What do people do online? Using data donation to understand digital behavior.

a workshop at the SPP Junior Researcher Meeting

Frieder Rodewald

University of Mannheim & Institute for Employment Research

Sebastian Prechsl

Institute for Employment Research & LMU Munich

October 22, 2025

Our Agenda

frodew.github.io/workshop_spp_annual_meeting/

1️⃣ What is digital trace data?

2️⃣ What is data donation? - The participant’s perspective.

3️⃣ What is data donation? - The researcher’s perspective.

Who are you?

Please raise your hand 🤚 if you …

- are familiar with the term digital trace data
- have worked with APIs
- have worked with data donation
- have worked with automated content analysis
- regularly use programming languages (e.g., R, Python)

About me: Frieder Rodewald

🎓 PhD, University of Mannheim & Institute for Employment Research

🧐 Research interests: "I study what people do online."

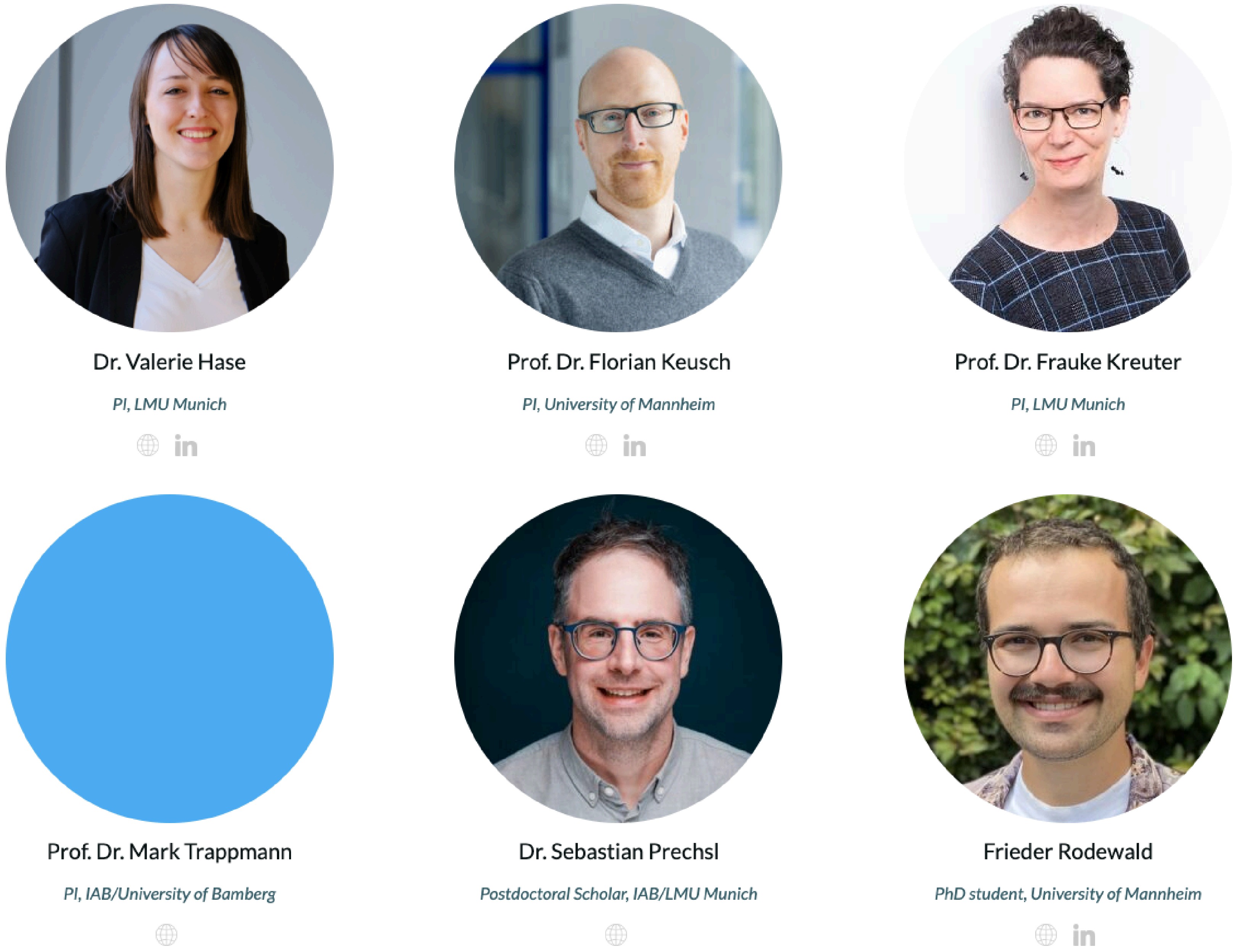
More info: github.com/frodew & frieder-rodewald.de

About me: Sebastian Prechsl

🎓 Postdoc at Institute for Employment Research & LMU

🧐 Research interests: Social inequalities (in the labor market), potentials of digital trace data for labor market research

👉 part of the SPP project Integrating Data Donations in Survey Infrastructure



Dr. Valerie Hase

PI, LMU Munich

Prof. Dr. Florian Keusch

PI, University of Mannheim

Prof. Dr. Frauke Kreuter

PI, LMU Munich

Prof. Dr. Mark Trappmann

PI, IAB/University of Bamberg

Dr. Sebastian Prechsl

Postdoctoral Scholar, IAB/LMU Munich

Frieder Rodewald

PhD student, University of Mannheim

Our Team

A huge thanks to Valerie Hase, for contributing to the conceptualizion of a previous data donation workshop at CompText in Vienna.

What is the goal of this workshop?

- ✅ Understanding digital data traces as a *type* of data
- ✅ Understanding data donation as a *method* of data access
- ✅ Working through key steps of data donation methods (participant & researcher view)
- ❓ Discussing when (not) to use data donation studies
- ❌ Detailed implementation (e.g., server set-up, coding data extraction scripts)

1️⃣ What is digital trace data?

🤔 Which examples for digital trace data you know?

What is digital trace data?

💡 **Definition:** *The recording and storing of activities on digital platforms to draw conclusions about digital and analog phenomena.*

This might include:

- Tweets, likes, shares on social media
- Geo data (locations, movements)
- Digital payments
- Spotify playlists

Where can we find/collect digital trace data?

- Apps (e.g., running apps)
- Social media platforms (e.g., Instagram)
- Payment systems (e.g., Paypal)
- Wearable devices (e.g., smart watch)

Which (latent) constructs can we measure?

