Escaping the Walled Garden? User Perspectives of Control in Data Portability for Social Media

JACK JAMIESON, NTT Communication Science Labs, Japan
NAOMI YAMASHITA, NTT Communication Science Labs & Kyoto University, Japan

Data portability–the capability to transfer one's data from one platform to another–has been described as an important tool for giving individuals more control over their data. It is defined in significant regulations such as the GDPR, and implemented by major online platforms. Unfortunately, there is a lack of research investigating internet users' perceptions and expectations of this technology in specific contexts, which is vital for building effective designs. One particularly important context for studying user perspectives is social media, since it is deeply embedded into daily life and is particularly complex regarding the value and portability of user data. This paper addresses that gap through a survey and interviews of social media users in the United States. We identify current attitudes and practices toward controlling their social media data, and examine participants' impressions about the extent to which data portability may enhance their control. Participants had generally favorable impressions, but had differing opinions about what forms of control are important and the extent to which those could be served by data portability. Based on the results, we propose future directions for improving users' control in the context of social media, such as fine-tuned filtering of data to be transferred and ways to coordinate transfers alongside social contacts.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; **Social media**.

Additional Key Words and Phrases: social media data, user empowerment, personal data, data protection, digital rights, social networking sites, user control, GDPR, CCPA, privacy

ACM Reference Format:
Jack Jamieson and Naomi Yamashita. 2023. Escaping the Walled Garden? User Perspectives of Control in Data Portability for Social Media. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 339 (October 2023), 27 pages. https://doi.org/10.1145/3610188

1 INTRODUCTION

Data portability is the capability to transfer one's data from one platform or service to another. It is mandated by regulations such as Europe's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), and similar regulations are in development or have been passed in several other regions [21, 63, 63, 64, 68]. Data portability has been implemented by major platforms, including a collaborative project by Apple, Meta, Google, Microsoft, and Twitter. These implementations serve to comply with regulations, but are also available to internet users even in regions without a legal requirement for data portability. Both regulators and implementers describe data portability as a means of giving control and choice to individual internet users, specifically by preventing individuals' data from being confined inside "walled-gardens" [66].

Authors' addresses: Jack Jamieson, jack@jackjamieson.net, NTT Communication Science Labs, Kyoto, Japan; Naomi Yamashita, naomiy@acm.org, NTT Communication Science Labs & Kyoto University, Kyoto, Japan.

2573-0142/2023/10-ART339 $15.00
https://doi.org/10.1145/3610188

In light of those claims and the increasing growth of data portability regulations and implementations, existing studies have looked into familiarity and general impressions regarding data portability regulations, and found that people are generally less aware of and enthusiastic about data portability compared to other data protection rights [19, 51, 72, 81, 81]. However, there is a lack of research investigating detailed perspectives in specific contexts, such as social media. Social media is a particularly important area for pursuing users' empowerment, since it is intensely embedded into daily life. A majority of United States internet users are concerned that they have little control over how companies use their data [4], and believe that big social media platforms have too much power and should be more regulated [92]. Data portability is a potential way to address this problem, however, the properties of social media may complicate this. For example, even though user data is extremely valuable to platforms, much of the value of social media for users is about interaction, rather than data itself. This was illustrated in Lindley et al.'s 2013 study [56] where they interviewed internet users who expressed that social media data was ephemeral and not worth preserving. Another study in 2015 reported that Facebook users had a variety of opinions about whether they would want to keep their data if they moved to a new platform, ranging from wanting to keep everything, to wanting to keep only certain content, to wanting to keep nothing [61]. In sum, perspectives of the value and desired portability of social media data seem to be varied and complex.

Altogether, even though data portability regulators and implementers emphasize its potential for enhancing users' control over their online data, we lack a good understanding of whether and how users perceive data portability might address aspects of control they regard as important. This study aims to address that research gap by examining how internet users in the United States view data portability in the context of social media, including their impressions of how current and future data portability systems may give them control over their online data.

To investigate user perceptions of data portability in the context of social media, We conducted a survey of a representative sample of USA residents (N = 295) as well as follow-up interviews (N = 26). Survey respondents were asked about their current social media use and attitudes toward their social media data. They were then presented with illustrative examples of data portability systems in order to stimulate reflection and elicit their impressions of current implementations of data portability and of its potential to enhance people's control of their online data in the future. During the follow-up interviews, participants expressed themselves in greater detail, such as explaining the degree to which they currently feel in control of their online data, defining what control means to them, and further describing the extent to which data portability addresses problems that are important to them.

