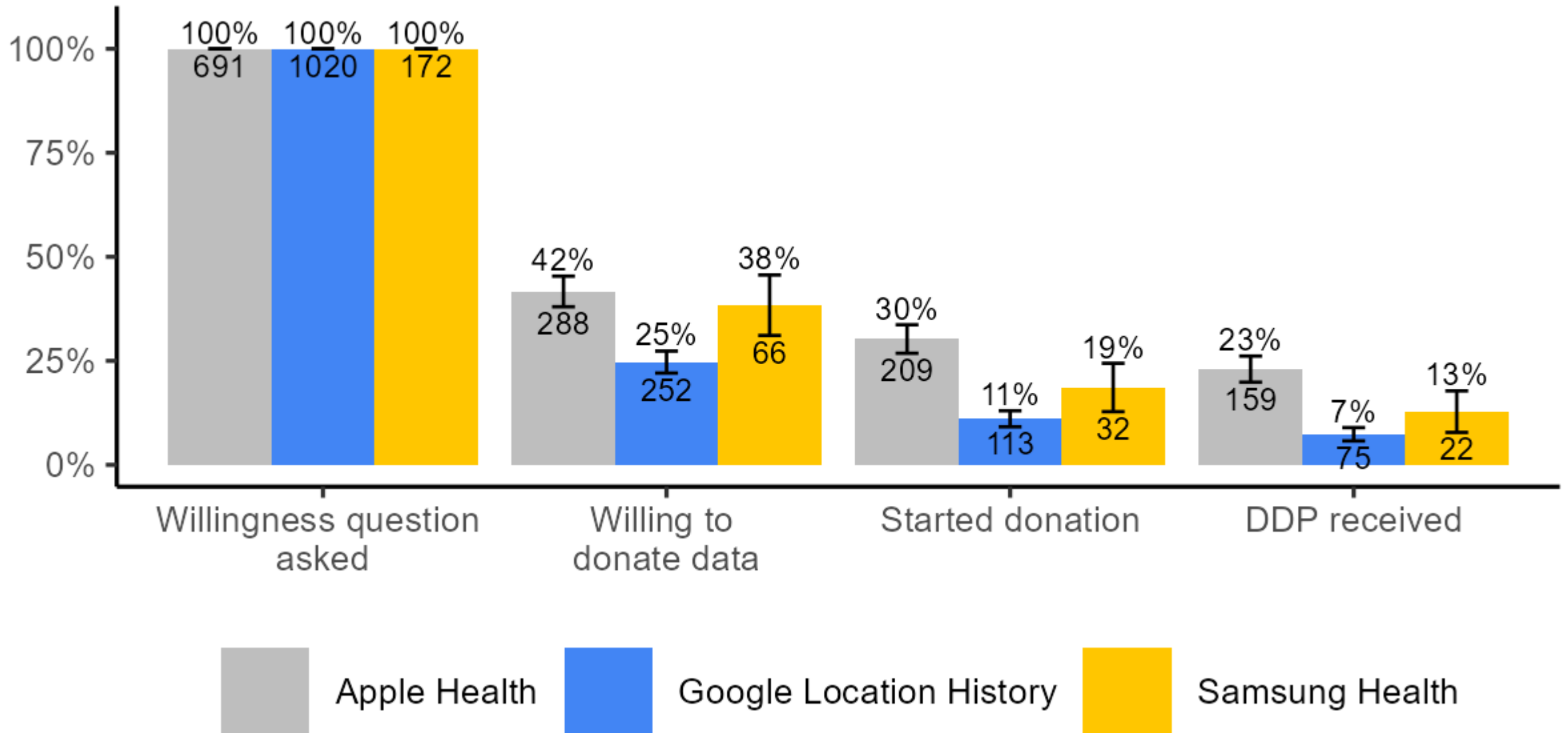
Datum	Aantal stappen
2017-01-01	2513
2017-01-02	3735
2017-01-03	3543
2017-01-04	4515
2017-01-05	2867
2017-01-06	3058
2017-01-07	2992

LISS participants (50+ years) donated

- Apple Health DDP
- Samsung Health DDP
- Google Location DDP

Study 2: Measuring Physical Activity in Older Adults

Participation by Health App (n=1883)



