

# Digital Traces via Data Donations

## Workshop DGPuK RezFo 2026

---

Session **2**: Data Donation Studies (Participant Perspective)

---

👉 Part of the SPP DFG Project [Integrating Data Donations in Survey Infrastructure](#)



# Agenda

1. What are data donation studies?

2. Participant perspective, including

📢 **Task 1:** Inspect your DDPs

📢 **Task 2:** Try donating your data

📢 **Task 3:** Extract personality profiles from data



Image by Hope House Press via Unsplash

# 1) What are data donation studies?



Source: Image by Markus Winkler via Unsplash

# Changes in legal contexts

- EU secures right to own data in Art. 15 of the General Data Protection Regulation ([GDPR](#))
  - *“The data subject shall have [...] access to the personal data”* (§ Art. 15, 1)
  - *“The controller shall provide a copy of the personal data”* (§ Art. 15, 3)
- According to § Art. 20, users must receive their data *“in a structured, commonly used and machine-readable format”* (§ Art. 20, 1)

# Changes in legal contexts

- EU secures right to own data in Art. 15 of the General Data Protection Regulation (GDPR)
  - *“The data subject shall have [...] access to the personal data”* (§ Art. 15, 1)
  - *“The controller shall provide a copy of the personal data”* (§ Art. 15, 3)
- According to § Art. 20, users must receive their data *“in a structured, commonly used and machine-readable format”* (§ Art. 20, 1)

👉 **Solution:** Platforms offer data download packages (DDPs), which users can request and download to inspect data.

👉 **Consequence:** Researchers uses DDPs as part of user-centric data donation studies.

Please raise your hand 🙋 if you...

*1) knew about this right?*

*2) ever tried to request your data?*



In Switzerland, only 8% of members of an online access panel did (Pfiffner & Friemel, 2023, p. 17)

# What are data donation studies?

**Definition** 💡 : Data donation studies are a **user-centric, privacy-by-design** method for collecting digital traces:

- Users have the right to request, access, and download data that platforms collect about them.
- They can make their *data packages (DDPs)* available to science via *data donation tools (DDTs)*, often in the context of surveys.
- Researchers use CSS methods to filter, anonymize, and aggregate this data **locally** on participants' devices.
- Participants can inspect/delete their data before any data is transferred.

# Which types of data do DDPs contain?

For platforms like Facebook, Instagram, or YouTube, for example... [\(Hase et al., 2024\)](#)

- *User profiles* (e.g., privacy settings)
- *Activities* (e.g., friends, likes, searches, exposure, analog movements)
- *Content and context* (e.g., ads watched, algorithmically inferred interests)



# How is data from DDPs different?

Compared to **APIs** ([Haim & Hase, 2023](#); [Ohme et al., 2024](#))...

- Control & informed consent of users
- Longitudinal data without “rate limits”
- Partly more knowledge about missing data (e.g., unit drop-out)
- Partly additional measurements (e.g., exposure data; non-public data)

# How is data from DDPs different?

Compared to **tracking** ([Haim & Hase, 2023](#); [Ohme et al., 2024](#))...

- Retrospective & often more longitudinal
- Partly additional measurements (e.g., algorithmic inferences)
- Anchored in current provisions in law (for issues with tracking, see [Edelson & McCoy, 2021](#))

👉 but similarly burdensome for participants!

# Summary: What are data donation studies?



- **Summary:** User-centric method for collecting digital traces: Users download their data packages (DDPs) from platforms and make them available to science via data donation tools (DDTs).
- **Further literature:**
  - Boeschoten et al. (2022)
  - Carrière et al. (2024)

## 2) Participant perspective



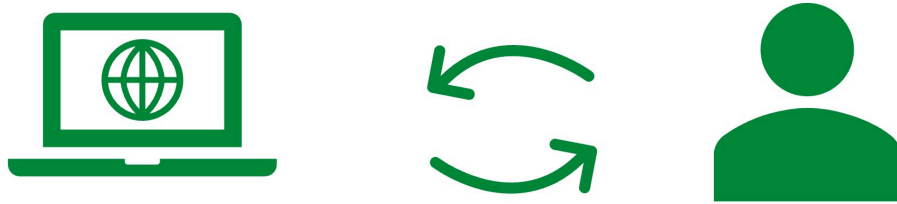
Source: Image by Markus Winkler via Unsplash