We structured our analysis of the survey and interview data using the following research questions:

- RQ1: What are participants' current perceptions and practices around controlling their social media data?
- RQ2: To what extent do participants' perceptions of current data portability systems intersect with the topic of controlling their social media data?
- RQ3: How do participants feel about the potential of data portability to enhance individuals' control over online data in the future?

We found that most participants described feeling a lack of control over their social media data. The main method through which they exerted control over their data was by being selective about what they posted in order to manage how other people interacted with their content. Most participants had favorable impressions of the data portability systems presented in the survey, and we identified several factors that shaped those impressions. Additionally, most agreed that data

portability would increase their control over their online data, but qualitative data revealed that the anticipated scope of this increased control tended to be small. Throughout the results, we identified multiple dimensions of control that could be impacted by data portability, such as protecting one's privacy, preserving access to one's data, and managing contact with social ties.

The results of this research provide an in-depth explanation of what aspects of controlling their data social media users are most concerned with and the extent to which they believe data portability can facilitate such control. We observed that users' desired forms of control were varied, which suggests that meaningful data portability designs must be tailored to specific contexts. Based on these results, we propose directions for future data portability designs to more substantially meet users' wants, including fine-tuned selection of data to be transferred and approaches for group coordination of transfers, in order to join new platforms alongside one's social media contacts.

This paper's contributions are timely because current implementations of data portability may define de facto norms that influence later generations. And further, other cases of GDPR-motivated design, such as cookie consent banners, have resulted in designs that are technically compliant (or sometimes not even that), but which undermine the spirit of regulations and provide poor user experience [33]. Thus, this study addressed a pressing need for a more substantial understanding of users' impressions and expectations of data portability, which are vital for developing systems that meaningfully achieve their potential for user empowerment.

2 BACKGROUND

2.1 Control and walled gardens in social media

Power and control are perennial themes in writing about social media, including accounts that social media empowers users [9, 14, 79] and critical perspectives emphasizing that social media platforms exert power over users [28, 80]. Of particular salience for data portability is the fact that the platform business model is centered around *walled gardens*, in which data contributed by users is captured as a valuable resource [69]. This model has contributed to a common view among social media users that trading their data for an online service is simply "how the world works" [25, p. 5]. Nonetheless, survey research has identified that most internet users feel they lack control over their online data [4] and are concerned about how their data is used online [4, 15, 35].

The Cambridge Analytica scandal [48] has provided an opportunity to observe how internet users responded to a breach of control over their online data. One study found that people with awareness about what happened between Facebook and Cambridge Analytica had heightened desires for data portability compared to others [78]. Other research has identified that many users responded by opting out. For example, a 2018 survey [20] reported that 38% of U.S. social media users had deleted at least one social media account because they did not trust it to treat their personal information properly, and another study found that 26% of U.S. Facebook users deleted the app from their phone in 2018 [71]. However, many see opting out as unattractive, particularly due to the social costs of missing out on social media [7, 25].

Switching to a different social media service presents an alternative to opting out. In the past five years, an increasing share of users (though still a minority) have begun to migrate toward alternatives such as radical free speech platforms [75, 96], federated web systems such as Mastodon [99], and other systems that present new structures for user interaction [59]. Many, though not all, of these alternatives are designed to increase users' control, such as by avoiding walled-garden style business models [32]. Even in the absence of design features focused on giving control to users, the ability to switch to a competitor is a potential avenue for enhancing individuals' power.

Many internet users already have accounts on multiple social media sites, which may make switching easier to some extent. However, different sites tend to serve different purposes and are

used for different aspects of self-presentation [85]. Further, social media users often endeavor to manage boundaries between different groups of social contacts [82], particularly when engaging with sensitive or potentially stigmatized topics [62, 97]. For example, among people who experience identity transitions, data from the past may be unwelcome reminders of a prior identity or difficult times [37] or may represent transitory periods of identity exploration that are not meant to be preserved [36]. Thus, as we consider the possibility that switching to a new site could be a means of exerting control over one's data, we recognize that this may complicate other forms of control over one's self-presentation and boundaries, and in some cases people actively want to detach from, rather than maintain access to their past data.