Bias in who donates

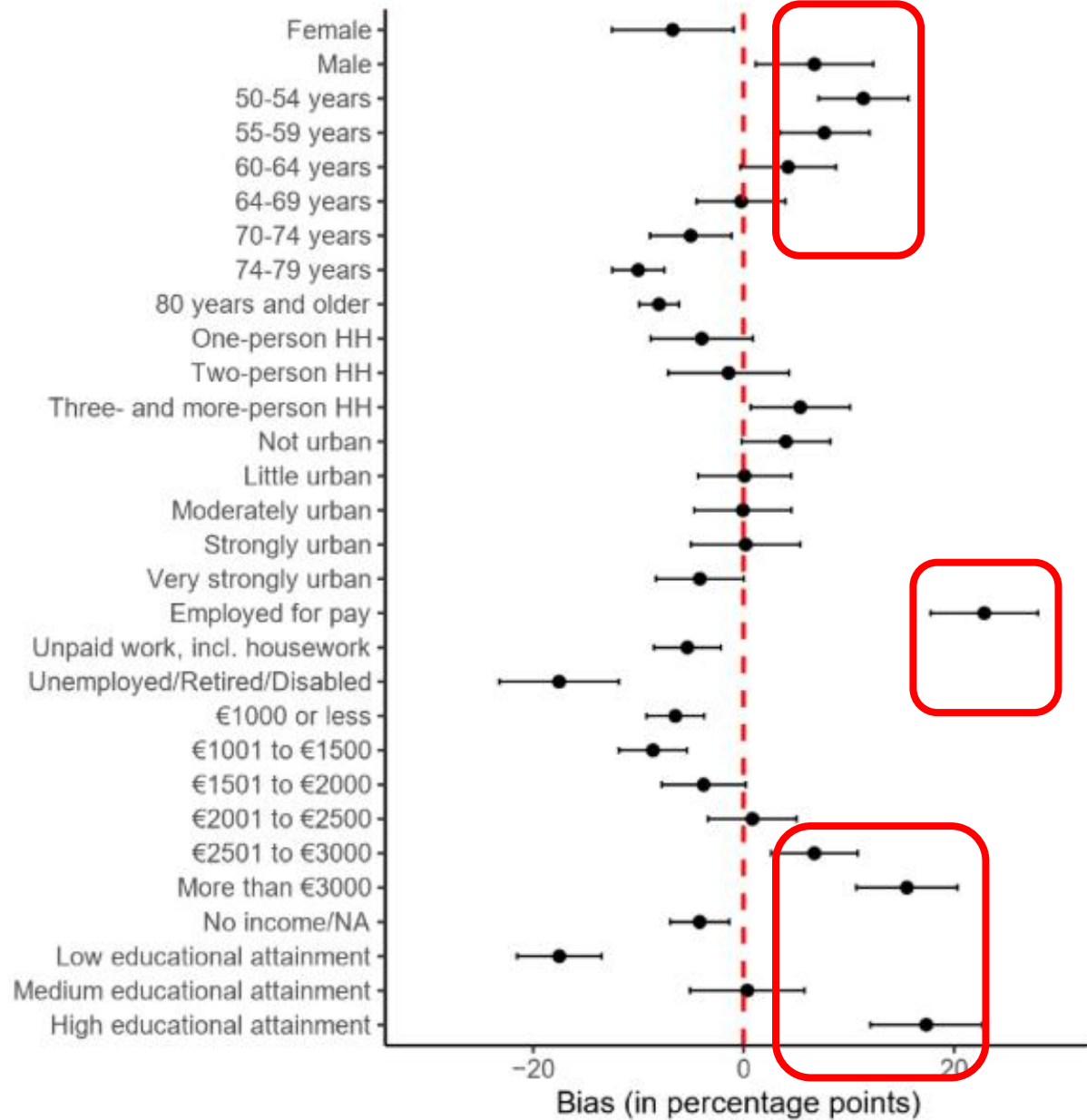


Figure 3. Bias in sociodemographic measures between donors and the full sample. Dots indicate bias point estimators, and bars indicate 95% CIs. HH: household; NA: not available.