# Data donation studies - participant perspective



Figure. Data donation study - participant perspective

# Step I: Data request



- 1** Request  
Users request their data from a platform

Figure. Data donation study - participant perspective

# Step I: Data request

CREATE A NEW EXPORT


1

Select data to include

1 of 1 selected


Products


Deselect all



YouTube and YouTube Music

Watch and search history, videos, comments and other content you've created on YouTube and YouTube Music [More info](#)



 Multiple formats

Show more products

Next step

Figure. Data donation study - participant perspective

# Step 1: Data request

2

Choose file type, frequency & destination

Destination

Transfer to:

Send download link via email

When your files are ready, you'll get an email with a download link. You'll have one week to download your files. [Learn more](#) about how to locate, access, and share your data.

Frequency

☒ Export once

1 export

☐ Export every 2 months for 1 year

6 exports

File type & size

File type:

.zip

Zip files can be opened on almost any computer.

File size:

2 GB

Exports larger than this size will be split into multiple files.

Figure. Data donation study - participant perspective



*How easy (or hard) did you find it to request your data? Did you encounter any obstacles? 🤔*

# Step I: Data request

Example problems from Hase et al. (2025)

# Step I: Data request

Different degrees in standardization for data requests ([Hase et al., 2024](#))...

- Verification procedure
- Specification of data (metrics, observation period)
- Notification on provision of DDP
- Duration of DDP availability

# Step II: Download



Figure. Data donation study - participant perspective

```
[{"header": "YouTube", "title": "Watched Avicii, Dua Lipa, Coldplay, Martin Garrix \u0026 Kygo, The Chainsmokers Style - Summer Vibes #145", "titleUrl": "https://www.youtube.com/watch?v\u003dUrNm7c9eiu4", "subtitles": [{"name": "Ru Deep Radio", "url": "https://www.youtube.com/channel/UCopEUD22VBdNd80muYA7sYA"}], "time": "2026-01-11T09:34:55.239Z", "products": ["YouTube"], "activityControls": ["YouTube watch history"]}, {"header": "YouTube", "title": "Watched hw_dungeoncrypt6o_h_en_110", "titleUrl": "https://www.youtube.com/watch?v\u003dnMo_AFTBFpA", "description": "Watched at 10:34 AM", "time": "2026-01-11T09:34:49.418Z", "products": ["YouTube"], "details": [{"name": "From Google Ads"}]}, {"header": "YouTube", "title": "Viewed Ads On YouTube Homepage", "time": "2026-01-11T09:34:45.242Z", "products": ["YouTube"], "details": [{"name": "From Google Ads"}]}, {"header": "YouTube", "title": "Watched Avicii, Dua Lipa, Coldplay, Martin Garrix \u0026 Kygo, The Chainsmokers Style - Summer Vibes #145", "titleUrl": "https://www.youtube.com/watch?v\u003dUrNm7c9eiu4", "subtitles": [{"name": "Ru Deep Radio", "url": "https://www.youtube.com/channel/UCopEUD22VBdNd80muYA7sYA"}], "time": "2026-01-11T09:34:55.239Z", "products": ["YouTube"], "activityControls": ["YouTube watch history"]}], [{"name": "search-history.json"}, {"name": "watch-history.json"}]]
```

## Task 1: Inspect your DDPs.

1. *Which variables can you identify?*
2. *How could you use these for research in your field?*
3. *What are questions you may have about these measurements?*

Feel free to work in groups of 2-3 people (around 15 minutes)

# Step II: Download

Different degrees in standardization for DDP content ([Hase et al., 2024](#))...

