SCM
Studies in Communication and Media

FULL PAPER

Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Warum spenden Menschen ihre Daten (nicht)? Ein Scoping-Review zu Faktoren, die die Teilnahme an Datenspende-Studien beeinflussen und Strategien zur Verbesserung dieser

Elisabeth Schmidbauer & Jörg Haßler

Elisabeth Schmidbauer (M. A.), LMU Munich, Department of Media and Communication, Akademiestraße 7, 80799 Munich, Germany. Contact: Elisabeth.Schmidbauer@ifkw.lmu.de. ORCID: https://orcid.org/0009-0002-0343-1343
Jörg Haßler (Assoc. Prof.), University of Oslo, Department of Media and Communication, Gaustadalléen 21, 0349 Oslo, Norway. Contact: joerg.hassler@media.uio.no. ORCID: https://orcid.org/0000-0003-2907-5228

FULL PAPER

Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Warum spenden Menschen ihre Daten (nicht)? Ein Scoping-Review zu Faktoren, die die Teilnahme an Datenspende-Studien beeinflussen und Strategien zur Verbesserung dieser

Elisabeth Schmidbauer & Jörg Haßler

Abstract: Data donations have recently emerged as research method, partly enabled by legislation such as the GDPR. A relevant challenge they face are low participation rates, which can introduce bias and limit the potential of data donations to harness digital trace data for studying behavior in various contexts. We conducted an interdisciplinary scoping literature review to map research on participation in data donation studies, outlining factors that might discourage participation and strategies to address these issues. To ensure methodological clarity, data donations are defined by clear criteria: The donation is an active and unpaid, but at times incentivized, transfer of intangible, reproducible, and retrospective, trace data. Furthermore, data donations are voluntary and informed, and primarily motivated by contributing to the common good. Based on this definition, which also served as the inclusion criteria for our study, we analyzed 30 studies. Included literature predominantly focuses on health research and hypothetical scenarios, exploring attitudes rather than actual donations. Although the investigated literature is diverse, several factors are discussed to influence willingness to donate. These include respondent characteristics such as gender differences and technical skills, situational factors such as privacy concerns and motivations including egoism, study-related attributes including required data types and who receives the data, and societal-level factors such as contextual influences and culture. Strategies to increase willingness include general approaches like offering benefits and the right communication, as well as study design–related strategies, particularly cultivating trust, e.g., through credible collaborations and appropriate privacy protection.

Keywords: Data donations, participation-bias, literature review, trace data

Zusammenfassung: Datenspenden haben sich, begünstigt durch gesetzliche Rahmenbedingungen wie die DSGVO, als Forschungsmethode etabliert. Eine zentrale Herausforderung sind jedoch niedrige Teilnahmequoten, die zu Verzerrungen führen können und das Potenzial von digitalen Spurdaten zur Analyse von Verhalten einschränken. Mithilfe eines interdisziplinären Scoping Literature Reviews systematisieren wir in diesem Beitrag den Forschungsstand zur Teilnahme an Datenspendenstudien sowie Einflussfaktoren und mögliche Strategien zur Steigerung der Teilnahmebereitschaft. Wir definieren Datenspenden als aktiven, unbezahlten – aber mitunter incentivierten – Transfer von immateriellen reproduzier-

baren, rückblickenden Daten. Zudem sind Spenden freiwillig, informiert, mit Gemeinwohlbezug. Auf Basis dieser Definition, die zugleich als Einschlusskriterium diente, analysierten wir 30 Studien. Die Literatur stammt überwiegend aus der Gesundheitsforschung, fokussiert auf hypothetische Szenarien und untersucht primär Einstellungen statt tatsächlicher Teilnahme. Dennoch lassen sich zentrale Einflussfaktoren identifizieren, darunter individuelle Merkmale wie Geschlecht und technische Fähigkeiten, situative Aspekte wie Datenschutzbedenken oder Egoismus, studienspezifische Eigenschaften wie Datentypen und -Empfänger sowie gesellschaftliche Kontexte. Generelle Strategien zur Erhöhung der Teilnahmebereitschaft umfassen beispielsweise Anreize zu schaffen und zielgerichtete Kommunikation sowie den Aufbau von Vertrauen, etwa durch glaubwürdige Kooperationen und transparente Datenschutzmaßnahmen.

Schlagwörter: Datenspenden, Teilnahmeverzerrung, Literaturübersicht, Spurendaten

1. Introduction

A considerable proportion of daily activities now occurs within the digital sphere. Online behavior, whether intentional or not, leaves digital trace data, defined as "records of activity […] undertaken through an online information system" (Howison et al., 2011, p. 769). These traces are rich data sources that enable more accurate and elaborate measurement of various phenomena (Araujo et al., 2017; Ohme et al., 2021; Scharkow, 2016; Wedel et al., 2024). They also help mitigate biases common in self-reported measures, such as reactivity (Baumgartner et al., 2023) or social desirability (Boeschoten et al., 2022).

Acquiring trace data is challenging because companies often withhold data and limit researchers' access (Breuer et al., 2022, p. 2). One approach to improve accessibility is data donation, in which individuals contribute their personal data to research. A prominent example is the donation of data download packages (DDPs). Under the European Union's General Data Protection Regulation (GDPR; EU Regulation 2016/679, 2016), users can request access to data platforms have stored about them (e.g., Pak et al., 2022, p. 135). Platforms provide this data in form of DDPs. These DDPs can then be shared with researchers, providing access to trace data on user behavior across platforms (Boeschoten et al., 2022; van Driel et al., 2022). For instance, Instagram DDPs may include data on likes, direct messages, and search queries (van Driel et al., 2022, p. 272).

Data gathered through donations can enhance research across disciplines, including communication and health, by improving data quality and providing insights into digital behavior. In communication studies, for example, Google DDP donations can be used to examine how citizens inform themselves about political topics before a referendum (Blassnig et al., 2023). In health research, data donations can support the measurement of physical activity and the analysis of related health outcomes (Keusch, Struminskaya, et al., 2024).

However, trace data, such as data donations, are subject to various biases and limitations (Amaya et al., 2020; Boeschoten et al., 2022; Hase & Haim, 2024). One stems from DDPs themselves. Because platforms create and control DDPs, their

structure and content can change at any time. This can introduce inconsistencies, ambiguities, and missing information, potentially compromising their reliability in research (Hase et al., 2024). To reduce errors in data donations, research emphasizes the need for best practices and clear recommendations (Carrière et al., 2025; Ohme & Araujo, 2022).

Much of the literature highlights data donations' susceptibility to non-response bias (Bietz et al., 2019; Ohme et al., 2023; Welbers et al., 2024). Participation rates are often low (Ohme et al., 2021; Pak et al., 2022), making it difficult to obtain representative samples (Ohme et al., 2023; Welbers et al., 2024). Improving understanding of the factors that predict non-response (e.g., Kmetty et al., 2024; Pfiffner & Friemel, 2023; Strycharz et al., 2024) is therefore essential for developing strategies to mitigate this bias.

This scoping review examines the literature on participation bias in data donation studies. It aims to identify respondent characteristics (e.g., age, gender), situational factors (i.e., motivations and study attributes), and societal influences shaping *why people do (not) donate their data*. It also seeks to identify strategies to increase participation.

2. Theory

2.1 What are data donations?

Data donation is a user-centric method of trace data collection in which participants share data with researchers (Boeschoten et al., 2022; Breuer et al., 2022; Haim & Hase, 2023; Ohme et al., 2023). As a distinct method, it must be differentiated from other forms of data sharing to understand the factors influencing participation. We therefore define data donations based on seven characteristics.

Data donations require participants' 1) *active involvement* in contributing data they hold (e.g., Bietz et al., 2019; Gomez Ortega et al., 2021). This distinguishes them from methods such as data linkage. For example, when accessing Twitter data via linkage, participants only provide their user handles in a survey, allowing researchers to access public profiles and additional information (Silber et al., 2022). In contrast, data donation enables participants to share Twitter DDPs, which contain more detailed and non-public information (Boeschoten et al., 2022; Silber et al., 2022; Stier et al., 2020). Some studies also allow participants to review and delete data before donating, further increasing active involvement (e.g., Breuer et al., 2022; Zannettou et al., 2023).

Additionally, data donations are 2) *unpaid*, as participants are not directly compensated for their data (Susha et al., 2019, p. 118). While incentives may be offered to compensate time and participation, they do not constitute payment for the data itself and thus preserve the notion of donation for the common good (Bietz et al., 2019, p. 8).

