

# *Introduction to Data Donation for Communication Research*

**Workshop at the European Communication Conference (ECC)**

26 September 2024, Ljubljana, Slovenia

## *Syllabus*

Introduction.....	2
Workshop programme .....	2
Registration.....	3
The Data Donation Module (DDM) .....	3
Presenters.....	7
Suggested reading.....	7

**Data Donation  
Lab** 



**University of  
Zurich**<sup>UZH</sup>

Digital Society Initiative



## ***Introduction***

The collection of data donations to access digital trace data for academic research has gained importance due to the increasing limitations of public APIs and web-scraping (Bruns, 2019; Freelon, 2018; Mancosu & Vegetti, 2020), and the introduction of the general data protection regulation (GDPR) in 2018 in the EU (Ausloos & Veale, 2021). To collect data donations in an academic context means that researchers approach users of a relevant platform and ask them to request and contribute their personal data to a research project. One way to do this is the *download-upload-approach* which utilizes the *takeout function* that most large companies such as Google, Instagram, or Spotify have integrated in their platforms. This function allows users to download a copy of the data the respective platform has stored about them (sometimes described as *data download package*; Boeschoten, Ausloos, et al., 2022) and subsequently submit (upload) these data to researchers. Currently, this approach still lacks best practices (Breuer et al., 2022; Ohme & Araujo, 2022) and few tools are available that researchers can adapt to implement a data donation collection based on this strategy (e.g., Araujo et al., 2022; Boeschoten, Mendrik, et al., 2022). The Data Donation Module (DDM) presented in the workshop addresses this gap and provides a solution to collect data donations following the download-upload-approach.

## ***Workshop Programme***

The workshop consists of three parts:

### **1. Introduction to data donation**

We will provide a brief overview on what data donation is, why it is a promising new approach for communication researchers, and what needs to be considered when designing a data donation project.

### **2. Introduction to the Data Donation Module**

We will present the Data Donation Module, an open-source web application that can be used to conduct data donation projects on all digital (media) platforms and other digital traces. The introduction includes a walkthrough of setting up a project, including crucial design decisions, recruitment of participants, monitoring of a project, and data analysis.

### **3. Best Practices**

We provide insights into various data donation projects and share the most important dos and don'ts that need to be considered from ethical approval to publishing your results.

## ***Registration***

Participation is free of charge. For organizational reasons we ask interested participants to register for the workshop. Due to the room capacity the number of participants is limited, and seats will be given based on a first come first serve basis. Walk-in participants are welcome as long as space is available.

Please register here: <https://forms.office.com/e/NXQB7NWEte>

**Date and time:** 26 September, 16:30 – 18:00

**Place:** The workshop will be held in the lecture room **FDV-20** at the conference site of ECC.

## ***The Data Donation Module (DDM)***

### **Design Principles**

The DDM provides a *secure, transparent* and *flexible* solution to collect data donations for academic research projects. DDM is developed open-source at the University of Zurich and the code is accessible on GitHub (<https://github.com/uzh/ddm>). The design of the DDM follows a particular set of principles:

***Institutional Deployment:*** DDM is designed to be used and deployed at an institutional level (e.g., at department- or university-level): one installation of DDM can be used by multiple researchers to conduct their data donation projects. This makes it possible to ensure data security as well as to define legal responsibilities. At the same time, the user management guarantees that the data of different projects are separated and only accessible by the respective researcher.

In addition to this institutional deployment, the DDM is also available as a service that can be used for individual projects.

***Usability:*** The application offers a graphical user interface (GUI) for both participants and researchers. Through this GUI, researchers can fully configure their data donation project without extensive programming skills. In addition, features such as an exception log and participation

statistics allow researchers to monitor their ongoing data donation collections through the same interface.

**Data Enrichment:** DDM enables the enrichment of data donations with additional data in three ways: (1) Information from data donations can be directly referenced in a questionnaire that is integrated in the application; (2) Participants can donate data from different platforms through the same interface; (3) Integration in standard surveys (“regular” surveys can link to a data donation project which can redirect back to the initial survey).

### **Core Functionality**

The core functionality of DDM can be described from two perspectives: First, the perspective of the participant, i.e., the data donor, and second, the perspective of the researcher who collects data donations for a research project.

#### ***Participant Perspective: Participation Flow***

From the participant’s perspective, donating data through DDM is a four-step process:

On the **briefing page** participants are briefed about the research project and the upcoming steps involved in the data donation.

In the subsequent **data donation** step, participants are presented with instructions on how to request and download their data from a digital platform of interest. On the same page, participants can upload these data into the application. During the upload, the data are filtered according to filter rules defined by the researchers. The filtered data are then displayed to the participants, based on which they can make an informed decision whether to donate their data. Up to this point, all data processing steps take place on the client side in the participant’s browser. If explicit consent is given to donate the data, the filtered data is submitted through an end-to-end encrypted connection to the server on which the DDM instance is hosted. Once received by the server, the data are encrypted and saved to a database.

Third, participants are presented with a **questionnaire**. This step is optional and is displayed if a questionnaire has been defined.

Lastly, participants are presented with a **debriefing page**.

#### ***Researcher Perspective: Data Donation Project Setup***

From the researcher’s perspective, the setup of the data donation collection directly relates to the steps of the participation flow. All settings can be defined in the application through a GUI.

For the **briefing page**, researchers can define how they will brief participants.