- **Internet use** (Parry et al. 2021) related to …
  - Well-being (Ohme et al. 2024) or voting (Bach et al. 2021)
- **News engagement** (Reiss 2023) related to …
  - News diversity (Jürgens and Stark 2022) or public opinion formation (Yan, Schroeder, and Stier 2022)
- **Movements** related to …
  - Mobility during pandemics (Li et al. 2021) or social networks (Sepulvado et al. 2022)

Why are digital traces becoming more popular?

- **Problems with self-reported data** (e.g., via survey)
  - Self-reported data subject to specific bias (Scharkow 2016; Parry et al. 2021)
  - Response rates in surveys are declining (Luiten, Hox, and de Leeuw 2020)
- **Availabillity**
  - Cheap (e.g., via APIs)
  - Large data sets ("big data")

⚠️ **Be careful**: These "advantages" are often claimed, but **not** empirically proven.

👉 Digital traces are neither necessarily less biased, cheaper, or larger.

(Dis-)advantages of digital trace data

- ✅ More fine-grained, often longitudinal measures due to timestamps
- ✅ Partly measurement of new variables (e.g., algorithmic inference)
- ❌ Bias due to errors in representation and measurement
- ❌ Implementation can be expensive and cumbersome
- ❌ More data does not mean better data!

How can we collect digital traces?

Platform- and user-centric methods

- **Platform-centric** (based on platform cooperation)
  - API (Jünger 2021)
  - Cooperation with platforms (Wagner 2023)
- **User-centric** (based on user cooperation and informed consent) or "follow the user" approaches (Caliandro 2024)
  - Data donation (Carrière et al. 2025)
  - Linkage (Sloan et al. 2020)
  - Sensors (Struminskaya et al. 2021)
  - Tracking (Christner et al. 2022)

🤔 Questions?

2️⃣ What is data donation?

The participant’s perspective.

Changes in legal contexts ⚖️

- EU secures right to own data in Art. 15 of the General Data Protection Regulation
  - "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 ✋

(Before a week ago…) 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 download packages (DDPs)* available to science, often in the context of web 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 YouTube, Instagram, or LinkedIn, 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** (Ohme et al. 2024)...

- Control & informed consent of users
- Longitudinal data without "rate limits"
- Partly additional measurements (e.g., exposure data; non-public data)

👉 but can be burdensome for participants!

Survey

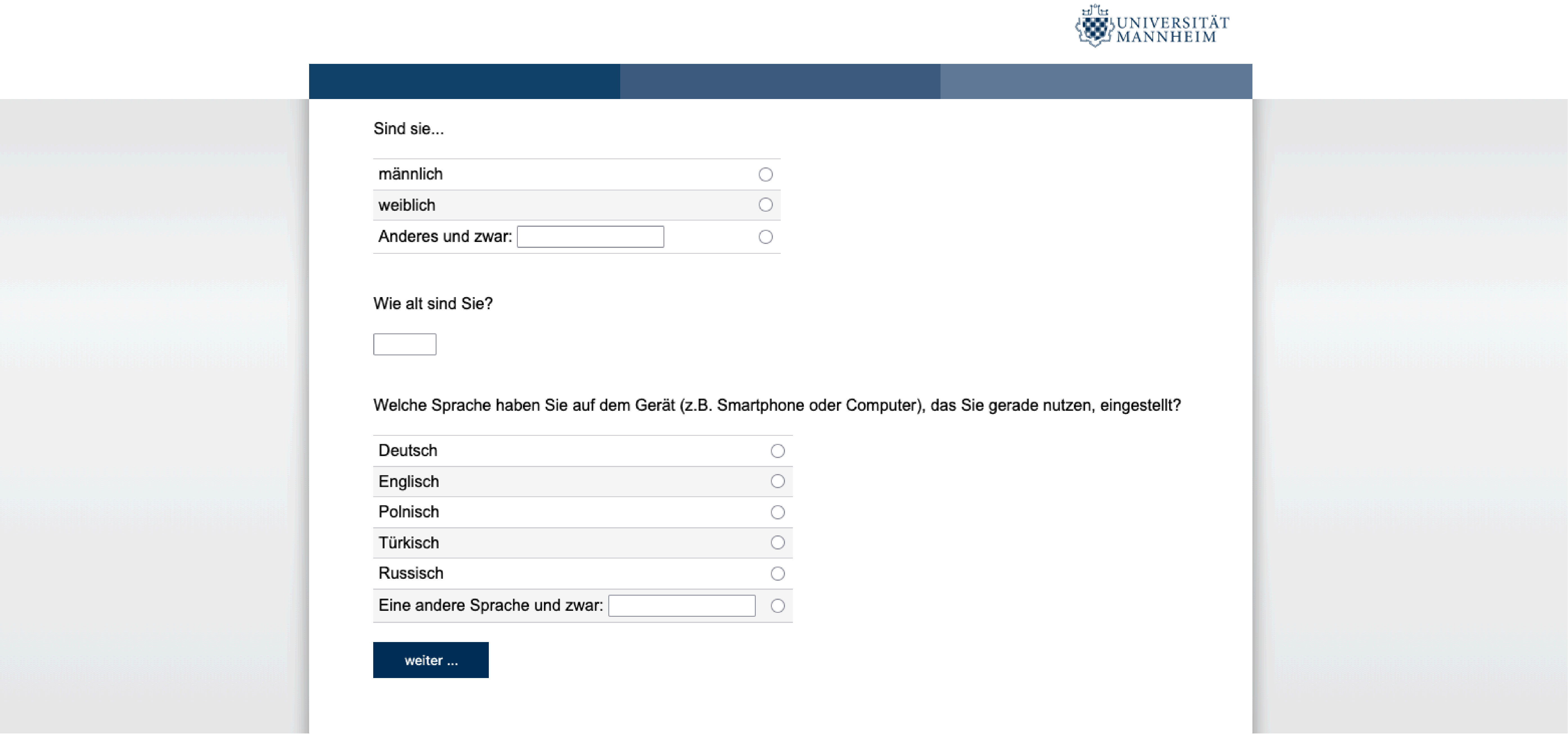
Request & Download Data

Extract Data

Inspect Data

Consent

Survey



UNIVERSITÄT MANNHEIM

Sind sie...

männlich	○
weiblich	○
Anderes und zwar:	○

Wie alt sind Sie?

Welche Sprache haben Sie auf dem Gerät (z.B. Smartphone oder Computer), das Sie gerade nutzen, eingestellt?