Next, data is 3) *not generated for research* (Bietz et al., 2019, p. 1). This study focuses on trace data, meaning historical records of participants' digital behavior (Pak et al., 2022, p. 135). Examples include data from fitness trackers (Pilgrim & Bohnet-Joschko, 2022), smartphone usage statistics (Baumgartner et al., 2023;

Ohme et al., 2021), social media data (e.g., Boeschoten et al., 2022; Pak et al., 2022) or Google search histories (Blassnig et al., 2023). Trace data is generated by systems and reflects user behavior (Howison et al., 2011), offering a more objective account than self-reports (Ohme et al., 2021). To qualify as a valid donation, data must be captured non-reactively, meaning it is not self-reported but provided as system-exported files (e.g., DDPs) or other materials containing recorded behavior, such as screenshots.

The next facet concerns the 4) *intangible and reproducible nature* of data donations, reflected in "data multiplicity" (Prainsack, 2019, p. 13). Unlike physical donations, data can be shared multiple times without loss to donors. However, donating data may feel less rewarding, as other forms of donation, such as blood donation, can evoke a 'warm glow' (e.g., Skatova & Goulding, 2019, p. 3).

Another aspect is that data is donated 5) *retrospectively*, meaning consent is given ex-post (Stier et al., 2020, p. 506), for data that already exists. The data were generated in the past, logged by platforms or services, and donated later. This distinguishes data donations from other user-centric trace data collection methods, such as tracking, where participants provide consent ex-ante and install tools that continuously collect data (Gil-López et al., 2023; Stier et al., 2020, p. 506). Retrospective donation can help mitigate reactivity during data collection (e.g., Baumgartner et al., 2023, p. 5).

Further, participants donate their data 6) *voluntarily and with informed consent* (e.g., Bietz et al., 2019, p. 1). They are informed about what data is requested, the procedures involved, and how the data will be handled, analyzed, and shared.

Finally, data donations are 7) *intended to serve the common good* and support societal well-being, as they are typically motivated by research rather than financial interests (Bietz et al., 2019; Hillebrand et al., 2023). However, the societal value of donated data depends on the type of research and, in particular, on researchers' ethical and responsible use of the data.

Summarizing, this study defines data donations as: 1) an *active* and 2) *unpaid* transfer of personal 3) *trace data not generated* for research to researchers. The data is 4) *intangible* and *reproducible,* and collected 5) *retrospectively*. Consent is 6) *voluntary and informed*, and 7) the study is intended to *serve the common good*.

To illustrate our definition more clearly, we provide examples of two data types and research designs that do not align with it and are therefore excluded from this study, namely genomic data and posthumous data sharing. First, our definition requires active data transfer by participants. In many cases, genomic data is already held by researchers or medical professionals, and participants only consent to its use, resembling data linkage rather than active donation. Second, our focus is on trace data reflecting past behavior. Although genomic data is highly personal, it does not capture behavior and is therefore not relevant to this review. While it can originate outside research contexts, such as medical diagnostics or ancestry services, it is often collected specifically for research, whereas our focus is on data generated independently of research. Third, unlike the fourth aspect of our definition, genomic data collection often involves a physical component, such as providing tissue samples (e.g., Mezinska et al., 2020, p. 3). This entails physical

loss and potential health risks, distinguishing it from the intangible nature of data donations as defined here.

A second excluded example is posthumous data provision. In this case, consent is not given retrospectively but earlier in life, when not all data donated at death may yet exist. This violates the fifth part of our definition.

2.2 Research on participation in data donation studies

Participation bias in data donation studies is difficult to assess due to varying definitions and uses of the term across disciplines. While common in communication research (e.g., Boeschoten et al., 2022; van Driel et al., 2022) and health research (e.g., Bietz et al., 2019), it also appears in fields such as human-computer interaction (e.g., Gomez Ortega et al., 2021), often with limited cross-disciplinary awareness.

Literature on health-related data donations is most prominent in research (Susha et al., 2019) and reports high rates of willingness to participate, ranging from 65% (Seltzer et al., 2019, p. 2) to 84% (Buhr et al., 2022, p. 9). In contrast, in communication research, the likelihood of donation tends to fall below the midpoint of a scale (e.g., Pfiffner & Friemel, 2023, p. 15) and donation rates are often quite low, around 20% (Hase & Haim, 2024; Keusch, Pankowska, et al., 2024; Strycharz et al., 2024). Additionally, different types of data might be prone to different perceptions and biases (e.g., Pak et al., 2022; Pfiffner & Friemel, 2023), as health data is perceived as more sensitive than social media data (Silber et al., 2022, p. 401).

Altogether, findings on (non-)participation in data donation studies are often not integrated. Thus, this review seeks to provide a comprehensive overview of research, mapping factors that may influence (non-)participation. By including multiple disciplines, it offers a broad perspective and insights beyond social media donations. Therefore, we ask:

> *RQ1: What characterizes studies on willingness to donate (e.g., research fields, methodology, hypothetical vs. actual donation)?*

2.3 What factors might influence data donations

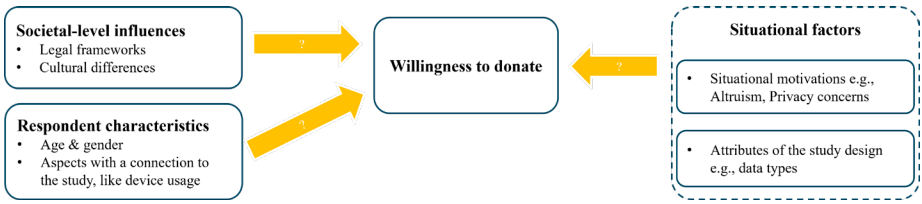
To examine participation in data donation, we draw on Groves et al.'s (1992) framework of survey non-response. It distinguishes three factors: (1) characteristics of the sample person, (2) societal-level factors, and (3) attributes of the survey design. As data donation studies typically lack face-to-face interaction, we exclude interviewer-related aspects such as interviewer attributes and respondent–interviewer interaction.

First, we adopt the category 'characteristics of the sample person', which we term *respondent characteristics*. These are stable demographic variables, such as age or gender, that are independent of the data donation context. Second, we examine *situational influences*, including '*attributes of the survey design*' as adopted from Groves et al. (1992) as one sub-dimension. This covers factors such as the

wording of the data request, which can affect willingness to donate (e.g., Silber et al., 2024). As a second sub-dimension, we add *situational motivations*, such as privacy concerns. Unlike stable demographic variables, these can vary depending on the situation and perceived risks or circumstances

Finally, we consider '*societal-level factors*', referring to broader influences such as legal frameworks (e.g., GDPR), national culture, or major events like health crises. These factors shape general attitudes toward research and can influence data-sharing behavior. Together, these categories structure our analysis of participation (see Fig. 1) and are discussed in the following sections.

Figure 1. Categories of analysis



2.3.1 Respondent characteristics influencing non-participation

First, we lay a focus on *respondent characteristics* that influence participation, also referred to as dispositional variables (Strycharz et al., 2024). Survey research shows that such characteristics, including sociodemographic traits like gender (Keusch, 2015, p. 189) and personal interest in the topic (p. 186), can affect participation in research.

Similarly, research on data donation has observed differences between donors and non-donors, for example, regarding political interest (Hase & Haim, 2024, p. 17), age and digital and algorithmic efficacy (Strycharz et al., 2024, p. 15), as well as general trust in research (Keusch, Pankowska, et al., 2024, p. 8) Therefore, we ask:

> *RQ2: Which respondent characteristics (e.g., sociodemographic characteristics, device usage) are discussed as potential determinants of willingness to donate?*

2.3.2 Situational factors influencing non-participation

Situational factors influencing non-participation can be divided into two subcategories: *Attributes of the survey design* and *situational motivations*. According to Groves et al. (1992, p. 477–478), '*attributes of the survey design*' refer to study characteristics and design choices that affect participation decisions. This factor is also relevant for data donation studies, as elements such as the type of data requested can influence willingness to donate (Kmetty et al., 2024; Silber et al., 2024).

We also assume that *situational motivations* are relevant. In data donation studies, participants share personal data, so privacy concerns may influence partici-

pation (Boeschoten et al., 2022; Ohme et al., 2021; Pfiffner & Friemel, 2023). The Privacy Calculus framework helps explain this behavior, suggesting that individuals weigh potential benefits and risks when deciding whether to disclose information, though decisions may not always be rational or consistent with general attitudes (Aquisti, 2004; Dinev et al., 2015).

This suggests that privacy concerns are not stable traits but context-dependent states. Decisions are made within specific situations and thus reflect situational motivations. Similarly, tendencies such as altruism, often considered a stable trait, may also depend on context. Skatova and Goulding (2019) show that motivations for data donation extend beyond general altruism and are more context-dependent. Accordingly, this study emphasizes situational factors influencing participation in data donation.