Study 3: Algorithmic Society (AlgoSoc) UvA

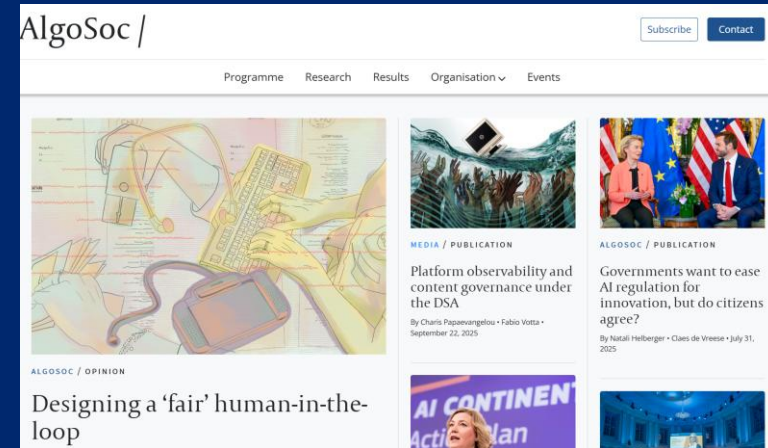
- LISS participants donated DDPs:



- Public **perceptions** and **experiences** with artificial intelligence (**AI**) and automated decision making (**ADM**) systems in **society**
- 2024 – 2026: 6 survey waves, 2 data donation waves ('24 & '26)