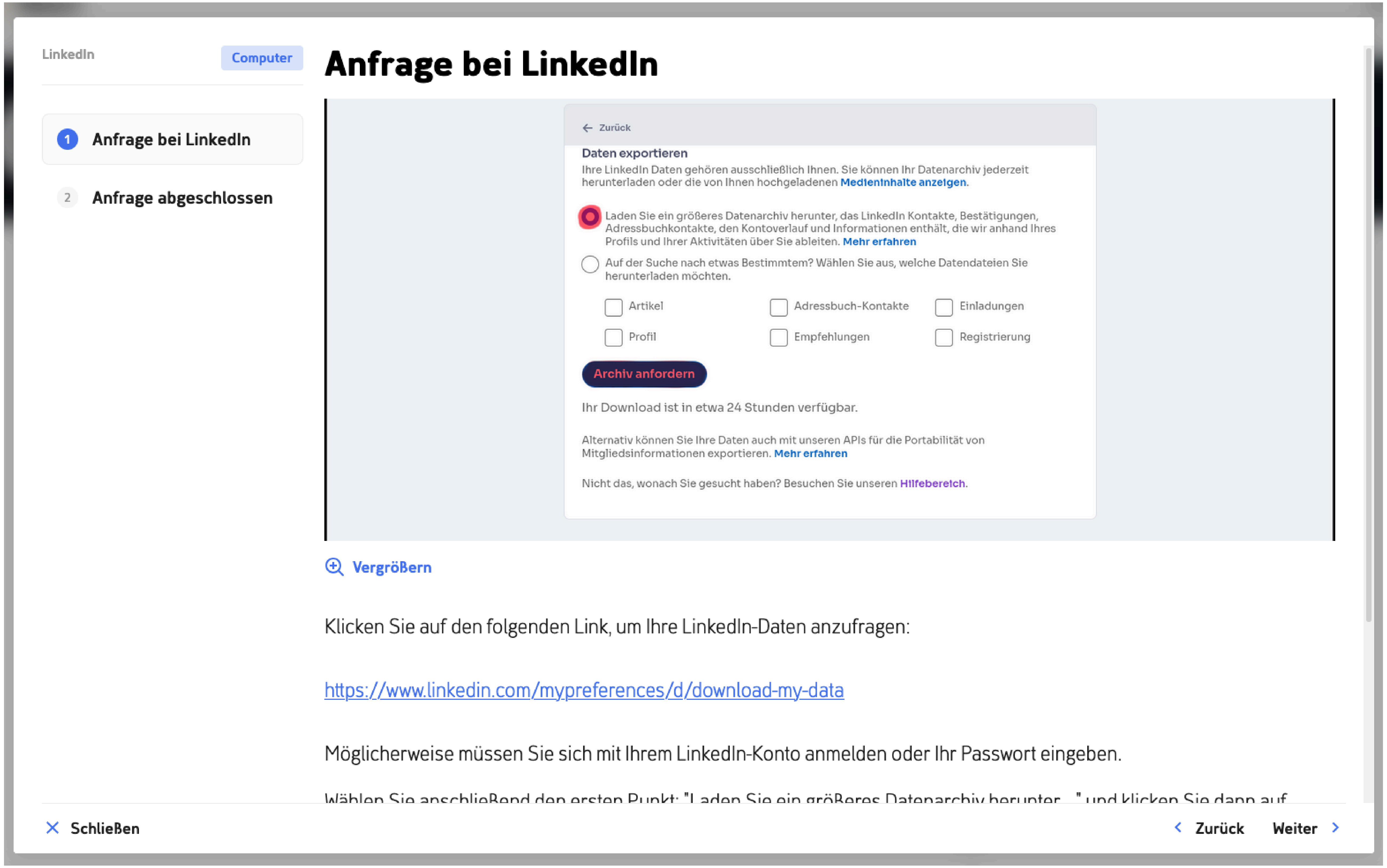
Deutsch	○
Englisch	○
Polnisch	○
Türkisch	○
Russisch	○
Eine andere Sprache und zwar:	○

Survey start page

Request & Download Data

Different degrees in standardization for data requests (Hase et al. 2024)...

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



LinkedIn

Computer

1 Anfrage bei LinkedIn

2 Anfrage abgeschlossen

Anfrage bei LinkedIn

← Zurück

Daten exportieren

Ihre LinkedIn Daten gehören ausschließlich Ihnen. Sie können Ihr Datenarchiv jederzeit herunterladen oder die von Ihnen hochgeladenen Medieninhalte anzeigen.

● Laden Sie ein größeres Datenarchiv herunter, das LinkedIn Kontakte, Bestätigungen, Adressbuchkontakte, den Kontoverlauf und Informationen enthält, die wir anhand Ihres Profils und Ihrer Aktivitäten über Sie ableiten. Mehr erfahren

○ Auf der Suche nach etwas Bestimmtem? Wählen Sie aus, welche Datendateien Sie herunterladen möchten.

□ Artikel □ Adressbuch-Kontakte □ Einladungen

□ Profil □ Empfehlungen □ Registrierung

Archiv anfordern

Ihr Download ist in etwa 24 Stunden verfügbar.

Alternativ können Sie Ihre Daten auch mit unseren APIs für die Portabilität von Mitgliedsinformationen exportieren. Mehr erfahren

Nicht das, wonach Sie gesucht haben? Besuchen Sie unseren Hilfebereich.

Vergrößern

Klicken Sie auf den folgenden Link, um Ihre LinkedIn-Daten anzufragen:

https://www.linkedin.com/mypreferences/d/download-my-data

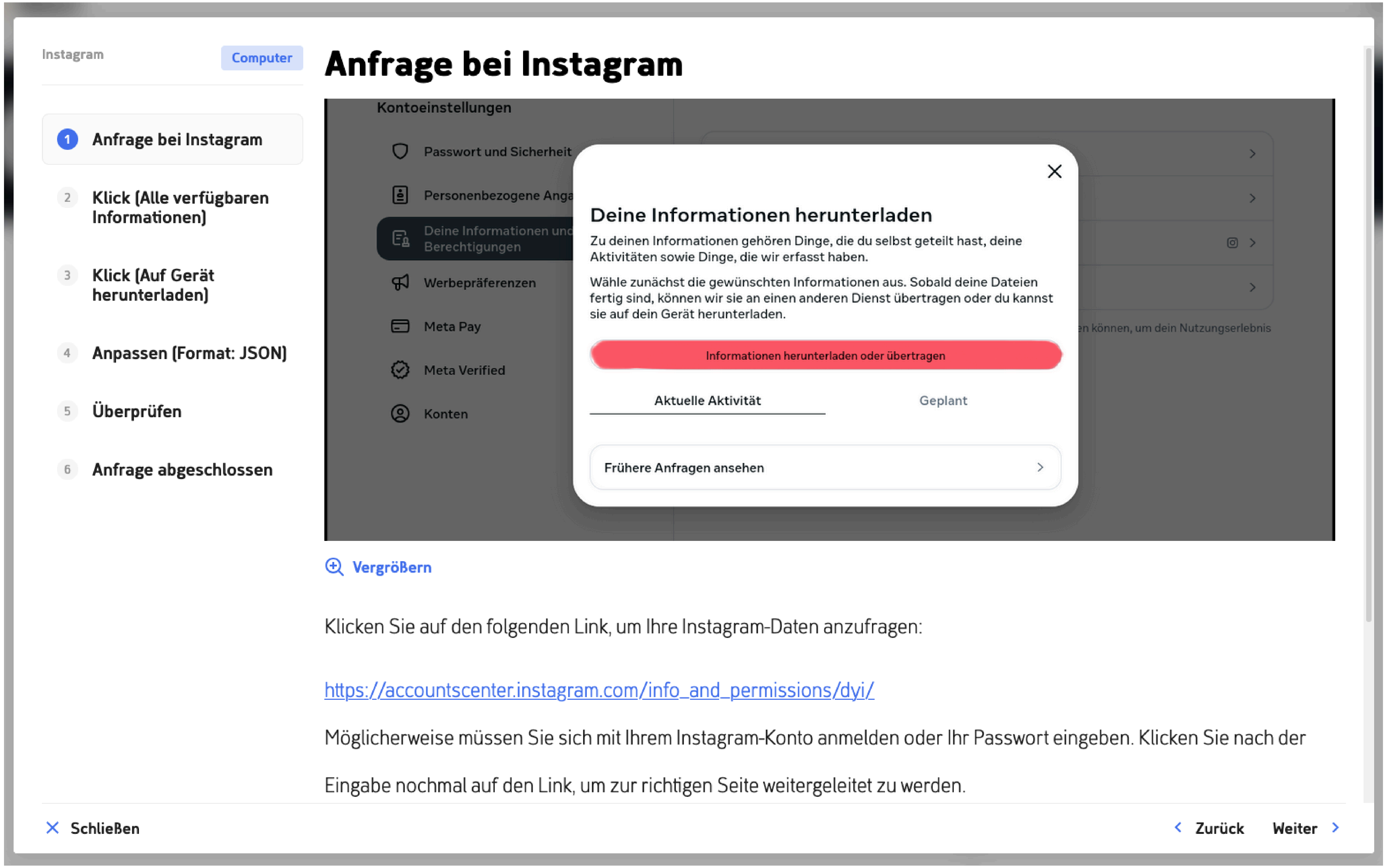
Möglicherweise müssen Sie sich mit Ihrem LinkedIn-Konto anmelden oder Ihr Passwort eingeben.

Wählen Sie anschließend den ersten Punkt: "Laden Sie ein größeres Datenarchiv herunter..." und klicken Sie dann auf

× Schließen

‹ Zurück Weiter ›

Request manual for LinkedIn on computer



1 Anfrage bei Instagram

2 Klick (Alle verfügbaren Informationen)

3 Klick (Auf Gerät herunterladen)

4 Anpassen (Format: JSON)

5 Überprüfen

6 Anfrage abgeschlossen

Anfrage bei Instagram

Deine Informationen herunterladen

Zu deinen Informationen gehören Dinge, die du selbst geteilt hast, deine Aktivitäten sowie Dinge, die wir erfasst haben.

Wähle zunächst die gewünschten Informationen aus. Sobald deine Dateien fertig sind, können wir sie an einen anderen Dienst übertragen oder du kannst sie auf dein Gerät herunterladen.

Klicken Sie auf den folgenden Link, um Ihre Instagram-Daten anzufragen:

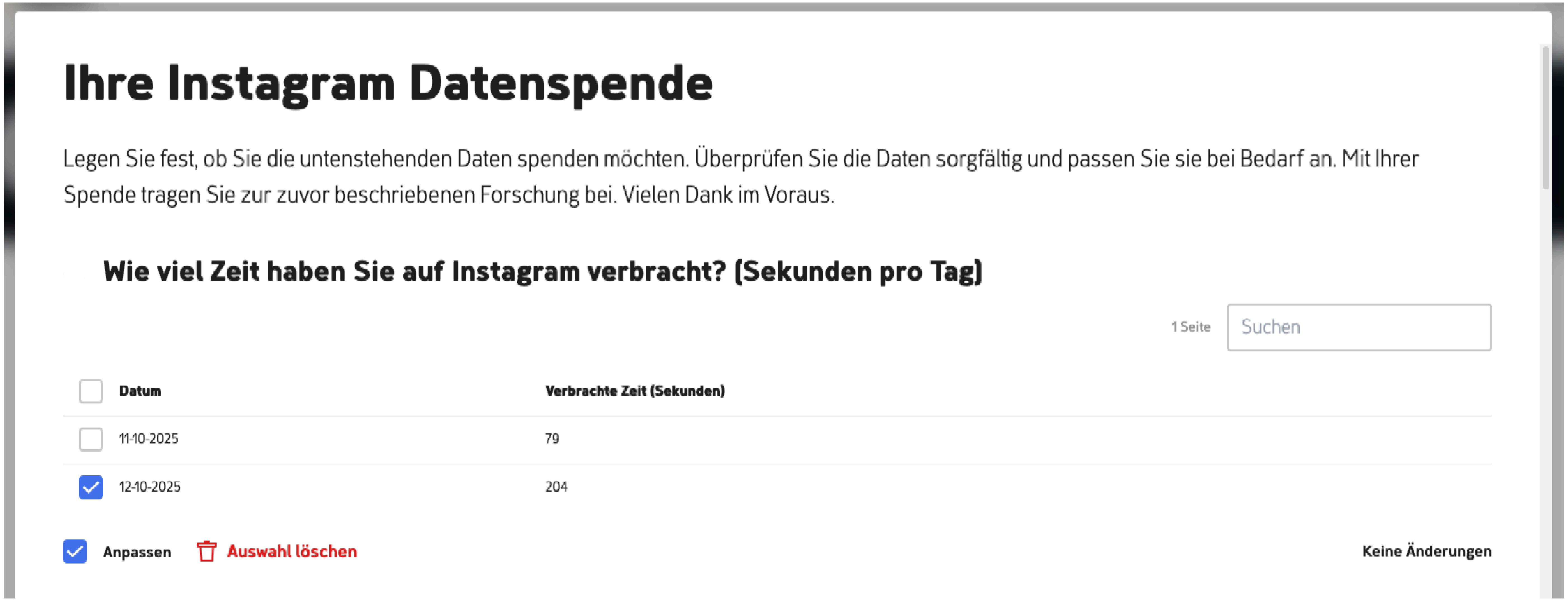
https://accountscenter.instagram.com/info_and_permissions/dyi/

Möglicherweise müssen Sie sich mit Ihrem Instagram-Konto anmelden oder Ihr Passwort eingeben. Klicken Sie nach der Eingabe nochmal auf den Link, um zur richtigen Seite weitergeleitet zu werden.