> *RQ3: What situational factors (i.e., situational motivations and attributes of the study) are discussed as influencing willingness to donate?*

2.3.3 Societal-level influences on participation

Finally, Groves et al. (1992) identify 'societal-level factors' as influences on survey participation, referring to the broader societal context of a request (p. 477). In data donation, cultural differences may shape openness to data sharing across populations. Legal frameworks such as the GDPR can influence both the feasibility of such research and public awareness of data access. Broader events, such as health crises, may also affect attitudes toward research and willingness to donate data. Consequently, the following research question is posed:

> *RQ4: What societal-level factors (e.g., cultural aspects, jurisdiction) are discussed as influencing willingness to donate?*

2.3.4 Which strategies might enhance participation?

Research on data donations has examined not only predictors of non-participation (Pak et al., 2022; Pfiffner & Friemel, 2023; Silber et al., 2022; Strycharz et al., 2024; Welbers et al., 2024) but also strategies to increase participation. In particular, ex-ante strategies targeting survey design are a key focus (de León et al., 2025; Hase & Haim, 2024). One approach is monetary incentives, which can be effective for certain data types (Kmetty et al., 2024; Silber et al., 2022) though they may introduce bias (Zannettou et al., 2023, p. 8).

Effective strategies must address the specific benefits and concerns of data donation, and some studies have begun to explore these (de León et al., 2025; Hase & Haim, 2024; Keusch, Pankowska, et al., 2024). For example, researchers may offer personalized feedback or insights about participants' data (Gomez Ortega et al., 2021, p. 498). Addressing privacy concerns is also essential, for instance, through strong data protection measures (e.g., Boeschoten et al., 2022, p. 408).

Nudges have also been proposed to increase survey participation. They are defined as "any aspect of the choice architecture that alters people's behavior in a

predictable way without restricting options or significantly changing economic incentives" (Thaler & Sunstein, 2008, p. 6). Like data donations, nudges aim to promote behavior that benefits individuals or society.

One example is framing, used in survey design to increase willingness to donate data. This can involve emphasizing the societal value of data donation (Hase & Haim, 2024) or highlighting the potential loss of valuable data if individuals do not donate (Keusch, Pankowska, et al., 2024). Although framing is not always effective, nudges remain a promising strategy. This leads to our final research question:

> *RQ5: What strategies (i.e., general or relating to study design) are proposed in the literature to increase participation willingness?*

3. Method

To address the research questions, we conducted a scoping review, a type of literature review used to map and organize diverse research fields (Arksey & O'Malley, 2005; Levac et al., 2010). Scoping reviews capture the breadth of research and identify overarching themes and trends (Arksey & O'Malley, 2005, p. 21). They also enable the integration of fragmented findings and the identification of factors influencing a concept (Peters et al., 2020, p. 2122). This approach is therefore well suited to examining factors influencing willingness to donate across disciplines. A scoping review is more appropriate than a systematic review for this study, as systematic reviews require narrowly defined questions on constructs, effect directions, or specific designs (Munn et al., 2018; Peters et al., 2020). This would be difficult given the diversity of research on willingness to donate and the exploratory focus of this study.

A research protocol (Booth et al., 2022; Peters et al., 2020; see online supplement 1 in OSF) was developed in advance according to PRISMA-P guidelines (Moher et al., 2015). The following section reports required elements and deviations from the protocol in line with PRISMA-ScR guidelines (Tricco et al., 2018).

3.1 Outcomes

The primary outcomes aim to address the research questions by identifying factors examined in the literature, such as RQ2) respondent characteristics, RQ3) situational influences, RQ4) societal-level influences on willingness to donate and RQ5) strategies to enhance participation. To address RQ1, we also examine the scope of the literature across disciplines, trends over time, types of data donated, and described donation processes.

3.2 Deviations from the protocol

We were unable to fully adhere to the research protocol, as outcomes such as actual donation and compliance rates were omitted. This was due to the inclusion of many hypothetical studies, where such measures were not applicable. Additio-

nally, willingness to donate was measured inconsistently across studies, for example, through categorical responses ("yes", "no", "don't know"; e.g., Buhr et al., 2022) or continuous scales from 0 to 100% (Hillebrand et al., 2023), limiting comparability. No measures of strategy effectiveness were reported.

Finally, the research question on situational influences was divided into two, placing greater emphasis on societal-level factors as a distinct category.

3.3 Eligibility criteria

To be included, studies had to be published between January 1, 2018, and April 17, 2023, when data collection occurred. The cutoff reflects the introduction of the GDPR, which likely facilitated data donation research. In the same year, iOS and Android introduced phone usage statistics (Android.com, 2018, September 25; Apple Inc., 2018, September 25), enabling access to behavioral data that can be donated, for example, via screenshots (Ohme et al., 2021).

Studies were selected based on inclusion criteria aligned with our definition: 1) an active and 2) unpaid transfer of personal 3) trace data not generated for research to researchers. The data must be 4) intangible and reproducible, 5) collected retrospectively, with 6) voluntary and informed consent, and 7) an aim to serve the common good.

To focus on literature relevant to the research questions, we excluded studies on 1) feasibility, 2) ethics, 3) legal reviews, 4) technical frameworks, and 5) applications of data donation without a focus on participation or bias. To comprehensively understand factors influencing willingness to donate, both empirical and theoretical work were considered. Grey literature was not systematically searched, but was included if identified.

3.4 Search strategy

To identify primary literature, we conducted database and Google Scholar searches using the same terms. The databases included 1) Web of Science, 2) ScienceDirect, 3) EBSCOhost (Business Source Complete, Academic Search Complete, Communication & Mass Media Complete, APA PsycInfo, APA PsycArticles, and Information Science & Technology), and 4) ACM Digital Library. Search results were downloaded as BibTeX files for screening. For Google Scholar, only the first five result pages were considered due to the high number of results (Haim et al., 2018).

To identify secondary literature, we consulted an expert (e.g., Booth et al., 2022), who suggested relevant studies after being briefed on the review's goals. We also conducted a backward citation search (Peters et al., 2020) by screening references of included studies and reviewing abstracts of unfamiliar but potentially relevant sources.

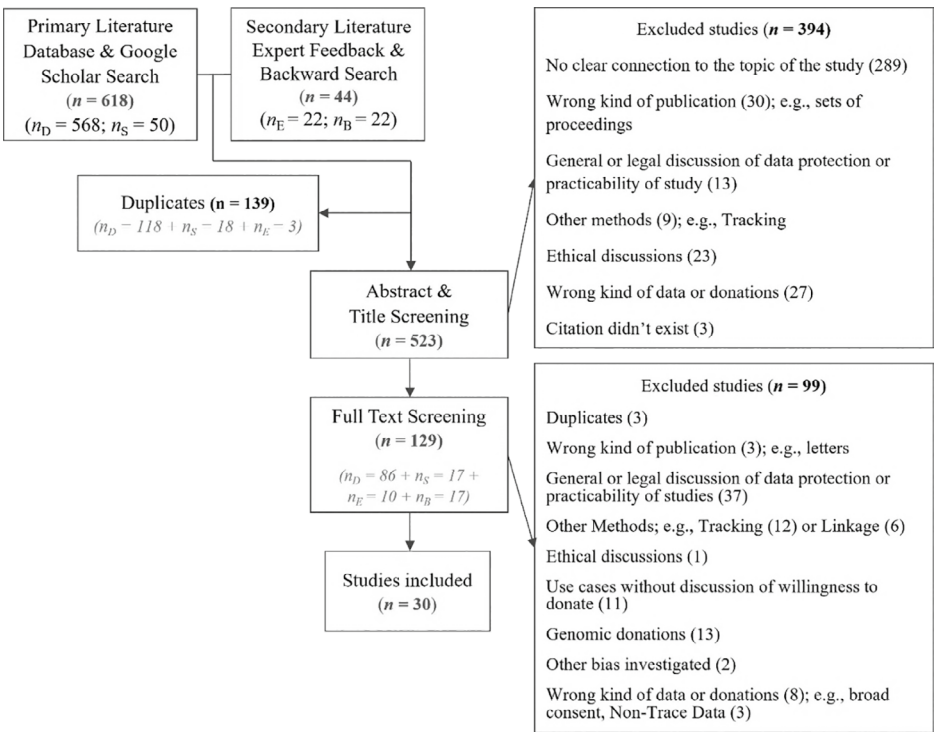
Search terms were developed in English and German and refined based on expert feedback: "Data don*" OR (donat* AND "Digital Trace") OR ("usercentric" AND donat*) OR ("user-centric" AND donat*) OR ("user centric" AND donat*) OR Datenspend* OR "Daten spend*". Where truncation was not possible, database-specific filters were applied.

