Follow the User?!

Data Donation Studies for Collecting Digital Trace Data

Session 2 : Data Donation Studies (Participant Perspective)

Frieder Rodewald (University of Mannheim) & Valerie Hase (LMU Munich)

👉 Part of the SPP DFG Project Integrating Data Donations in Survey Infrastructure

Agenda

1. **What are data donation studies?**

2. **Participant perspective**, including

    2.1 📢 **Task 1**: Inspect your DDPs

    2.2 📢 **Task 2**: Try donating your 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 ✋ : Who has ever tried to request their data from an online platform? 🤔

What are data donation studies?

Definition 💡: Data donation studies are a user-centric 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 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)

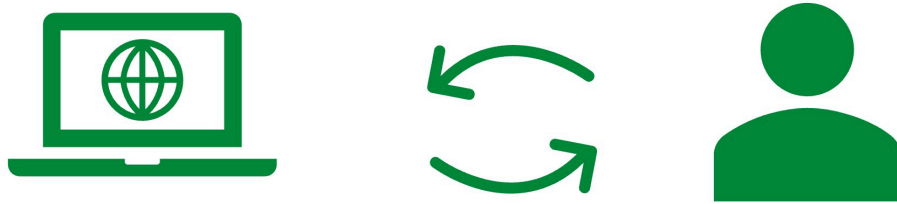 - Carrière et al. (2024)

2) Participant perspective

Source: Image by Markus Winkler via Unsplash

Data donation studies – participant perspective

1 Request
Users request their data from a platform

2 Download
Users download their data once it becomes available

3 Data donation
Users donate their data to research via data donation tools

Figure. Data donation study - participant perspective

Step 1: Data request

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

Figure. Data donation study - participant perspective

Step 1: Data request

Figure. Data donation study – participant perspective

Step 1: Data request

Figure. Data donation study - participant perspective

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

Step 1: 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

1 <u>Request</u>
Users request their data from a platform

2 <u>Download</u>
Users download their data once it becomes available

Figure. Data donation study - participant perspective
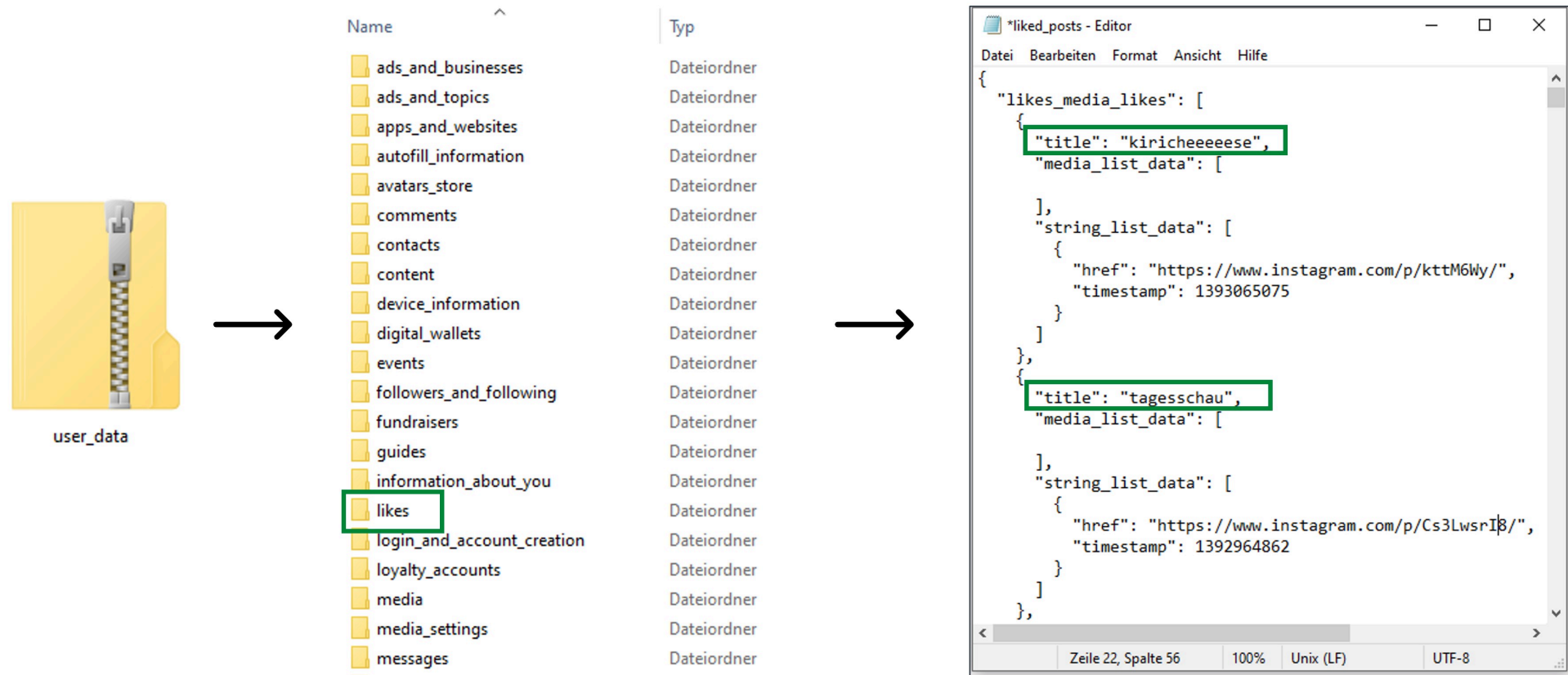
Step II: Download

Figure. Data donation study - participant perspective

📢 Task 1: Inspect your DDPs.

Which variables can you identify? How could you use these for research in your field?

Feel free to work in groups of 2-3 people.

Step II: Download

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

- Documentation

  - DDP structure?

  - Measurements?

- Completeness & scope

  - Missing data?

  - Limited time frames?

Step III: Data donation

1 <u>Request</u>
Users request their data from a platform

2 <u>Download</u>
Users download their data once it becomes available

3 <u>Data donation</u>
Users donate their data to research via data donation tools

Figure. Data donation study - participant perspective

Step III: Data donation

Figure. Data donation study - participant perspective

Step III: Data donation

Ihre YouTube Datenspende

Legen Sie fest, ob Sie die untenstehenden Daten spenden möchten. Überprüfen Sie die Daten sorgfältig und passen Sie sie bei Bedarf an. Mit Ihrer Spende tragen Sie zur zuvor beschriebenen Forschung bei. Vielen Dank im Voraus.

0 Welche Kanäle haben Sie abonniert?

1 Seite   Suchen

Abonnierter Kanal

DER SPIEGEL

Cafe Music BGM channel

ZEIT ONLINE

Figure. Data donation study – participant perspective

Step III: Data donation

Figure. Data donation study - participant perspective

📢 Task 2: Try donating your data.

Upload and inspect your data. Did you encounter any difficulties in sharing your data?

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 4 .

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

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

Hase, V., Ausloos, J., Boeschoten, L., Pfiffner, 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

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