Request manual for Instagram on computer

⁘ Extract Data

Inspect Data



Ihre Instagram 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.

Wie viel Zeit haben Sie auf Instagram verbracht? (Sekunden pro Tag)

□	Datum	Verbrachte Zeit (Sekunden)
□	11-10-2025	79
☑	12-10-2025	204

Data overview on data donation platform

☑ Data donation

🤔 You might have already requested and downloaded your data in preparation for today. Did you encounter any difficulties in requesting and downloading your data?
We will dive into the content of your data a bit later.

🤔 Questions?

3️⃣ What is data donation?

The researcher’s perspective.

🤔 What are methodological decisions researchers have to take in data donation studies?

🔑 Key decisions

- Which theoretical questions do I want to answer?
- How do I operationalize key variables via my data donation tool?
- How do I integrate the tool in surveys & recruit participants?

Survey

Request & Download Data

Extract Data

Inspect Data

Consent

Decisions we took

Frame and Motivate:
Testing strategies to increase
participation in data donation studies.

Frieder Rodewald
University of Mannheim
f.rodewald@uni-mannheim.de
https://orcid.org/0009-0009-6859-5761

Florian Keusch
University of Mannheim
https://orcid.org/0000-0003-1002-4092

Valerie Hase
Ludwig-Maximilian-University Munich
https://orcid.org/0000-0001-6656-4894

Sebastian Prechsl
Institute for Employment Research, Germany
Ludwig-Maximilian-University Munich
https://orcid.org/0000-0001-9033-7317

Frauke Kreuter
Ludwig-Maximilian-University Munich
https://orcid.org/0000-0002-7339-2645

Mark Trappmann
Institute for Employment Research, Germany

Preprint Version
Version: 09. October 2025

- **Goal**: Feasability to implement data donation in a labor market panel, develop best-use practices, and understand non-response bias
- **Issues**: Data availability, low response rates and bias (e.g., (Keusch et al. 2024))
  - Behavioral intentions as "willingness to donate" high (79-52% of survey respondents)
  - Actual behavior as "participation in data donation" low (37-12% of survey respondents)
  - Well known intention-behavior gap; where 🤑 seems to help (Kmetty et al. 2025)
- **Sample**: A non-probability panel (online access panel)

Survey

- Survey concerning…
  - online platform usage (YouTube, Instagram, and LinkedIn),
  - labor market characteristics,
  - and common indicators for non-participation.

Data request & download

Data extraction

🛠️ Strategy to make the extraction work

1. Take a look at the DDP; download it, best for multiple time periods and for different languages
2. Understand the structure of the JSON or CSV.
3. Get an example file running.
4. Write the code for the extraction script.
5. Test your script, first locally and then in the wild.
6. Adapt your script.

📢 Task: Try it yourself.

Take a look at your downloaded data. What do you see; anything caught your eye?

Feel free to work in groups of 2-3 people for 5 minutes.

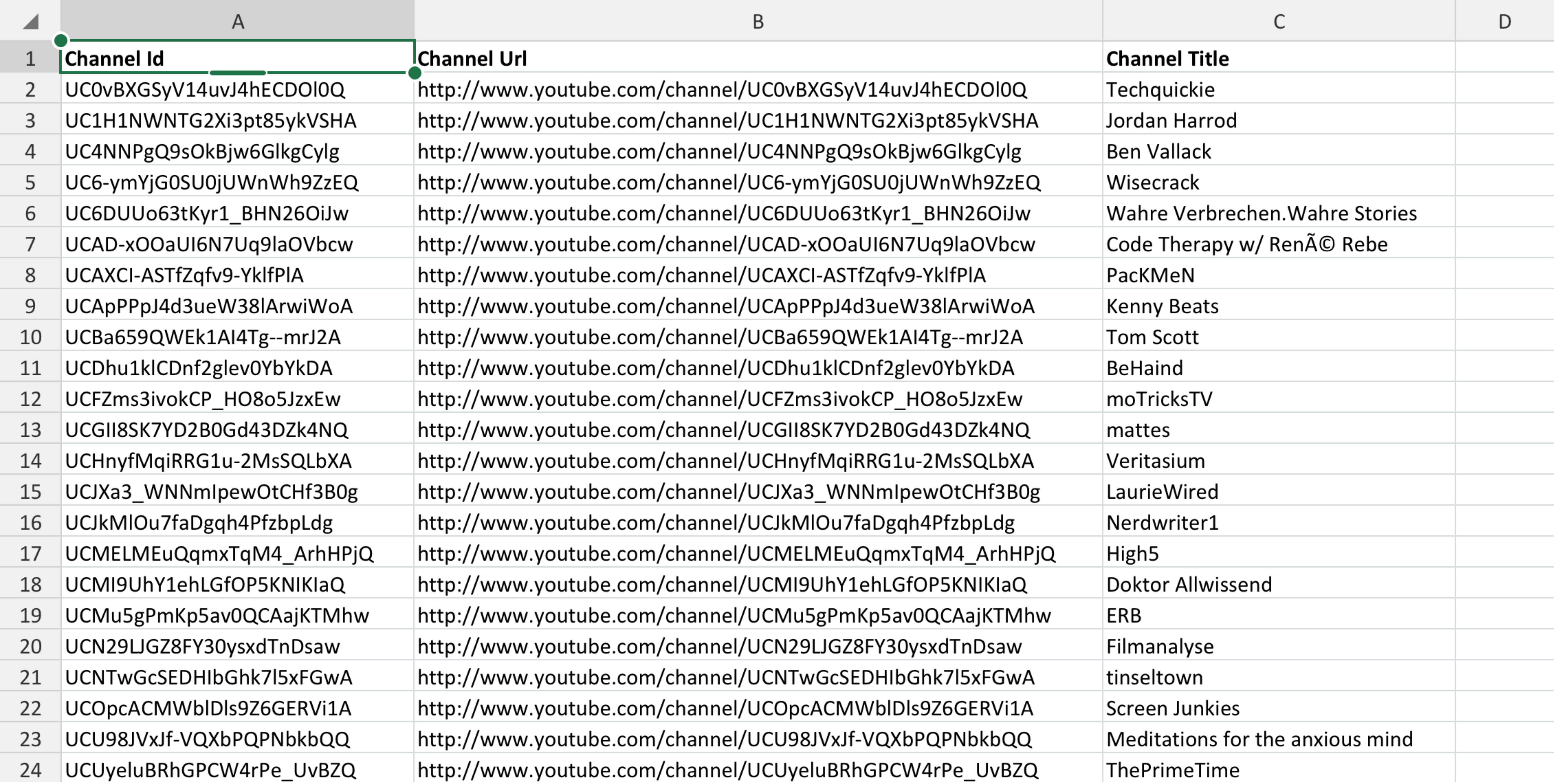
Different degrees in standardization for DDP content (Hase et al. 2024)...