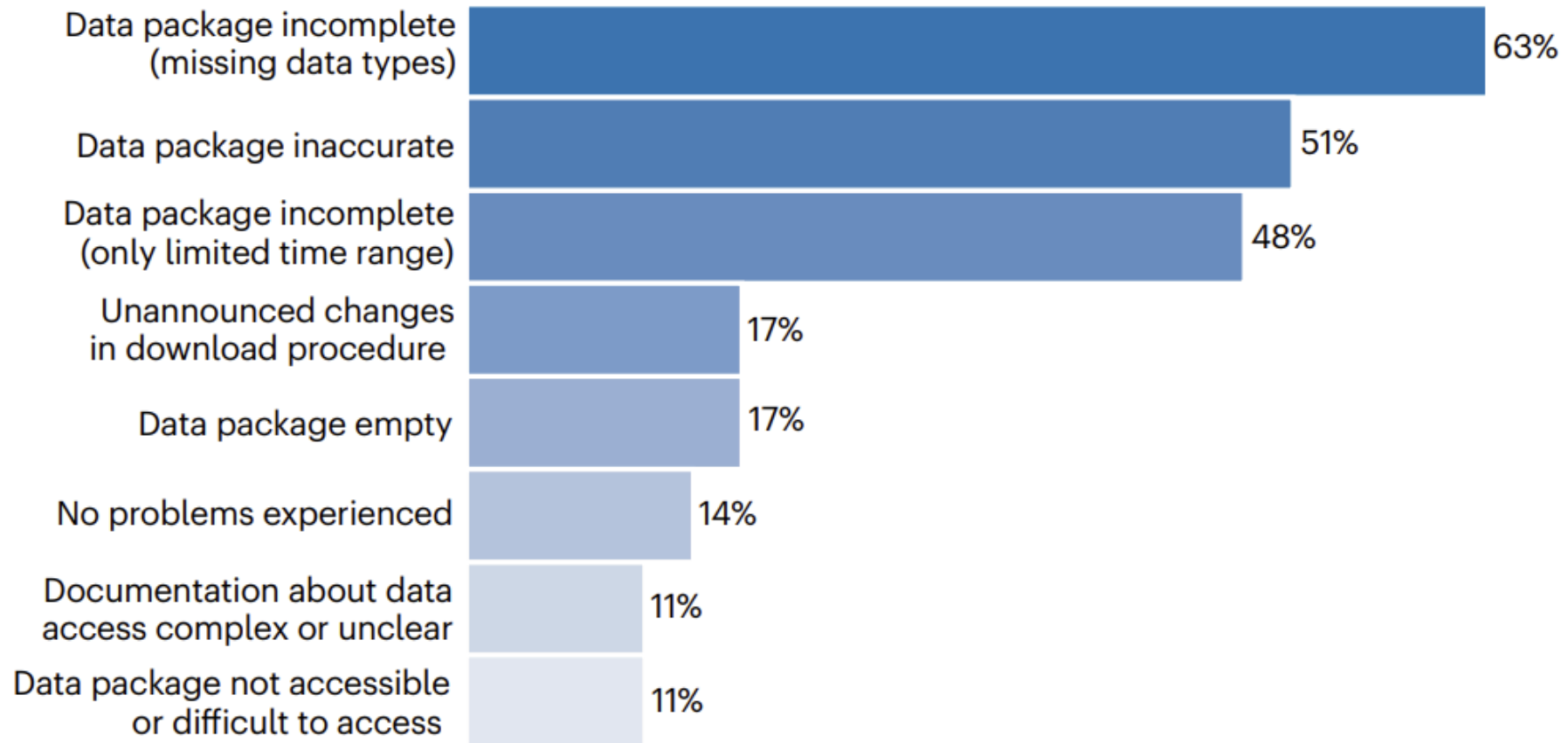
## Documentation & changes over time

- What does each file “mean”?
- How was this data measured?

## Completeness & scope

- Is data missing and, if so, why?
- Why are some data points limited to specific time frames?
- How do we handle multi-lingual DDPs?

# Step II: Download



**Fig. 1 | Common problems in platform data donations experienced by researchers.** Data from a [June 2024 survey](#) among 51 data donation researchers.