3.5 Selection process

The first stage of the selection process (Fig. 2), abstract screening, was conducted using Rayyan.ai (Kellermeyer et al., 2018). BibTeX citations from database searches were uploaded, and papers were manually labeled as included for full-text screening or excluded based on their abstracts. Missing abstracts were retrieved and screened manually. Additional features of Rayyan.ai[1], such as word clouds, were not used. Google Scholar results and secondary literature were documented in a Microsoft Word file, where eligibility decisions were also recorded.

In the second stage, full texts were screened in Citavi[2], and eligible studies underwent backward citation screening. After the selection process (Fig. 2), 30 articles were included (see online Appendix 2). Exclusions labeled "Citation didn't exist" refer to three database entries for which no corresponding publications could be identified.

Figure 2. Selection process



Notes. n_D = Database; n_S = Google Scholar; n_E = Expert feedback; n_B = Backward search

1 https://www.rayyan.ai/
2 https://www.citavi.com/de

3.6 Data charting

Data charting included both empirical findings and theoretically discussed factors and strategies reported in the papers. Categories were predefined (online supplement 1 in OSF) and revised after charting the first five articles.

We excluded categories such as goals or research questions, population, sample size, key findings, strategy effectiveness, donated data, initial response rate, and compliance rate, as these outcomes were often unavailable due to the hypothetical nature of many studies. We added the categories "type of publication" and "applied strategies or incentives to enhance participation," and refined others to better capture the construct of interest (online supplement 3 in OSF).

Study fields were coded by the first author, primarily based on the journals. When unclear, the paper's main topic and author affiliations were considered, and a second field was assigned when necessary.

3.7 Data analysis

The analysis strategy was not specified in the research protocol due to the unpredictable diversity of the data. We conducted an inductive thematic analysis (Braun & Clarke, 2006), a method used in scoping reviews (e.g., Severin-Nielsen, 2023) to identify themes, here understood as influencing factors. Relevant text passages were abstracted, summarized, and consolidated into overarching factors (Tab. 1). These were included as subcategories to describe the identified factors (online supplement 4 in OSF).

Table 1. Example of data analysis

	Abstraction	Factor	Subcategories
"Patients also expressed concerns about potential data and privacy breaches" (Seltzer et al., 2019, p. 3)	Privacy concerns	Privacy concerns	Security concerns, negative experiences
"Willingness toward sharing PHD is related with privacy and security considerations." (Karampela et al., 2019, p. 9)	Privacy & security concerns		
"Privacy-Related Experiences: - Was the victim of fraud and/or identity theft" (Hendricks-Sturrup et al., 2022, p. 5)	Privacy-related negative experiences		

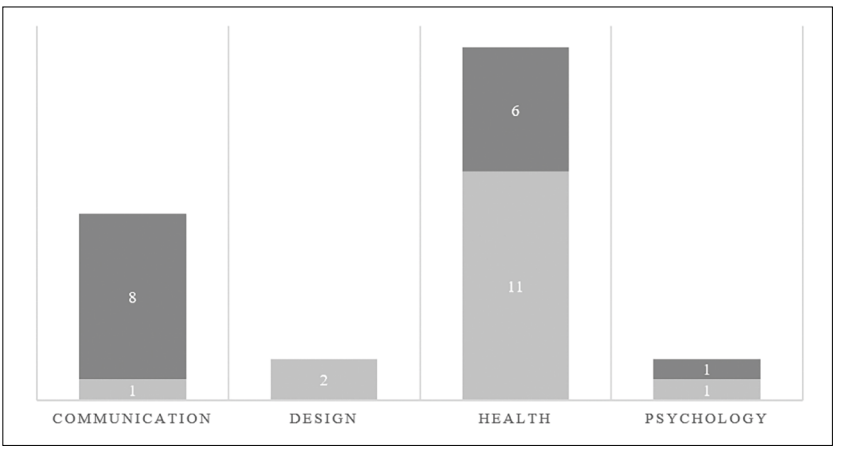
4. Results

4.1 Descriptive overview of studies' characteristics

Of the 30 publications, 26 are journal articles, three are conference papers (Diethei et al., 2021; Gomez Ortega et al., 2021; Maus et al., 2020), and one is an essay (Bietz et al., 2019). Online appendix 5 details study categorizations.

Regarding RQ1, which examines disciplinary background, methodology, and focus on hypothetical versus actual donations, health research dominates the sample (n = 17), followed by communication (n = 9), which often has a methodological focus.

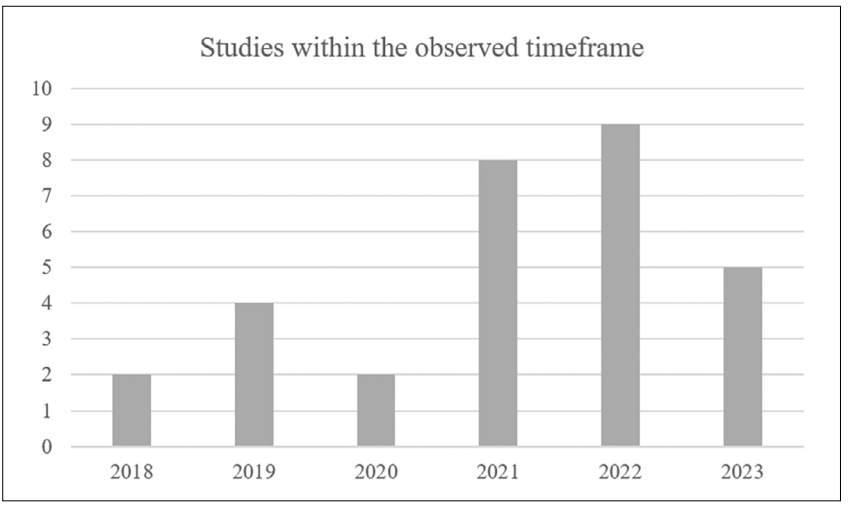
Figure 3. Disciplines of included studies



Note. The upper section of the bar in darker color indicates studies with a methodological focus.

Examining the publication years (Fig. 4), an upward trend up to 2023 can be shown. The decrease seen in 2023 is likely due to data collection ending in April.

Figure 4. Publications over time



Note. The timeframe investigated for this study lays within January 2018 to April 2023

Regarding methodological approaches (Tab. 2), surveys are most common (n = 12), followed by actual data donation studies (n = 9), interviews (n = 3), theoretical discussions (n = 3), experiments (n = 2), and focus groups (n = 1). Table 2 provides a more detailed breakdown, with the "Total" row indicating how many studies primarily used each method.

Table 2. Detailed classification of methodological approaches included studies used

	Data donations	Experiment	Qualitative interview	Survey	Theoretical work	Focus groups
+ Theoretical discussion	1					
+ Diary	1					
+ Survey	5					
–	2					
–		2				
+ Content analysis			1			
+ Participant observation			1			
–			1			
–				12		
+ Discussion group					1	
–					2	
–						1
Total	9	2	3	12	3	1

Notes. "+" indicates the method in the column was combined with the one in the row; "–" means no additional method was used. E.g., "5" in the *Data Donation* column and "+*Survey*" row means five studies primarily used data donation and also included a survey. The "2" in the *Data Donation* column and "–" row shows two studies only using data donations.

Further analysis shows that most studies (n = 18) examine hypothetical willingness to donate data. This is followed by actual donations (n = 7) and purely theoretical perspectives (n = 3) and studies investigating hypothetical scenarios along with an actual donation (n = 2).

Investigating the kinds of data the studies (hypothetically) requested (Tab. 3), we can see that health research dominates once more.

Table 3. Data types included studies investigated

Investigated data types	*N*
Health data	18
Social media	9
Finance data	4
Google searches	4
Other (e.g., 'machine data', voting history)	3
Smartphone usage (e.g., battery, screen-time)	3
Mobility (e.g., ride sharing, GPS data)	2
Musik streaming	2
Communication (e.g., mail records, text messages)	2

Note. One study can investigate more than one data type.

Concerning the donation processes (Tab. 4), the precision of questions and descriptions varies. Most studies (n = 14) do not explicitly state which donation process they assume, often just asking for willingness to donate a certain data type (e.g., Pilgrim & Bohnet-Joschko, 2022) or concerns when donating data (Hendricks-Sturrup & Lu, 2023).

Table 4. Donation methods investigated

Investigated donation method	*N*
Unspecified	14
Data upload through websites or surveys	6
Specialized tool (DDPs)	5
Through apps	5

Note. One study can investigate more than one donation process.

4.2 Respondent characteristics

Table 5 shows how many studies discussed respondent characteristics as potentially influential on willingness to donate, addressing the second research question, and showcases exemplary subcategories. Online appendix 5 shows assignments, which studies count into which category.