- Documentation
    - DDP structure?
    - Measurements?
- Completeness & scope
    - Missing data?
    - Limited time frames?
    - Language sensitive?

🚨 Key issues (Hase et al. 2024)

- Missing documentation by platforms (e.g., file structure)
- Sudden changes in DDPs
- Differences across languages & devices
- Insufficient in-tool classification (e.g., LLM integration)

Example: Extract list of subscriptions



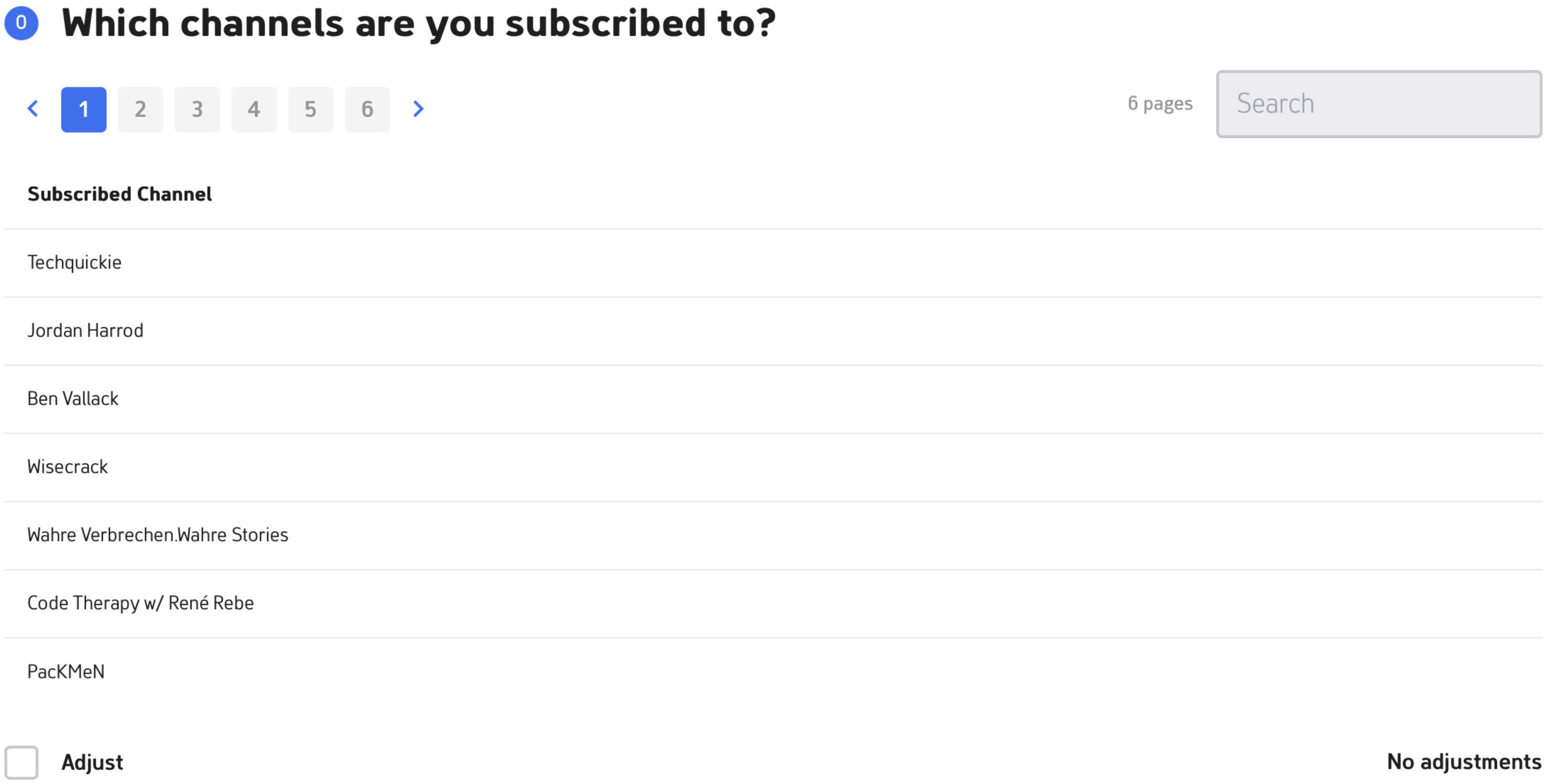
	A	B	C	D
1	**Channel Id**	**Channel Url**	**Channel Title**	
2	UC0vBXGSyV14uvJ4hECDOl0Q	http://www.youtube.com/channel/UC0vBXGSyV14uvJ4hECDOl0Q	Techquickie	
3	UC1H1NWNTG2Xi3pt85ykVSHA	http://www.youtube.com/channel/UC1H1NWNTG2Xi3pt85ykVSHA	Jordan Harrod	
4	UC4NNPgQ9sOkBjw6GlkgCylg	http://www.youtube.com/channel/UC4NNPgQ9sOkBjw6GlkgCylg	Ben Vallack	
5	UC6-ymYjG0SU0jUWnWh9ZzEQ	http://www.youtube.com/channel/UC6-ymYjG0SU0jUWnWh9ZzEQ	Wisecrack	
6	UC6DUUo63tKyr1_BHN26OiJw	http://www.youtube.com/channel/UC6DUUo63tKyr1_BHN26OiJw	Wahre Verbrechen.Wahre Stories	
7	UCAD-xOOaUI6N7Uq9laOVbcw	http://www.youtube.com/channel/UCAD-xOOaUI6N7Uq9laOVbcw	Code Therapy w/ RenÃ© Rebe	
8	UCAXCI-ASTfZqfv9-YklfPlA	http://www.youtube.com/channel/UCAXCI-ASTfZqfv9-YklfPlA	PacKMeN	
9	UCApPPpJ4d3ueW38lArwiWoA	http://www.youtube.com/channel/UCApPPpJ4d3ueW38lArwiWoA	Kenny Beats	
10	UCBa659QWEk1AI4Tg--mrJ2A	http://www.youtube.com/channel/UCBa659QWEk1AI4Tg--mrJ2A	Tom Scott	
11	UCDhu1klCDnf2glev0YbYkDA	http://www.youtube.com/channel/UCDhu1klCDnf2glev0YbYkDA	BeHaind	
12	UCFZms3ivokCP_HO8o5JzxEw	http://www.youtube.com/channel/UCFZms3ivokCP_HO8o5JzxEw	moTricksTV	
13	UCGII8SK7YD2B0Gd43DZk4NQ	http://www.youtube.com/channel/UCGII8SK7YD2B0Gd43DZk4NQ	mattes	
14	UCHnyfMqiRRG1u-2MsSQLbXA	http://www.youtube.com/channel/UCHnyfMqiRRG1u-2MsSQLbXA	Veritasium	
15	UCJXa3_WNNmIpewOtCHf3B0g	http://www.youtube.com/channel/UCJXa3_WNNmIpewOtCHf3B0g	LaurieWired	
16	UCJkMlOu7faDgqh4PfzbpLdg	http://www.youtube.com/channel/UCJkMlOu7faDgqh4PfzbpLdg	Nerdwriter1	
17	UCMELMEuQqmxTqM4_ArhHPjQ	http://www.youtube.com/channel/UCMELMEuQqmxTqM4_ArhHPjQ	High5	
18	UCMI9UhY1ehLGfOP5KNIKIaQ	http://www.youtube.com/channel/UCMI9UhY1ehLGfOP5KNIKIaQ	Doktor Allwissend	
19	UCMu5gPmKp5av0QCAajKTMhw	http://www.youtube.com/channel/UCMu5gPmKp5av0QCAajKTMhw	ERB	
20	UCN29LJGZ8FY30ysxdTnDsaw	http://www.youtube.com/channel/UCN29LJGZ8FY30ysxdTnDsaw	Filmanalyse	
21	UCNTwGcSEDHIbGhk7l5xFGwA	http://www.youtube.com/channel/UCNTwGcSEDHIbGhk7l5xFGwA	tinseltown	
22	UCOpcACMWblDls9Z6GERVi1A	http://www.youtube.com/channel/UCOpcACMWblDls9Z6GERVi1A	Screen Junkies	
23	UCU98JVxJf-VQXbPQPNbkbQQ	http://www.youtube.com/channel/UCU98JVxJf-VQXbPQPNbkbQQ	Meditations for the anxious mind	
24	UCUyeluBRhGPCW4rPe_UvBZQ	http://www.youtube.com/channel/UCUyeluBRhGPCW4rPe_UvBZQ	ThePrimeTime	