Valkenburg et al. (2025). It is time to ensure research access to platform data



# Step III: Data donation



Figure. Data donation study - participant perspective

# Step III: Data donation



## Donate your data

You are about to go through a few steps to donate your data. Before donating, you can view your data and then give your consent to share it.  
The page may take a moment to load — thank you for your patience.

Continue

Figure. Data donation study - participant perspective

# Step III: Data donation

## YouTube Data Donation

Please select your downloaded YouTube ZIP file.

E.g. data.zip

Choose file

Note: The process to extract the correct data from the file is done on your own device. No data is stored or sent yet.

Continue

Figure. Data donation study - participant perspective

# Step III: Data donation

## 1 Watch History

< 1 2 3 4 5 6 7 >

19 pages

Timestamp	Video Title	Video ID	Channel Name	Channel ID	Activity Controls
2026-01-11 09:34:55.239	Avicii, Dua Lipa, Coldplay, Martin Garrix & Kygo, The Chainsmokers Style - Summer Vibes #145	UrNm7c9eiu4	Ru Deep Radio	UCopEUD22VBdNd80muYA7sYA	YouTube watch history
2026-01-11 09:34:49.418	hw_dungeoncrypt6o_h_en_110	nMo_AFTBFpA			Web & App Activity, YouTube watch history, YouTube search history
2026-01-11 09:05:45.500	Tom Odell - Another Love (Lyrics)	Jkj36B1YuDU	7clouds	UCNqFDjYTexJDET3rPDrmJKg	YouTube watch history
2026-01-11 09:05:38.861	Bring your ideas to life with Base44	OugxD0argKI			Web & App Activity, YouTube watch history, YouTube search history
2026-01-11 09:05:32.534	Einmal satt auf die Hand: BIG Toppings. BIG Taste. BIG Sensation. Von Wagner.	Or47Afc4Xqs			Web & App Activity, YouTube watch history, YouTube search history
2026-01-11 09:05:10.530	Mega Hits 2025 🍷 The Best Of Vocal Deep House Music Mix 2025 🍷 Summer Music Mix 2025 #4	N6DW31S_oYl	Helios Deep	UCw527CsYYIcRVR3W2ywMVOA	YouTube watch history

Figure. Data donation study - participant perspective

## Task 2: Try donating your data.

*Upload and inspect your data. If you could not download the data, use the ZIP folder “example data” (see landing page) for this session.*

1. Are there any surprising/missing values?
2. Did you encounter any difficulties in deleting/sharing your data?

Use this link: <https://next.eyra.co/a/nWPJC4?p=999> - replace number after  $p=$  with random number. Feel free to work in groups of 2-3 people.

### Task 3: Extract personality profiles from data.

*Inspect the ZIP folder “example” data for this session (manually or upload it to the data donation study - data you see may then differ). Imagine you are a commercial company.*


*Based on the trace data, how would you profile this person in terms of...*

1. sociodemographic characteristics
2. their interests
3. ads you may target them with

Use this link: <https://next.eyra.co/a/nWPJC4?p=999> - replace number after  $p=$  with random number. Feel free to work in groups of 2-3 people.