Table 5. Respondent characteristics influencing willingness to donate

	Emp.	Theor.	Exemplary subcategories
Sociodemographic	13	2	Age, gender, education, ethnicity, migration background, income
Technical skills	3	6	Media & technology literacy, technological affinity
Device usage	6	2	Frequency of App use, platform usage
Characteristics related to the study	5	2	Politics: Political interest & political efficacy Other: Health status, news consumption
Other characteristics	5	0	Trust, Big Five personality traits, other donation behavior
Research compliance	2	2	Participation in experiments, attitudes to surveys
Device ownership	4	0	Type of device, operating system

Notes. One study can investigate more than one factor. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions.

First, sociodemographic factors such as gender, age, and income are frequently discussed as influencing willingness to donate (n = 15). Hypothetical studies often report negative effects of age (Hillebrand et al., 2023; Karampela et al., 2019; Kaspar, 2020; Pilgrim & Bohnet-Joschko, 2022). When gender effects are significant, men tend to be more willing to donate data in both actual (e.g., Silber et al., 2022) and hypothetical contexts (Karampela et al., 2019; Pilgrim & Bohnet-Joschko, 2022).

Technical skills (n = 9) and device usage (n = 8) are also commonly identified as relevant respondent characteristics.

Study-related characteristics (n = 7) further influence participation. For example, individuals with a disease may be more willing to contribute to related research (Karampela et al., 2019). Physical activity can also matter, as individuals may hesitate to share fitness data if they perceive themselves as insufficiently active or not fitting certain health ideals (Toepoel et al., 2021, p. 4).

Finally, other personal characteristics (n = 5), such as Big Five traits, general research compliance (n = 4), and device ownership (n = 4), are also linked to willingness to donate.

4.3 Situational factors

Table 6 provides an overview of situational motivations that potentially influence willingness to donate. This is complemented by Table 7, which summarizes the findings on the attributes of the study design. Together, they answer the third research question.

Table 6. Situational motivations influencing willingness to donate

	Emp.	Theor.	Exemplary subcategories
Privacy concerns	18	4	Concerns on data storage, negative privacy related experiences
Egoism	10	3	Positive: Direct or future benefit Negative: Direct loss
Altruism	10	2	General: Desire to help others and contribute to social welfare Collectivism: Benefit for direct community Impure altruism: Indirect reciprocity Reluctant altruism: Perceived need for donation

Notes. One study can investigate more than one factor. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions

Most studies identify privacy concerns (n = 22) as a key factor influencing willingness to donate. These include concerns about data use and prior experiences with privacy breaches (e.g., fraud). Most studies find a negative effect on actual donation (e.g., Ohme et al., 2021; Pak et al., 2022) or willingness to donate (e.g., Hillebrand et al., 2023; Karampela et al., 2019). Some report effects only for specific data types or inconclusive results (Hendricks-Sturrup et al., 2022), likely due to varying operationalizations.

Egoism (n = 13) is also frequently discussed. Participation may increase when personal benefits are expected but decrease when participants anticipate risks or losses, such as concerns about consequences of sharing online searches.

Altruism (n = 12) is another key factor with several forms. These include pure altruism, the desire to contribute to the common good (Skatova & Goulding, 2019), and collectivism, the intention to support one's community (Diethei et al., 2021, p. 7). Impure altruism reflects self-interested motives, such as expectations of indirect reciprocity, where individuals believe others will also donate, enabling mutually beneficial research (Hillebrand et al., 2023; Skatova & Goulding, 2019). This aligns with findings that individuals are more willing to share data when they expect others to do so (Pilgrim & Bohnet-Joschko, 2022, p. 9). Finally, reluctant altruism occurs when individuals donate because they perceive a need and believe others will not contribute (Skatova & Goulding, 2019, p. 3).

Table 7. Attributes of the study design influencing willingness to donate

	Emp.	Theor.	Exemplary subcategories
Data types	8	4	Sensitivity of data, data granularity, perceived relevance of data
Recipient of the data	9	2	Trust in recipient of the data, attitudes towards recipient of the data, type of institution receiving data
Understanding & impact	9	1	Understanding: Perceived value of study for society or research
			Impact: Perceived impact of one's contribution, uncertainty about value of contribution
Conditions of sharing	8	1	Recipient of the data: Commercial vs. academic
			Donation process: Type of donation process, trust in process
Purpose of the study	5	2	Attitudes towards the purpose of the study
Control in the process	4	1	Autonomy and sovereignty, perceived behavioral control
Transparency	3	1	Accountability for data and research
Technical problems	1	3	
Presentation & explanation	1	2	Question positioning and posing, evaluation of survey
Effort of participation	0	3	Forgetting to participate
Informed consent	2	1	Understanding the consent, type of consent

Notes. One study can investigate more than one factor. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions

A key study design factor is the type and sensitivity of requested data (n = 12). Health data is often perceived as more sensitive than social media data, reducing willingness to donate (Silber et al., 2022). However, perceived relevance also matters, as the research value of health data may be more apparent to participants. Trust in the data recipient (n = 11) is another important factor.

Understanding and perceived impact (n = 10) also influence participation. Participants may be more willing to donate if they understand the study's purpose and believe their contribution is meaningful, highlighting the importance of clear communication.

Willingness further depends on sharing conditions (n = 9), such as whether researchers are from academic or industry contexts and how trustworthy the dona-

tion process is. The study's purpose (n = 7) also matters, as positive attitudes toward it can increase participation.

Control over the donation process (n = 5) is another factor. Allowing participants to review or delete data can enhance perceived control and encourage participation.

Less frequently discussed factors include transparency (n = 4), technical issues (n = 4), and clarity of presentation and explanation (n = 3). Effort required to participate (n = 3) may also reduce motivation or lead to non-participation. Finally, informed consent (n = 3), including clarity and potential time constraints, can influence willingness to donate.

4.4 Societal-level influences

Societal-level influences played a minor role in literature. However, those that were discussed are summarized in Table 8 in response to the fourth research question.

Table 8. Societal-level influences on willingness to donate

	Emp.	Theor.	Exemplary subcategories
Contextual sensitivity	3	3	Relevance of time of (health-) crisis, concrete concerns for the community
Country & culture	2	1	East vs. West Germany, rural vs. urban, degree of digitalization of the country

Notes. One study can investigate more than one societal-level influence. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions

We identified contextual sensitivity (n = 6) as a factor influencing willingness to donate. In particular, crises, such as pandemics, can increase participation, as heightened health concerns may shape behavior.

Country and culture (n = 3) also emerged as relevant influences. Factors such as levels of digitalization and whether individuals live in rural or urban areas may affect willingness to donate data.

4.5 Strategies

Finally, answering the fifth research question, we want to shed light on strategies that might augment willingness to donate, once more divided into general and those specifically relating to study design (Tab. 9 & 10).

Table 9. General strategies to augment willingness to donate

		n	Exemplary subcategories
Benefits	Offering benefits	5	Insight into data, visualizations
	Monetary incentives	7	Sweepstakes, direct payments
	Community	2	Establishing community for exchange or comparison
Communication	Address altruism	5	Emphasizing the societal relevance of study
	Emphasize the relevance	4	Providing information on value of contribution, emphasizing the need for donation, applicability of study results for the greater good
	Express appreciation	2	Involving participants in the research process, acknowledging participants contribution
Nudges	Reminder	3	Sending reminders
	Framing	2	Framing contribution of data as a donation, pro-social framing
	Individualizing messages	1	Addressing participants motivations
	Timing	1	Posing request after participants already shared some data, event-based requests
Manipulation	Personal recruitment	1	Utilizing the higher compliance typically observed in in-person interactions
	Hidden request	1	Intransparent requests leading to inadvertently or unknowingly agreement to participate

Note. One study can investigate more than one strategy.

To increase willingness to donate, studies may offer benefits (n = 4), appealing to egoistic motivations. These include personalized products (Gómez Ortega et al., 2021, p. 498) and monetary incentives (n = 7), though optimal incentive levels remain unclear (e.g., Silber et al., 2022). Establishing communities (n = 2) for sharing or comparing donated data has also been suggested.

Effective communication is essential. Emphasizing study relevance (n = 4) and appealing to altruism (n = 5) can encourage participation, while expressing appreciation (n = 2) may target egoistic motives.

Nudges offer additional strategies, such as reminders (n = 3) to reduce forgetfulness and framing (n = 2) to activate altruistic motivations. Other approaches include tailoring messages to individual motivations (n = 1) and timing requests strategically (n = 1), for example, when participants are already sharing data.