Source: <https://monitor.algosoc.org/en/>

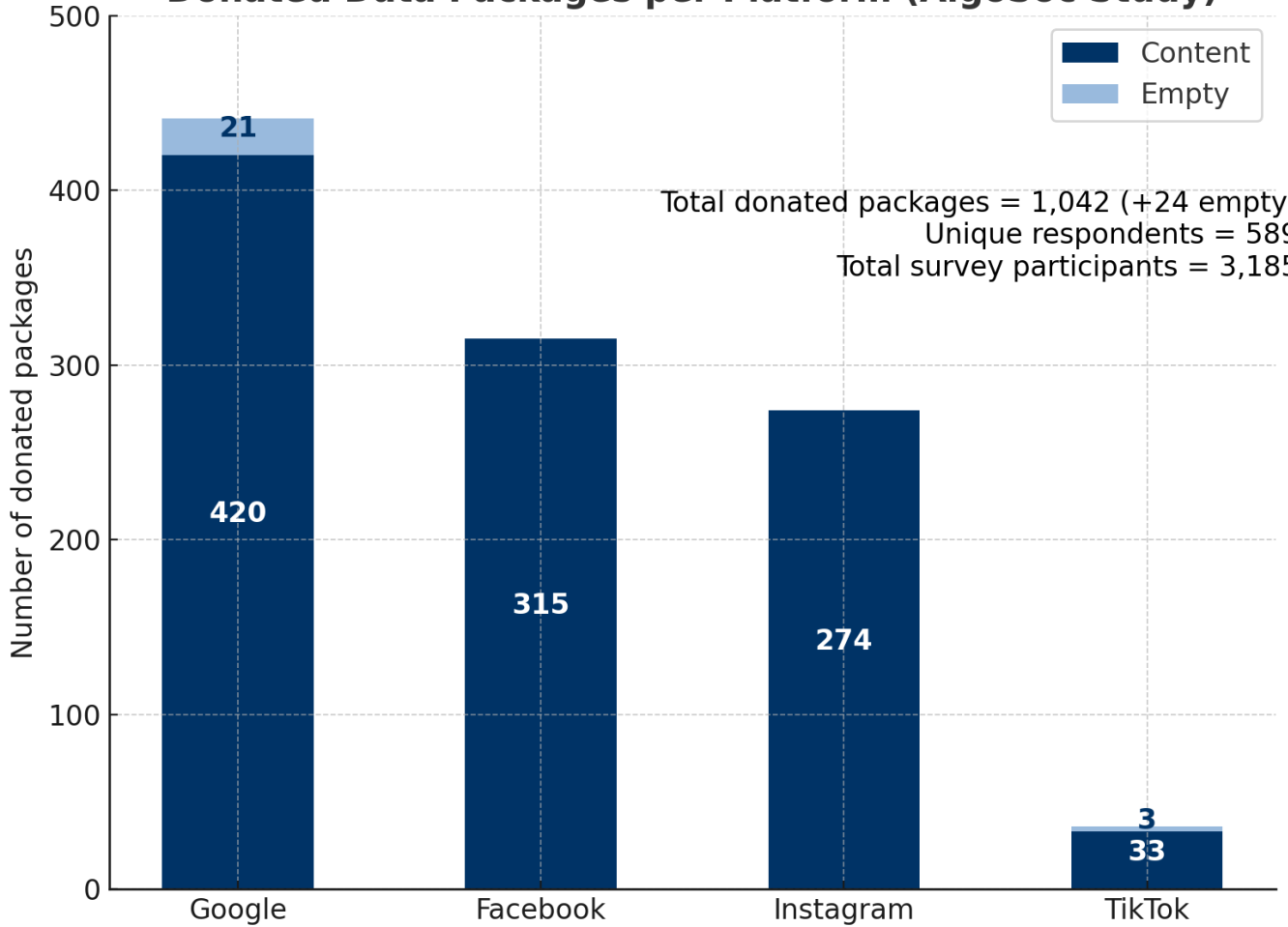


AlgoSoc.org

Study 3: Algorithmic Society (AlgoSoc) UvA



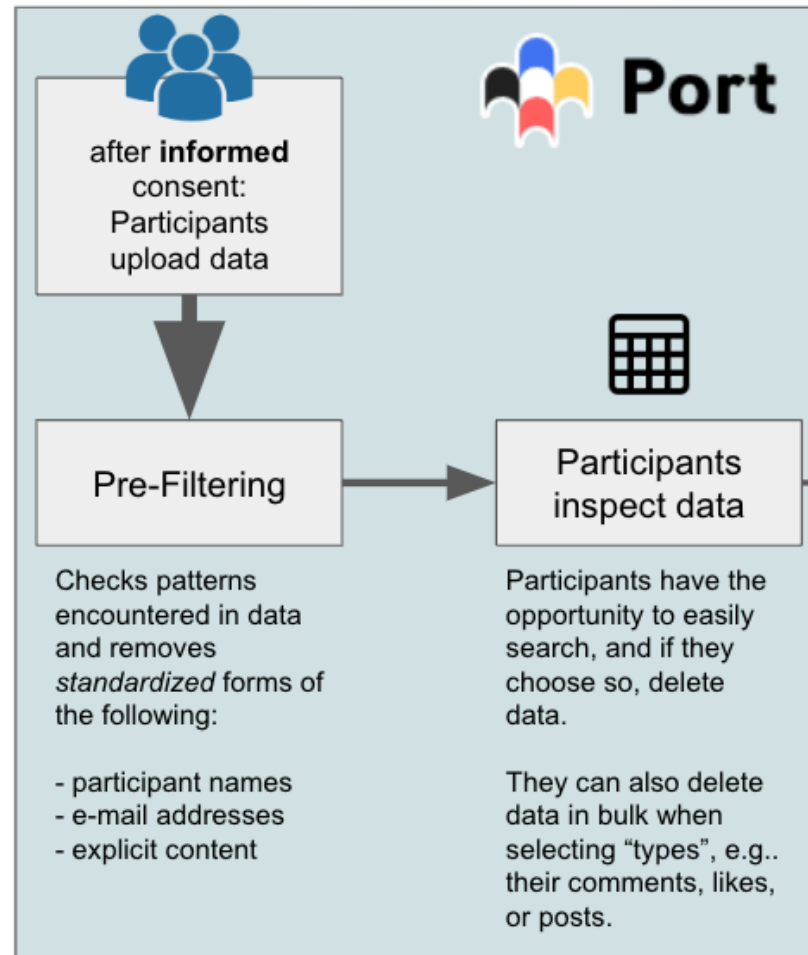
Donated Data Packages per Platform (AlgoSoc Study)