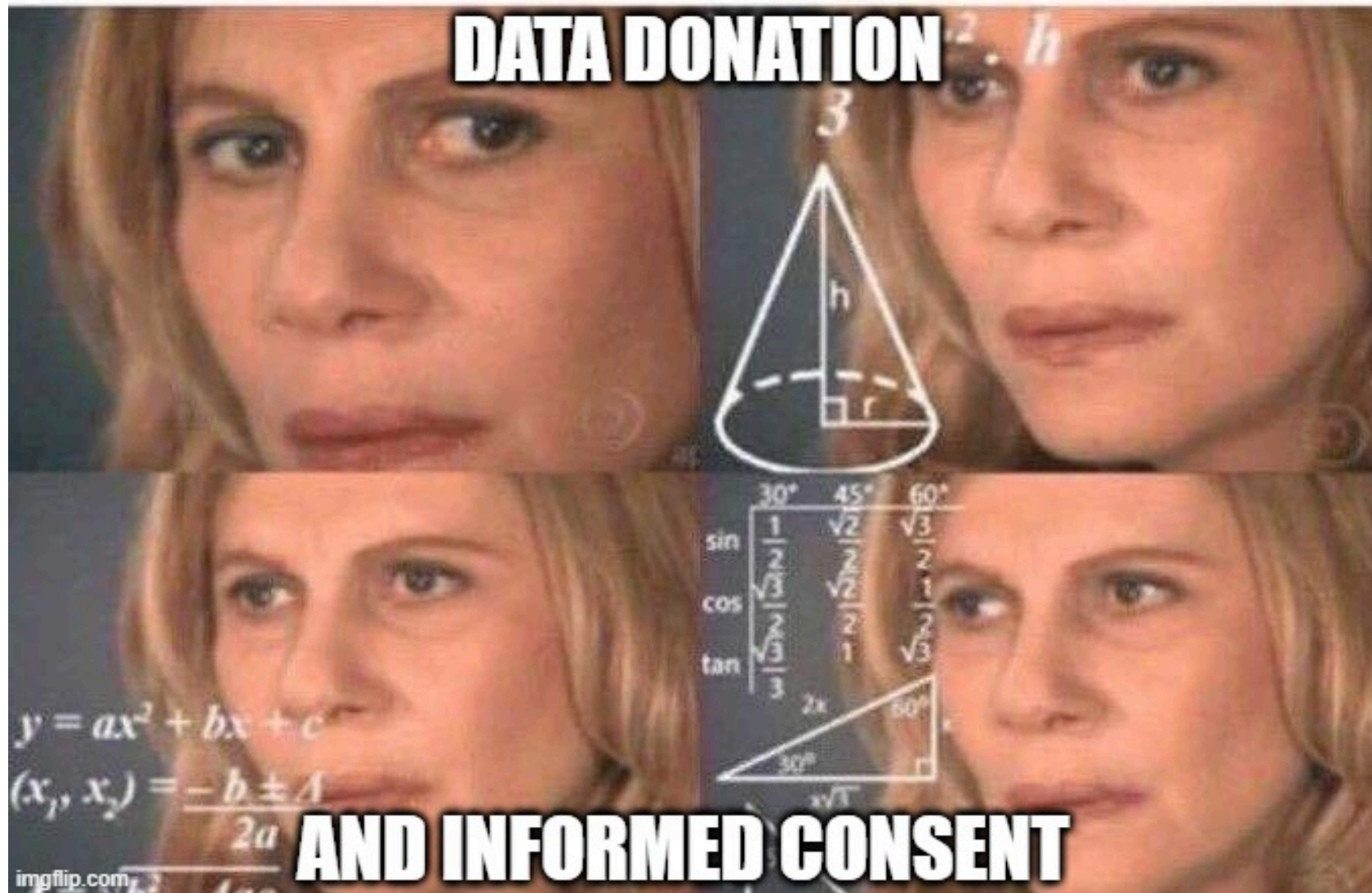
# Bias in Data Donation Studies

Similar to tracking studies, data donation characterized by (Hase et al., 2024; Hase & Haim, 2024; Keusch et al., 2024):

- Errors in representation (e.g., coverage error, non-response error)
  - Measurement errors (e.g., missing data, specification error)
- 👉 Not necessarily less biased than survey data - but subject to different bias?
- 👉 We will talk about this in Session  .

# Informed Consent in Data Donation Studies

Similar to tracking (Kreuter et al., 2020): Is consent really “informed”? (Groot Kormelink et al., 2025; Hase & Haim, 2024)





# Informed Consent in Data Donation Studies

Similar to tracking (Kreuter et al., 2020): Is consent really “informed”? (Groot Kormelink et al., 2025; Hase & Haim, 2024)

- in walk-through interviews, participants “*fundamentally misunderstood [this] key principle*” (p. 8)
- Anecdotal evidence: sending data via email, sending different data, sending someone else’s data, etc.
- Almost no one deletes data (except for searches!)