Some studies also discuss manipulative techniques, such as personal recruitment (n = 1) or hidden requests (n = 1), though these raise ethical concerns (Pilgrim & Bohnet-Joschko, 2022, p. 8).

Finally, strategies related to study design (Tab. 10) are closely linked to the identified design attributes influencing willingness to donate (Tab. 7).

Table 10. Strategies relating to study design attributes to augment willingness to donate

	n	Exemplary subcategories
Cultivating trust	14	Collaborations with trustworthy partners, no commercial use of data
Privacy protection & emphasis	10	Local processing, data minimizing, anonymizing, respecting legal guidelines
Reducing burden & effort	9	Execution: Offering instructions and technological support
		Preparations: Embedded requests, simplifying processes & interfaces
Autonomy & control	8	Data types: Offering participants the possibility to filter and delete parts of their data
		Other: Usage of familiar devices, possibility to drop out
Informed consent	7	Intelligibility of consent: Participants should understand consent forms
		Appropriateness of consent: Transparent and potentially selective consent
Data types	5	Asking for less granular and sensitive data

Notes. One study can investigate more than one strategy. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions

First, cultivating trust (n = 14) is widely recommended, for example, through collaboration with trusted partners (Diethei et al., 2021, p. 11) or involvement of ethics boards (Boeschoten et al., 2022, p. 409). Trust improves perceptions of data recipients and transparency and can mitigate privacy concerns (Juga et al., 2021).

Given the prevalence of privacy concerns, strong privacy protection and clear communication of safeguards are essential (n = 10). Aligning measures with participants' expectations can increase trust and willingness to donate.

Reducing effort and burden (n = 9) is also important to prevent dropouts. This includes simplifying procedures and offering technical support to address limited skills.

Increasing autonomy and control (n = 8) is another key strategy. Allowing participants to review, select, or delete data can enhance control, awareness, and digital literacy (Buhr et al., 2022, p. 10; Gómez Ortega et al., 2021, p. 497). However, greater perceived control may improve attitudes toward sharing while reducing willingness to share (Juga et al., 2021, p. 53). Clear and appropriate informed consent is also recommended (n = 7).

Finally, requesting appropriate data types (n = 5), particularly less sensitive data, can increase willingness to donate. While research needs may limit this, data

minimization should be considered (Hakobyan et al., 2025; Ohme & Araujo, 2022, p. 1).

5. Discussion

The aim of this scoping review was to examine research on participation in data donation studies and factors influencing willingness to donate.

First, we analyzed what characterizes the literature on willingness to donate. To date, the number of studies is relatively small, with more than half of the studies under investigation originating from a health context. This is important, as concerns about data donation may vary across fields. In health contexts, data is often perceived as particularly sensitive, but its potential value and benefits for others (Kulzer et al., 2022) may also be more salient.

Most studies examine hypothetical willingness to donate and use a variety of measurement approaches. This reliance on hypothetical designs is a limitation, as such studies may overestimate actual participation rates. However, since this study was conducted, more research has begun to include actual data donations (e.g., Keusch, Pankowska, et al., 2024; Strycharz et al., 2024). While surveys provide insight into attitudes, stated intentions do not always translate into behavior. Research on the privacy paradox (e.g., Norberg & Horne, 2007) highlights this gap. For example, individuals often report concerns about sharing data with companies but still do so for minor benefits. In the context of data donation, the opposite pattern may occur: Individuals may express support for donating data for research but refrain from doing so, possibly due to the absence of immediate personal benefits.

When measuring hypothetical willingness to donate, wording is crucial. A lack of specificity about the donation process can influence responses. Some studies ask about willingness without clarifying how or under what conditions data would be donated, which is problematic given that study attributes affect decisions. For example, the method of data collection, such as via a platform or questionnaire, can lead to different levels of willingness (Toepoel et al., 2021, p. 8).

Second, we examined respondent characteristics. Studies most often focus on sociodemographic factors, with men generally more likely to donate, as well as technical skills and device usage. Study-related characteristics are also relevant. For instance, individuals may choose not to donate fitness data due to low activity levels (Toepoel et al., 2021, p. 4), which may introduce social desirability bias in participation decisions rather than in the data itself.

Regarding the third research question on situational factors, privacy concerns, egoism, and altruism emerge as key motivations. However, general altruism alone has limited explanatory power, as other mechanisms and forms of prosocial behavior also shape participation (Skatova & Goulding, 2019, p. 12).

Altruism can take different forms, including impure, reluctant, and collectivist altruism. The latter is particularly relevant, as individuals may struggle to see how their contribution benefits broader society. Framing donations as supporting an immediate community can make participation more tangible. For example, indi-

viduals who have lost someone to suicide may better recognize the value of their contribution when it supports prevention research (Sleigh, 2018, p. 4).

Regarding study design attributes, the literature highlights data types, the data recipient, and participants' understanding of the study and its impact. These can be interpreted through the Theory of Contextual Integrity (Nissenbaum, 2010), which conceptualizes privacy as appropriate information flow based on context-specific norms. These norms involve actors (e.g., recipients), attributes (e.g., data types), and transmission principles (e.g., conditions of sharing). Willingness to donate may therefore depend on whether study attributes align with participants' expectations of appropriate data flows, which can also foster trust.

Our findings suggest several practical recommendations for designing data donation studies. Researchers should follow established best practices to ensure ethical and methodological standards (Carrière et al., 2024; Ohme & Araujo, 2022). At the same time, they must carefully select and prepare requested data while considering structural constraints, particularly those related to DDPs, whose content and format are controlled by platforms (Hase et al., 2024).

To address technical difficulties and reduce effort, lowering barriers to participation should be a priority. This includes clear instructions and support services, such as help desks, to guide participants.

Aligning data requests with participant preferences, for example, by avoiding highly sensitive data, can build trust and support data minimization. Whitelists can help achieve this (Hakobyan et al., 2025; Ohme & Araujo, 2022). Allowing participants to remove parts of their data can further increase perceived control and autonomy.

If the study design cannot be adapted due to specific data requirements, communication becomes even more important. Transparency about data handling can help address privacy concerns, although it may also increase their salience (Marreiros et al., 2017, p. 10). Communicating the non-commercial nature of the research and adhering to ethical standards can strengthen trust. Emphasizing the societal relevance of the study and the impact of individual contributions can further increase willingness to participate. Offering meaningful incentives, such as personal data insights, may also enhance motivation.

Furthermore, nudges may be an effective strategy to increase willingness to donate data, for example, through reminders or careful timing of requests. In contrast, manipulative strategies such as hidden requests can be considered forms of "sludge," meaning illegitimate interventions that provide no benefit to individuals or society (Thaler, 2018, p. 431). Such practices should be avoided, as they may undermine autonomy and comfort, erode trust in research, and reduce willingness to participate in the future. Study design strategies should therefore prioritize transparency, non-manipulation, and responsible use of behavioral interventions.

Altogether, strategies to increase willingness to donate require further investigation. While they cannot address all biases inherent in data donations, such as sampling bias from hard-to-reach populations, they can improve participation among accessible groups. If the goal is to increase response rates within these groups, the strategies and factors outlined here should be considered. However,

there is no universal solution, and not all ex-ante strategies are effective in every context (Hase & Haim, 2024; Keusch, Pankowska, et al., 2024). Complementary approaches, such as retrospective statistical correction, may therefore be valuable (Pak et al., 2022).

It remains important to identify which strategies are effective in specific contexts to encourage donations. This is particularly relevant in communication research, where data donations enable the study of digital behavior and key societal issues, such as algorithmic personalization and online participation, that are difficult to examine accurately with other methods.

Future research requires a more consistent understanding of variables such as technical skills and willingness to donate. However, these measures must remain context-sensitive. For example, the skills needed for screenshot donations differ from those required for handling DDPs, so measurement tools should reflect these specific demands. Developing comparable measures for each task would improve consistency across studies.

Furthermore, privacy concerns also vary by data type and sensitivity. Concerns about private chat data in DDPs may center on potential data leaks, whereas concerns about social media data may relate to repercussions for expressed opinions.

This study has several limitations. First, data donation is a relatively new research method, and the field is evolving rapidly. Since this study was conducted, additional literature has emerged, particularly in communication research (e.g., de León et al., 2025; Haim & Hase, 2024; Keusch, Pankowska, et al., 2024; Silber et al., 2024; Strycharz et al., 2024). The review should therefore be seen as a snapshot of the field at a specific point in time. As research continues to grow, updating this work or conducting full reviews may prove valuable, as demonstrated by the review conducted by Xiong et al. (2025).

Methodologically, data charting, extraction, and analysis were conducted by only one of the authors, despite recommendations for multiple coders to enhance reliability (Booth et al., 2022; Levac et al., 2010). This was mitigated in part by consulting an expert on suggestions for literature at key stages. The review also did not include forward citation searching. Although backward reference screening was conducted, future reviews should combine both approaches to improve coverage.