Study 3: Algorithmic Society (AlgoSoc) UvA



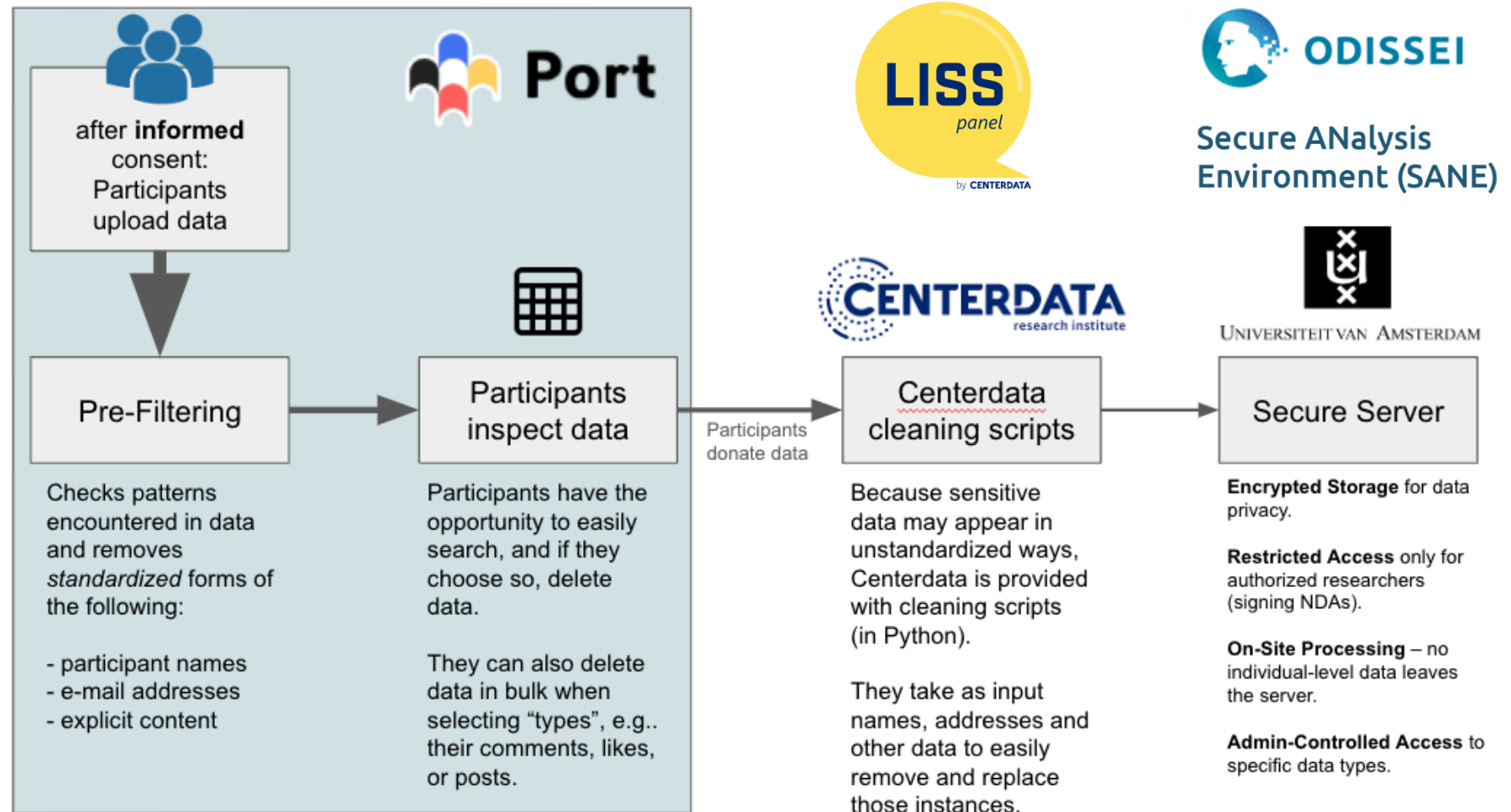
Additional privacy measures



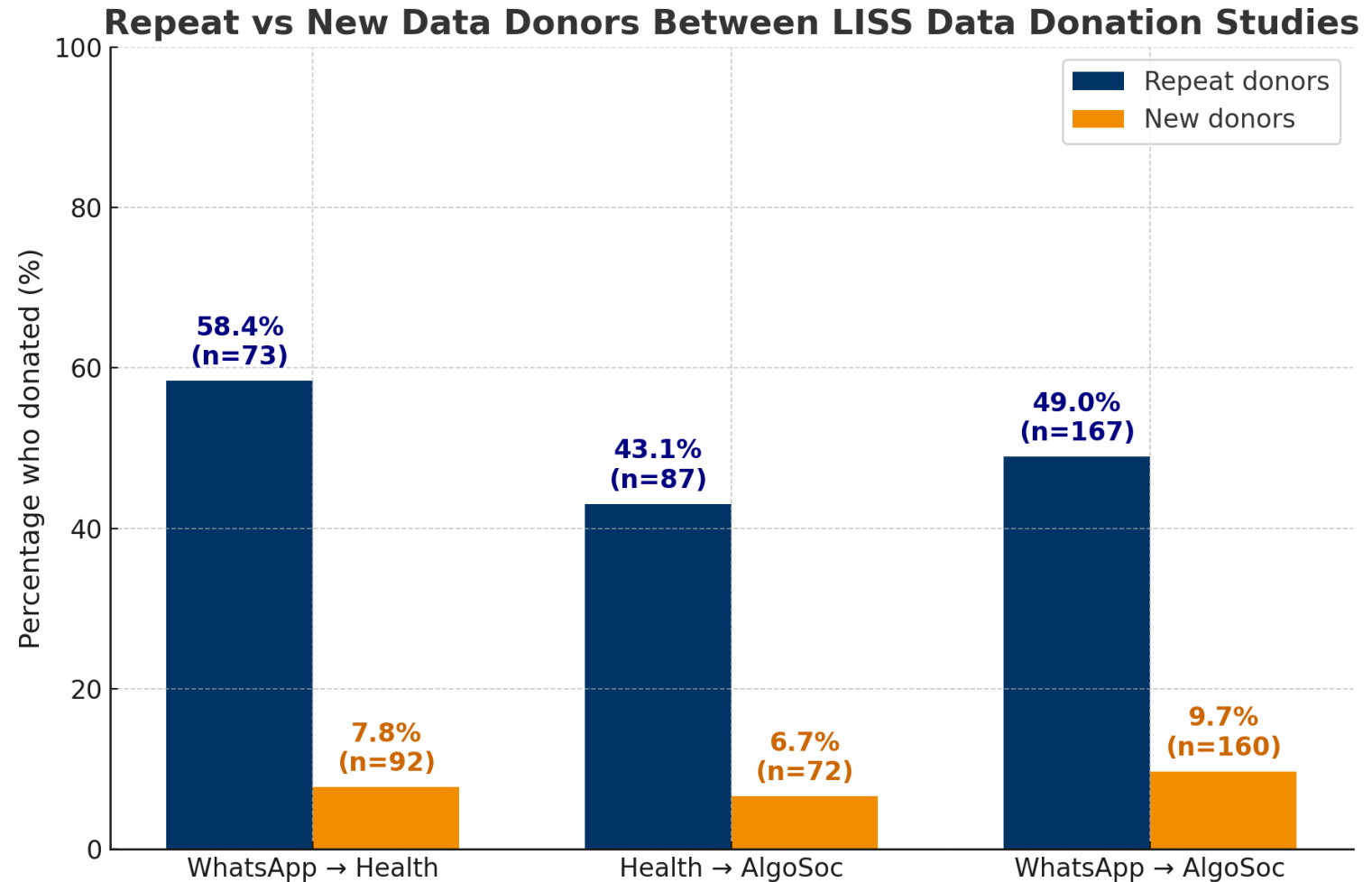
Study 3: Algorithmic Society (AlgoSoc) UvA



Additional privacy measures



Repeating and new donors over the three studies



Proportion of LISS panel members who **donated data again (blue)** or **for the first time (orange)** in subsequent data donation studies. **Repeat donations** vary between 40 - 60%. **New donations** are about 6–10%.

Lessons learned



- Feasible for probability-based panels
- Lower participation than surveys, but still valuable
- Multi-step friction is the main barrier