subscriptions.csv (before processing)

```
...
    "subscriptions": {
        "extraction_function": ef.extract_subscriptions,
        "possible_filenames": ["Abos.csv", "subscriptions.csv"],
        "title": {
            "en": "Which channels are you subscribed to?",
            "de": "Welche Kanäle haben Sie abonniert?",
            "nl": "Op welke kanalen ben je geabonneerd?",
        },
    },
...
```

```python
def extract_youtube_content_from_zip_folder(zip_file_path, possible_filenames):
    """Extract content from YouTube data export zip file using filenames"""

    try:
        with zipfile.ZipFile(zip_file_path, "r") as zip_ref:
            # Get the list of file names in the zip file
            filenames = zip_ref.namelist()
            # Look for matching files
            for possible_filename in possible_filenames:
                for filename in filenames:
                    if possible_filename in filename:
                        try:
                            # Process based on file extension
                            if filename.endswith(".json"):
                                with zip_ref.open(filename) as json_file:
                                    json_content = json.loads(json_file.read())
                                    return json_content
                            elif file_name.endswith(".csv"):
                                with zip_ref.open(file_name) as csv_file:
                                    csv_content = pd.read_csv(csv_file)
```

```python
def extract_subscriptions(subscriptions_csv):
    """Extract YouTube channel subscriptions"""

    # Define column name
    if "Kanaltitel" in subscriptions_csv.columns:
        channel_column = "Kanaltitel"
    else:
        channel_column = "Channel Title"

    # Define description
    channel_name = "Subscribed Channel"

    # Create DataFrame with just the channel names
    subscriptions_df = pd.DataFrame({channel_name: subscriptions_csv[channel_column]})
    return subscriptions_df
```



Which channels are you subscribed to?

Subscribed Channel
Techquickie
Jordan Harrod
Ben Vallack
Wisecrack
Wahre Verbrechen.Wahre Stories
Code Therapy w/ René Rebe
PacKMeN

□ Adjust

No adjustments

subscriptions.csv (after processing)

Data inspection

☑ Data donation

- A data donation platform helps to guide them through the process
- The process should be made as easy as possible for participants

📚 Workshop Takeaways

- There are many ways in which researchers can learn about people’s online behavior through digital trace data
- People can request, download and finally donate their (anonymized) data form most online platforms
- Data donation can be burdensome for participants
- Data quailty heavily depends on the platform and what kind of data you extract from participants

🙃 Thank you for participating; happy to talk with you about data donation (and anything else) throughout the next days.

EXTRA: Can I extend the data?

- Manual annotation by participants during data donation
- APIs/scraping to extend collected data
- Text-as-data methods for classification

EXTRA: Errors in representation

For example …

- **Coverage error**: Who is (not) represented in the sampling frame? (e.g., social media users vs. YouTube users)
- **Sampling error**: Who is (not) represented in the sample? (e.g., non-probability samples)
- **Non-response error**: Who does (not) want to participate in the data donation?
- **Compliance error**: Who is (not) able to participate in the data donation?

🤔 What do you think: Which participant characteristics may correlate with non-response or non-compliance?

EXTRA: What’s next for data donation studies?

Advancing the method

- 📸 Multimodal & cross-platform data (Wedel, Ohme, and Araujo 2024)
- In-tool, local classification (e.g., local ML/LLMs?)
- Workflow/UX-perspective

Data as a political tool

- Platforms do (willingly?) not provide data according to the GDPR/DSA (Hase et al. 2024)
- The EU has started to sanction platforms like X/TikTok
- DSA may become the subject of larger geo-political debates with the USA (Seiling, Ohme, and De Vreese 2025)

🚀 Can we improve & apply the method?

- Can the method actually be applied for empirical research? (few examples, like (Thorson et al. 2021; Wojcieszak et al. 2024))
- Requires interdisciplinary perspectives (e.g., addressing bias, integration in probability-based panels)

🤔 Questions?

References

Bach, Ruben L., Christoph Kern, Ashley Amaya, Florian Keusch, Frauke Kreuter, Jan Hecht, and Jonathan Heinemann. 2021. "Predicting Voting Behavior Using Digital Trace Data." *Social Science Computer Review* 39 (5): 862–83. https://doi.org/10.1177/0894439319882896.