Additionally, Google Scholar screening was limited to the first five pages. While relevance declined after the first three pages, applying a fixed cutoff risks omitting relevant studies. Future reviews should consider more flexible stopping criteria based on relevance and saturation.

Second, the relatively small number of included studies (n = 30) is a limitation, resulting from the strict definition of data donation. This narrow scope was necessary, as data collection and sharing procedures are central to understanding participation. Given the limited literature, five studies with unclear procedures were included if they plausibly involved data donation and contributed to understanding willingness to donate.

Finally, most studies originate from a health context, which may limit generalizability. In health research, individuals tend to be more willing to donate data (e.g., Buhr et al., 2019) and may view it as a social duty (Skatova & Goulding,

2019). This may not apply to other contexts, such as information-seeking research (Blassnig et al., 2023), where willingness may differ. Findings should therefore be interpreted as context-dependent.

At the same time, the interdisciplinary nature of data donation research offers opportunities. Communication research in this area is still emerging and depends on recent developments in data access. Insights from other fields, where data sharing has been studied longer, can provide valuable input, encourage new perspectives, and support reflection on how findings apply to communication research.

Differences in definitions across disciplines require scrutiny to ensure a shared understanding, especially in cross-disciplinary collaborations. Establishing common terminology and methodological standards is essential for effective cooperation. When this is achieved, such collaborations can be highly productive. For example, social media data can be relevant not only to communication research but also to health research on mental well-being (Sleigh et al., 2018).

6. Conclusion

In conclusion, this study provides an overview of factors influencing willingness to donate and offers insights into strategies that may increase participation in data donation studies. It highlights participation bias in this context and presents actionable recommendations for study design.

To encourage participation, researchers should prioritize trust, transparency, and ethical standards such as data minimization. While not all studies can fully align with participant preferences, reducing effort and providing support remain important. Additionally, nudges and improved communication can be effective. In communication research, emphasizing societal relevance and individual benefits may further enhance engagement.

Considering these aspects can help make data donation studies more participant-centered, ethically sound, and impactful.

Nonetheless, beyond participation bias, other biases related to trace data must also be considered (Amaya et al., 2020; Keusch et al., 2023). Further research is needed to assess these limitations. Researchers should carefully select methods, samples, and measurements, particularly during study design, to minimize bias.

Overall, data donations offer strong potential for communication research by leveraging digital trace data to better understand societal phenomena and behavior, while helping address limitations of traditional methods. To fully realize this potential, future research should continue to examine participation motivations to improve methodology, mobilization, and representativeness.

Generative AI declaration

During the preparation of this manuscript, we used the following AI tools: Deepl Write was used for language editing and enhancing the clarity and flow of the drafted text. All AI-generated outputs were carefully reviewed and edited by the authors.

Disclosure statement

All authors declare that they have no conflicts of interest.

Funding

This research project is funded by the Bavarian Research Institute for Digital Transformation (bidt), an institute of the Bavarian Academy of Sciences and Humanities.

Supplementary material

Supplementary material for this article is available on the Open Science Framework (OSF): https://osf.io/rczpg/overview?view_only=c769a960b0914d30a3d16eb92872053e.

References

Amaya, A., Biemer, P. P., & Kinyon, D. (2020). Total error in a big data world: Adapting the TSE framework to big data. *Journal of Survey Statistics and Methodology*, *8*(1), 89–119. https://doi.org/10.1093/jssam/smz056

Android.com. (2018, September 25). *Technology for today's world: Helping you reclaim a sense of balance*. Retrieved from https://blog.google/products/android/technology-todays-world-helping-you-reclaim-sense-balance/

Apple Inc. (2018, September 25). *iOS 12 führt neue Funktionen zum Vermeiden von Unterbrechungen und zur Verwaltung der Screen Time ein* [iOS 12 introduces new function to avoid interruptions and managing screen time]. Retrieved from https://www.apple.com/de/newsroom/2018/06/ios-12-introduces-new-features-to-reduce-interruptions-and-manage-screen-time/

Aquisti, A. (2004, May 17). Privacy in electronic commerce and the economics of immediate gratification. *Proceedings of the 5th ACM Conference on Electronic Commerce*, New York, NY, USA. https://doi.org/10.1145/988772.988777

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures*, *11*(3), 173–190. https://doi.org/10.1080/19312458.2017.1317337

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Baumgartner, S. E., Sumter, S. R., Petkevič, V., & Wiradhany, W. (2023). A novel iOS data donation approach: Automatic processing, compliance, and reactivity in a longitudinal study. *Social Science Computer Review*, *41*(4), 1456–1472. https://doi.org/10.1177/08944393211071068

Bietz, M., Patrick, K., & Bloss, C. (2019). Data donation as a model for citizen science health research. *Citizen Science: Theory and Practice*, *4*(1). https://doi.org/10.5334/cstp.178

Blassnig, S., Mitova, E., Pfiffner, N., & Reiss, M. V. (2023). Googling referendum campaigns: Analyzing online search patterns regarding Swiss direct-democratic votes. *Media and Communication*, *11*(1), 19–30. https://doi.org/10.17645/mac.v11i1.6030

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

Booth, A., Sutton, A., Clowes, M., & Martyn-St James, M. (2022). *Systematic approaches to a successful literature review* (Third edition). SAGE.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

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*, *26*(14), 2649–2668. https://doi.org/10.1080/1369118X.2022.2097015

Buhr, L., Schicktanz, S., & Nordmeyer, E. (2022). Attitudes toward mobile apps for pandemic research among smartphone users in Germany: National survey. *JMIR MHealth and UHealth*, *10*(1), 1–17. https://doi.org/10.2196/31857

Carrière, T. C., Boeschoten, L., Struminskaya, B., Janssen, H. L., Schipper, N. C. de, & Araujo, T. (2025). Best practices for studies using digital data donation. *Quality & Quantity*, *59*(Suppl 1), 389–412. https://doi.org/10.1007/s11135-024-01983-x

de León, E., Votta, F., Jongerius, S., Mulder, J., Struminskaya, B., Araujo, T., & de Vreese, C. (2025). Can tailored recruitment messaging increase digital trace data donation compliance? *International Journal of Social Research Methodology*, 1–17. https://doi.org/10.1080/13645579.2025.2568497

Diethei, D., Niess, J., Stellmacher, C., Stefanidi, E., & Schöning, J. (2021). Sharing heartbeats: Motivations of citizen scientists in times of crises. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, & S. Drucker (Eds.), *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). ACM. https://doi.org/10.1145/3411764.3445665

Dinev, T., McConnell, A. R., & Smith, H. J. (2015). Research commentary—Informing privacy research through information systems, psychology, and behavioral economics: Thinking outside the "APCO" box. *Information Systems Research*, *26*(4), 639–655. https://doi.org/10.1287/isre.2015.0600

European Parliament and the Council of the European Union. (2016, April 27). *On the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General data protection regulation)*. Retrieved from https://eur-lex.europa.eu/eli/reg/2016/679/oj

Gil-López, T., Christner, C., León, E. de, Makhortykh, M., Urman, A., Maier, M., & Adam, S. (2023). Do (not!) track me: Relationship between willingness to participate and sample composition in online information behavior tracking research. *Social Science Computer Review*, *41*(6), 2274–2292. https://doi.org/10.1177/08944393231156634

Gomez Ortega, A., Bourgeois, J., & Kortuem, G. (2021). Towards designerly data donation. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (pp. 496–501). ACM. https://doi.org/10.1145/3460418.3479362#

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, *56*, 475–495. https://doi.org/10.1086/269338

Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the filter bubble? *Digital Journalism*, *6*(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Haim, M., & Hase, V. (2023). Computational Methods und Tools für die Erhebung und Auswertung von Social-Media-Daten [Computational methods and tools for social media data collection and analysis]. In S. Stollfuß, L. Niebling, & F. Raczkowski (Eds.), *Handbuch Digitale Medien und Methoden* (pp. 1–20). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36629-2_41-1

Hakobyan, O., Hillmann, P.-J., Martin, F., Böttinger, E., & Drimalla, H. (2025). Development and evaluation of Dona, a privacy-preserving donation platform for messaging data from WhatsApp, Facebook, and Instagram. *Behavior Research Methods*, *57*(3), 94. https://doi.org/10.3758/s13428-024-02593-z

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), 1–29. https://doi.org/10.5117/CCR2024.2.2.HASE

Hendricks-Sturrup, R., & Lu, C. Y. (2023). An assessment of perspectives and concerns among research participants of childbearing age regarding the health-relatedness of data, online data privacy, and donating data to researchers: Survey study. *Journal of Medical Internet Research*, *25*. https://doi.org/10.2196/41937