# Summary: Participant perspective

- **Summary:** Key steps from the participant perspective ...

1. Data request
2. Data download
3. Data donation

- **Further literature:**

- Boeschoten et al. ([2022](#))
- Carrière et al. ([2024](#))

# Questions? 🤔

# References

- Boeschoten, L., Mendrik, A., Van Der Veen, E., Vloothuis, J., Hu, H., Voorvaart, R., & Oberski, D. L. (2022). Privacy-preserving local analysis of digital trace data: A proof-of-concept. *Patterns*, 3(3), 100444.  
<https://doi.org/10.1016/j.patter.2022.100444>
- Carrière, T. C., Boeschoten, L., Struminskaya, B., Janssen, H. L., De Schipper, N. C., & Araujo, T. (2024). Best practices for studies using digital data donation. *Quality & Quantity*. <https://doi.org/10.1007/s11135-024-01983-x>
- Edelson, L., & McCoy, D. (2021). Facebook is obstructing our work on disinformation. Other researchers could be next. *The Guardian*. <https://www.theguardian.com/technology/2021/aug/14/facebook-research-disinformation-politics>
- Groot Kormelink, T., Houwing, F., Struminskaya, B., Boeschoten, L., Schipper, N. de, & Welbers, K. (2025). Meaningful informed consent? How participants experience and understand data donation. *Information, Communication & Society*, 0(0), 1–18. <https://doi.org/10.1080/1369118X.2025.2540915>
- Haim, M., & Hase, V. (2023). Computational Methods und Tools für die Erhebung und Auswertung von Social-Media-Daten. In S. Stollfuß, L. Niebling, & F. Raczkowski (Eds.), *Handbuch Digitale Medien und Methoden* (pp. 1–20). Springer Fachmedien Wiesbaden. [https://link.springer.com/10.1007/978-3-658-36629-2\\_41-1](https://link.springer.com/10.1007/978-3-658-36629-2_41-1)
- Hase, V., Ausloos, J., Boeschoten, L., Pffner, N., Janssen, H., Araujo, T., Carrière, T., De Vreese, C., Haßler, J., Loecherbach, F., Kmetty, Z., Möller, J., Ohme, J., Schmidbauer, E., Struminskaya, B., Trilling, D., Welbers, K., &

- Haim, M. (2024). Fulfilling Data Access Obligations: How Could (and Should) Platforms Facilitate Data Donation Studies? *Internet Policy Review*, 13(3). <https://doi.org/10.14763/2024.3.1793>
- Hase, V., & Haim, M. (2024). Can We Get Rid of Bias? Mitigating Systematic Error in Data Donation Studies through Survey Design Strategies. *Computational Communication Research*, 6(2), 1. <https://doi.org/10.5117/CCR2024.2.2.HASE>
- Keusch, F., Pankowska, P. K., Cernat, A., & Bach, R. L. (2024). Do You Have Two Minutes to Talk about Your Data? Willingness to Participate and Nonparticipation Bias in Facebook Data Donation. *Field Methods*, 36(4), 279–293. <https://doi.org/10.1177/1525822X231225907>
- Kreuter, F., Haas, G.-C., Keusch, F., Bähr, S., & Trappmann, M. (2020). Collecting Survey and Smartphone Sensor Data With an App: Opportunities and Challenges Around Privacy and Informed Consent. *Social Science Computer Review*, 38(5), 533–549. <https://doi.org/10.1177/0894439318816389>
- Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital Trace Data Collection for Social Media Effects Research: APIs, Data Donation, and (Screen) Tracking. *Communication Methods and Measures*, 18(2), 124–141. <https://doi.org/10.1080/19312458.2023.2181319>
- Pfiffner, N., & Friemel, Thomas. N. (2023). Leveraging Data Donations for Communication Research: Exploring Drivers Behind the Willingness to Donate. *Communication Methods and Measures*, 17(3), 227–249. <https://doi.org/10.1080/19312458.2023.2176474>