- Repeat participation improves efficiency
- Whitelist vs Blacklist = strategic trade-off

Survey Data

LISS
CenterData

Waar zou u de Nederlandse media in het algemeen plaatsen op een schaal van "zeer links" tot "zeer rechts"?

Nederlandse media in het algemeen zijn:

Zeer links

Links

Enigszins links

Neutraal: niet links noch rechts

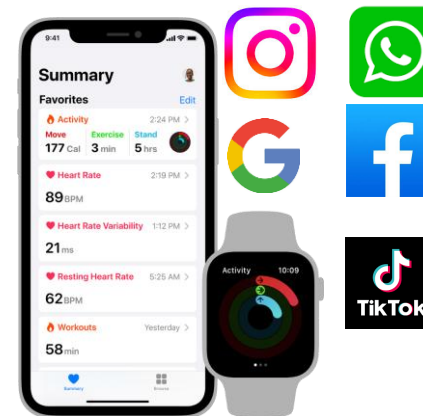
Enigszins rechts

Rechts

Zeer rechts

Vorige Verder

Digital Traces



Hybrid research designs



Questions?

More information:

Centerdata:
www.centerdata.nl

LISS data archive (non commercial use):
www.lissdata.nl



CENTERDATA

research institute

Joris Mulder
LISS quality & innovation Lead

joris.mulder@centerdata.nl