The **data donation setup** is built around so-called *File Blueprints* and *Uploaders*. A File Blueprint relates to a single file that will be extracted from a data donation and defines the data validation and data extraction steps for this file. If participants are expected to upload a zip-container instead of a single file, multiple File Blueprints can be bundled into an Uploader and the files within the zip-container will only be extracted if they match a File Blueprint. Furthermore, researchers can define the instructions for participants to guide them through the data donation.

In the **questionnaire setup**, DDM allows to include data points contained in a data donation in the question or item text. For example, if YouTube watch histories are collected, the last watched video can be included in the question text. This feature can also be used to provide participants with feedback on their own media usage (e.g., in the form of numbers or graphs).

Lastly, the content of the **debriefing page** can be freely configured, and researchers may set up a redirection link to, for example, an external survey.

## Conclusion

Compared to existing tools, DDM adopts a similar design pattern (Araujo et al., 2022; Boeschoten, Mendrik, et al., 2022) but differs from these solutions in three important points:

First, DDM is explicitly targeted at the deployment at an institutional level. This can open the collection of data donations to groups of researchers who so far did not have the (financial or knowledge) resources to conduct data donation studies. If collecting data donations proves to be a sustainable and future-proof data collection approach, the hosting of infrastructures that enable data donation collections must be delegated to an institutional level in order to use financial resources effectively and ensure that cybersecurity measures protecting the servers and databases on which citizens' data donations are stored and processed are up-to-date. This burden should not be placed on every communication scholar that is interested in collecting data donations, just as not every research project hosts its own online survey application.

Second, data donation projects in DDM can be fully configured through a graphical user interface, making it accessible to researchers who lack programming skills. Although more and more communication scholars are using programming languages such as Python or Javascript, the majority still does not have these skills. Hence, to make the collection of data donations more accessible for researchers, it is important to offer technical solutions that can be used by everyone and do not rely on advanced IT-skills. This is also important for the broader establishment and validation of the collection of data donations as a methodological approach.

Third, DDM offers the possibility to reference the donated data in a questionnaire integrated within the application. This enables the enrichment of single data points and thus one of the main advantages of user-centric digital trace data collection approaches.

## References

- Araujo, T., Ausloos, J., van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., van de Velde, B., Vreese, C. de, & Welbers, K. (2022). OSD2F: An Open-Source Data Donation Framework. *Computational Communication Research*, 4(2), 372–387. <https://doi.org/10.5117/CCr2022.2.001.ArAU>
- Ausloos, J., & Veale, M. (2021). Researching with Data Rights. *Technology and Regulation*, 136–157. <https://doi.org/10.26116/techreg.2020.010>
- Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, 4(2), 388–423. <https://doi.org/10.5117/CCr2022.2.002.BoEs>
- 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 (New York, N.Y.)*, 3(3), 100444. <https://doi.org/10.1016/j.patter.2022.100444>
- Breuer, J., Kmetty, Z., Haim, M., & Stier, S. (2022). User-centric approaches for collecting Facebook data in the 'post-API age': experiences from two studies and recommendations for future research. *Information, Communication & Society*, 1–20. <https://doi.org/10.1080/1369118X.2022.2097015>
- Bruns, A. (2019). After the 'APIcalypse': social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Freelon, D. (2018). Computational Research in the Post-API Age. *Political Communication*, 35(4), 665–668. <https://doi.org/10.1080/10584609.2018.1477506>
- Mancosu, M., & Vegetti, F. (2020). What You Can Scrape and What Is Right to Scrape: A Proposal for a Tool to Collect Public Facebook Data. *Social Media + Society*, 6(3), 205630512094070. <https://doi.org/10.1177/2056305120940703>
- Ohme, J., & Araujo, T. (2022). Digital data donations: A quest for best practices. *Patterns (New York, N.Y.)*, 3(4), 100467. <https://doi.org/10.1016/j.patter.2022.100467>

## **Presenters**

**Nico Pfiffner** (PhD, 2024 U of Zurich) is a senior teaching and research associate at the Department of Communication and Media Research (IKMZ) at the University of Zurich. His current work focuses on the question of how digital trace data can be made accessible for academic research through data donations from citizens. He is interested in both the technical implementation of data donation collections as well as the associated practical and epistemological challenges and implications. He is the main developer of the Data Donation Module (DDM) and consults projects from various disciplines on the use of the tool.

**Thomas N. Friemel** (PhD, 2008 U of Zurich) is an associate professor at the Department of Communication and Media Research at the University of Zurich. His research focuses on health communication and the social context of media use and effects. He is interested in the development of new research methods and research designs to create relevant knowledge for the digital transformation of our society. This mission includes the development of technical infrastructures such as the Data Donation Module (DDM), the establishment of organizational structures such as the Data Donation Lab, and outreach activities such as the Data Donation Day.

## **Suggested reading**

If you would like to prepare for the workshop you might be interested in the following articles on data donation in general and the data donation module.

Pfiffner, N., & Friemel, T. 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>

Pfiffner, N., Witlox, P., & Friemel, T. M. (2024). Data Donation Module: A Web Application for Collecting and Enriching Data Donations. *Computational Communication Research*, 6(2). <https://journal.computationalcommunication.org/article/view/8597>

*Data Donation Workshop @ ECC 2024*  
[datadonation.uzh.ch/en/ecc-workshop2024](https://datadonation.uzh.ch/en/ecc-workshop2024)

[datadonation@ikmz.uzh.ch](mailto:datadonation@ikmz.uzh.ch)

**Data Donation  
Lab** 