Caliandro, Alessandro. 2024. "Follow the User: Taking Advantage of Internet Users as Methodological Resources." *Convergence: The International Journal of Research into New Media Technologies*, December, 13548565241307569. https://doi.org/10.1177/13548565241307569.

Carrière, Thijs C., Laura Boeschoten, Bella Struminskaya, Heleen L. Janssen, Niek C. de Schipper, and Theo Araujo. 2025. "Best Practices for Studies Using Digital Data Donation." *Quality & Quantity* 59 (1): 389–412. https://doi.org/10.1007/s11135-024-01983-x.

Christner, Clara, Aleksandra Urman, Silke Adam, and Michaela Maier. 2022. "Automated Tracking Approaches for Studying Online Media Use: A Critical Review and Recommendations." *Communication Methods and Measures* 16 (2): 79–95. https://doi.org/10.1080/19312458.2021.1907841.

Hase, Valerie, Jef Ausloos, Laura Boeschoten, Nico Pfiffner, Heleen Janssen, Theo Araujo, Thijs Carrière, et al. 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.

Jünger, Jakob. 2021. "A Brief History of APIs." In *Handbook of Computational Social Science, Volume 2*, 1st ed., 17–32. London: Routledge.

Jürgens, Pascal, and Birgit Stark. 2022. "Mapping Exposure Diversity: The Divergent Effects of Algorithmic Curation on News Consumption." *Journal of Communication*, March, jqac009. https://doi.org/10.1093/joc/jqac009.

Keusch, Florian, Paulina K. Pankowska, Alexandru Cernat, and Ruben L. Bach. 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): 279–93. https://doi.org/10.1177/1525822X231225907.

Kmetty, Zoltán, Ádám Stefkovics, Júlia Számely, Dongning Deng, Anikó Kellner, Edit Pauló, Elisa Omodei, and Júlia Koltai. 2025. "Determinants of Willingness to Donate Data from Social Media Platforms." *Information, Communication & Society* 28 (7): 1324–49. https://doi.org/10.1080/1369118X.2024.2340995.

Li, Xiao, Haowen Xu, Xiao Huang, Chenxiao Guo, Yuhao Kang, and Xinyue Ye. 2021. "Emerging Geo-Data Sources to Reveal Human Mobility Dynamics During COVID-19 Pandemic: Opportunities and Challenges." *Computational Urban Science* 1 (1): 22. https://doi.org/10.1007/s43762-021-00022-x.

Luiten, Annemieke, Joop Hox, and Edith de Leeuw. 2020. "Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study Across Countries and Surveys." *Journal of Official Statistics* 36 (3): 469–87. https://doi.org/10.2478/jos-2020-0025.

Ohme, Jakob, Theo Araujo, Laura Boeschoten, Deen Freelon, Nilam Ram, Byron B. Reeves, and Thomas N. Robinson. 2024. "Digital Trace Data Collection for Social Media Effects Research: APIs, Data Donation, and (Screen) Tracking." *Communication Methods and Measures* 18 (2): 124–41. https://doi.org/10.1080/19312458.2023.2181319.

Parry, Douglas A., Brittany I. Davidson, Craig J. R. Sewall, Jacob T. Fisher, Hannah Mieczkowski, and Daniel S. Quintana. 2021. "A Systematic Review and Meta-Analysis of Discrepancies Between Logged and Self-Reported Digital Media Use." *Nature Human Behaviour* 5 (11): 1535–47. https://doi.org/10.1038/s41562-021-01117-5.

Reiss, Michael V. 2023. "Dissecting Non-Use of Online News – Systematic Evidence from Combining Tracking and Automated Text Classification." *Digital Journalism* 11 (2): 363–83. https://doi.org/10.1080/21670811.2022.2105243.

Scharkow, Michael. 2016. "The Accuracy of Self-Reported Internet Use—A Validation Study Using Client Log Data." *Communication Methods and Measures* 10 (1): 13–27. https://doi.org/10.1080/19312458.2015.1118446.

Seiling, Lukas, Jakob Ohme, and Claes De Vreese. 2025. "Wird Europa Den DSA in Verhandlungen Mit Trump Opfern?" *Tagesspiegel*, March.

Sepulvado, Brandon, Michael Lee Wood, Ethan Fridmanski, Cheng Wang, Matthew J. Chandler, Omar Lizardo, and David Hachen. 2022. "Predicting Homophily and Social Network Connectivity From Dyadic Behavioral Similarity Trajectory Clusters." *Social Science Computer Review* 40 (1): 195–211. https://doi.org/10.1177/0894439320923123.

Sloan, Luke, Curtis Jessop, Tarek Al Baghal, and Matthew Williams. 2020. "Linking Survey and Twitter Data: Informed Consent, Disclosure, Security, and Archiving." *Journal of Empirical Research on Human Research Ethics* 15 (1-2): 63–76. https://doi.org/10.1177/1556264619853447.

Struminskaya, Bella, Peter Lugtig, Vera Toepoel, Barry Schouten, Deirdre Giesen, and Ralph Dolmans. 2021. "Sharing Data Collected with Smartphone Sensors." *Public Opinion Quarterly* 85 (S1): 423–62. https://doi.org/10.1093/poq/nfab025.

Thorson, Kjerstin, Kelley Cotter, Mel Medeiros, and Chankyung Pak. 2021. "Algorithmic Inference, Political Interest, and Exposure to News and Politics on Facebook." *Information, Communication & Society* 24 (2): 183–200. https://doi.org/10.1080/1369118X.2019.1642934.

Wagner, Michael W. 2023. "Independence by Permission." *Science* 381 (6656): 388–91. https://doi.org/10.1126/science.adi2430.

Wedel, Lion, Jakob Ohme, and Theo Araujo. 2024. "Augmenting Data Download Packages – Integrating Data Donations, Video Metadata, and the Multimodal Nature of Audio-visual Content." *Methods&#44; Data&#44; Analyses (Online First)*, October, 32 Pages. https://doi.org/10.12758/MDA.2024.08.

Wojcieszak, Magdalena, Ericka Menchen-Trevino, Bernhard Clemm Von Hohenberg, Sjifra De Leeuw, João Gonçalves, Sam Davidson, and Alexandre Gonçalves. 2024. "Non-News Websites Expose People to More Political Content Than News Websites: Evidence from Browsing Data in Three Countries." *Political Communication* 41 (1): 129–51. https://doi.org/10.1080/10584609.2023.2238641.

Yan, Pu, Ralph Schroeder, and Sebastian Stier. 2022. "Is There a Link Between Climate Change Scepticism and Populism? An Analysis of Web Tracking and Survey Data from Europe and the US." *Information, Communication & Society* 25 (10): 1400–1439. https://doi.org/10.1080/1369118X.2020.1864005.