Hendricks-Sturrup, R., Zhang, F., & Lu, C. Y. (2022). A survey of research participants' privacy-related experiences and willingness to share real-world data with researchers. *Journal of Personalized Medicine*, *12*(11). https://doi.org/10.3390/jpm12111922

Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. *Information and Organization*, *33*(1). https://doi.org/10.1016/j.infoandorg.2023.100452

Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems*, *12*(12), 767–797. https://doi.org/10.17705/1jais.00282

Juga, J., Juntunen, J., & Koivumäki, T. (2021). Willingness to share personal health information: Impact of attitudes, trust and control. *Records Management Journal*, *31*(1), 48–59. https://doi.org/10.1108/RMJ-02-2020-0005

Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Connected Health user willingness to share personal health data: Questionnaire study. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/14537

Kaspar, K. (2020). Motivations COVID-19 for social distancing and app use as complementary measures to combat the pandemic: Quantitative survey study. *Journal of Medical Internet Research*, *22*(8). https://doi.org/10.2196/21613

Kellermeyer, L., Harnke, B., & Knight, S. (2018). Covidence and Rayyan. *Journal of the Medical Library Association*, *106*(4). https://doi.org/10.5195/jmla.2018.513

Keusch, F. (2015). Why do people participate in web surveys? Applying survey participation theory to internet survey data collection. *Management Review Quarterly*, *65*(3), 183–216. https://doi.org/10.1007/s11301-014-0111-y

Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2023). Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a German national survey. *Sociological Methods & Research*, *52*(2), 841–878. https://doi.org/10.1177/0049124120914924

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 Face-

book data donation. *Field Methods*, *36*(4), 279–293. https://doi.org/10.1177/1525822X231225907

Keusch, F., Struminskaya, B., Mulder, J., & Jongerius, S. (2024). *Measuring physical activity in older adults through data donation: Consent rates, donation success, & bias*. https://doi.org/10.31235/osf.io/fdg8t

Kmetty, Z., Stefkovics, Á., Számely, J., Deng, D., Kellner, A., Pauló, E., Omodei, E., & Koltai, J. (2024). Determinants of willingness to donate data from social media platforms. *Information, Communication & Society*, *28*(7), 1324–1349. https://doi.org/10.1080/1369118X.2024.2340995

Kulzer, B., Heinemann, L., & Roos, T. (2022). Patients' experience of new technologies and digitalization in diabetes care in Germany. *Journal of Diabetes Science and Technology*, *16*(6), 1521–1531. https://doi.org/10.1177/19322968211041377

Levac, D., Colquhoun, H., & O'Brien, K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(69). https://doi.org/10.1186/1748-5908-5-69

Marreiros, H., Tonin, M., Vlassopoulos, M., & Schraefel, M. C. (2017). "Now that you mention it": A survey experiment on information, inattention and online privacy. *Journal of Economic Behavior & Organization*, *140*, 1–17. https://doi.org/10.1016/j.jebo.2017.03.024

Maus, B., Salvi, D., & Olsson, C. M. (2020). Enhancing citizens trust in technologies for data donation in clinical research: Validation of a design prototype. *Companion Proceedings of the 10th International Conference on the Internet of Things (IoT 2020)*. Advance online publication. https://doi.org/10.1145/3423423.3423430

Mezinska, S., Kaleja, J., Mileiko, I., Santare, D., Rovite, V., & Tzivian, L. (2020). Public awareness of and attitudes towards research biobanks in Latvia. *BMC Medical Ethics*, *21*(1). https://doi.org/10.1186/s12910-020-00506-1

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1). https://doi.org/10.1186/2046-4053-4-1

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1). https://doi.org/10.1186/s12874-018-0611-x

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press. https://doi.org/10.1515/9780804772891

Norberg, P., & Horne, D. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Ohme, J., & Araujo, T. (2022). Digital data donations: A quest for best practices. *Patterns*, *3*(4). https://doi.org/10.1016/j.patter.2022.100467

Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2023). 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

Ohme, J., Araujo, T., de Vreese, C., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, *9*(2), 293–313. https://doi.org/10.1177/2050157920959106

Pak, C., Cotter, K., & Thorson, K. (2022). Correcting sample selection bias of historical digital trace data: Inverse probability weighting (IPW) and type II Tobit model. *Com-*

munication Methods and Measures, *16*(2), 134–155. https://doi.org/10.1080/19312458.2022.2037537

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis*, *18*(10), 2119–2126. https://doi.org/10.11124/JBIES-20-00167

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

Pilgrim, K., & Bohnet-Joschko, S. (2022). Effectiveness of digital forced-choice nudges for voluntary data donation by health self-trackers in Germany: Web-based experiment. *Journal of Medical Internet Research*, *24*(2), 1–13. https://doi.org/10.2196/31363

Prainsack, B. (2019). Data donation: How to resist the iLeviathan. In J. Krutzinna & L. Floridi (Eds.), *The ethics of medical data donation* (pp. 9–22). Springer. https://doi.org/10.1007/978-3-030-04363-6_2

Scharkow, M. (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

Seltzer, E., Goldshear, J., Guntuku, S. C., Grande, D., Asch, D. A., Klinger, E. V., & Merchant, R. M. (2019). Patients' willingness to share digital health and non-health data for research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, *19*(1). https://doi.org/10.1186/s12911-019-0886-9

Severin-Nielsen, M. K. (2023). Politicians' social media usage in a hybrid media environment: A scoping review of the literature between 2008–2022. *Nordicom Review*, *44*(2), 172–193. https://doi.org/10.2478/nor-2023-0010

Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P., Stier, S., & Weiß, B. (2022). Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behaviour. *Journal of the Royal Statistical Society Series a: Statistics in Society*, *185*(Supplement_2), S387–S407. https://doi.org/10.1111/rssa.12954

Silber, H., Breuer, J., Felderer, B., Gerdon, F., Stammann, P., Daikeler, J., Keusch, F., & Weiß, B. (2024). Asking for traces: A vignette study on acceptability norms and personal willingness to donate digital trace data. *Social Science Computer Review*. Advance online publication. https://doi.org/10.1177/08944393241305776

Skatova, A., & Goulding, J. (2019). Psychology of personal data donation. *PloS One*, *14*(11), 1–20. https://doi.org/10.1371/journal.pone.0224240

Sleigh, J. (2018). Experiences of donating personal data to mental health research: An explorative anthropological study. *Biomedical Informatics Insights*, *10*. https://doi.org/10.1177/1178222618785131

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*, *38*(5), 503–516. https://doi.org/10.1177/0894439319843669

Strycharz, J., Meppelink, C., Zarouali, B., Araujo, T., & Voorveld, H. (2024). The blind spot in data donations: Who is (not) willing to donate digital data in social scientific research. *Computational Communication Research*, *6*(2), 1–27. https://doi.org/10.5117/CCR2024.2.3.STRY

Susha, I., Grönlund, Å., & van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly*, *36*(1), 112–128. https://doi.org/10.1016/j.giq.2018.11.002

Thaler, R. H. (2018). Nudge, not sludge. *Science*, *361*(6401), 431. https://doi.org/10.1126/science.aau9241

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale Univ. Press.

Toepoel, V., Luiten, A., & Zandvliet, R. (2021). Response, willingness, and data donation in a study on accelerometer possession in the general population. *Survey Practice*, *14*(1), 1–19. https://doi.org/10.29115/SP-2021-0005

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). Prisma extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

van Driel, I. I., Giachanou, A., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of social media data donations. *Communication Methods and Measures*, *16*(4), 266–282. https://doi.org/10.1080/19312458.2022.2109608

Wedel, L., Ohme, J., & Araujo, T. (2024). Augmenting data download packages – Integrating data donations, video metadata, and the multimodal nature of audio-visual content. Advance online publication. https://doi.org/10.12758/mda.2024.08

Welbers, K., Loecherbach, F., Lin, Z., & Trilling, D. (2024). Anything you would like to share: Evaluating a data donation application in a survey and field study. *Computational Communication Research*, *6*(2), 1–28. https://doi.org/10.5117/CCR2024.2.5.WELB

Xiong, Y., van der Wal, A., & Beyens, I. (2025). Improving participation in data donation studies: A systematic review of factors driving participation and evidence-informed best practices. *Social Science Computer Review*. Advance online publication. https://doi.org/10.1177/08944393251395958

Zannettou, S., Nemeth, O.-N., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2023, January 12). Analyzing user engagement with TikTok's short format video recommendations using data donations. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery*, New York, NY, USA (pp. 1–16). https://doi.org/10.1145/3613904.3642433