2.2 Data Portability

The most prominent example of data portability is defined in article 20 of Europe's General Data Protection Regulation (GDPR) [16] which describes that individuals have the following rights:

- To receive data that has been provided to a data controller (e.g., an online platform) "in a structured, commonly used and machine-readable format."
- "To transmit those data to another controller without hindrance from the controller to which the personal data have been provided."
- "To have the personal data transmitted directly from one controller to another, where technically feasible."

The third clause is significant. Providing access to a structured download is simpler than facilitating a direct transfer between servers. However, a direct transfer may be beneficial if the quantity of data is too large for users to download [70], which may be particularly important for being inclusive of regions and individuals with infrastructure constraints [73].

Similar data portability regulations exist or are under consideration in other regions such as Australia [21], Brazil [64], California [63], Canada [68], and Japan [63]. Although California's CCPA provides a clear requirement for data portability within that state, there is no nationwide regulation within the United States [66]. Data portability has been implemented in major platforms (e.g., Google TakeOut, Facebook's Transfer your information) and in the Data Transfer Project (an open-source collaboration among Apple, Meta, Google, Microsoft, and Twitter). Through these implementations, data portability tools are available even in regions that are not subject to data portability regulation.

A report published by the European Commission explains that "the primary aim of data portability is enhancing individual's control over their personal data and making sure they play an active role in the data ecosystem" and that, "By affirming individuals' personal rights and control over the personal data concerning them, data portability also represents an opportunity to 're-balance' the relationship between data subjects and data controllers" [70, p. 4]. According to this report, empowerment arises from "preventing 'lock-in"' and promoting "controlled and limited sharing by users of personal data between organisations and thus enrich[ing] services and customer experiences" [70, p. 5]. Similarly, in a 2020 submission to the U.S. Federal Trade Commission, Facebook asserted that they support the principle of data portability because it "gives people control and choice, while also promoting innovation" [22].

Some scholars have expressed skepticism towards these claims. Fukuyama [29] asserted that data portability is "not a way of addressing the political threat that platform power poses" because the most important data held by platforms is not the data directly provided by users, but instead metadata about users' interactions, which is not subject to the right to data portability. Additionally, Lambert argued that the rapid growth of platforms such as TikTok, which do not import user data from other platforms, demonstrates that a lack of data portability is not a barrier to using

new platforms [52]. Accordingly, there is a need for empirical research to understand how data portability might impact users' behavior and experiences.

Much of the empirical research about data portability has involved making data portability requests and evaluating the extent to which data controllers comply with regulatory requirements [e.g., 6, 51, 84, 86, 95, 100]. In general, this research has identified middling compliance, such as a study identifying that 71.4% of providers failed at least one requirement [84]. The same study found the ability to import data was particularly lacking, with no providers directly offering to import data obtained through a right to data portability request (exported from Facebook), and only 23.2% offering at least minimal import capability. Overall, this research shows that support for data portability is somewhat underdeveloped, especially regarding importing.

Several studies have investigated awareness of GDPR rights among European residents, and have consistently found data portability to be the least known and exercised [19, 72, 81]. A panel survey in the Netherlands [81] found that respondents had lower intentions to use data portability than they did toward any other GDPR right, and viewed it as the least effective GDPR right for protecting internet users' privacy. Similarly, Kuebler-Wachendorff et. al [51] surveyed German internet users and found that less than one-third of participants had heard of the right to data portability, and that it was the most difficult GDPR right for participants to comprehend. Nonetheless, data portability appeared to be important to respondents, 66.4% of whom indicated that the inability to transfer data between platforms was a major inhibiting factor regarding switching to a new platform.

We found two studies in which participants engaged with data portability more substantially, although they focused more on user impressions of downloaded data rather than the act of transferring data to a new platform. In Karegar et. al [49], ten participants exported data from their Google accounts, visualized it in a tool created by the authors, and were then interviewed. Most participants indicated that rather than transferring directly, they would prefer to export their data to their hard drive and then filter out some information before uploading it to a new service. Second, Bowyer et al. [11] tasked 11 internet users with making data portability requests by emailing data providers and then interviewed them about the outcomes, finding that participants were more likely to experience a decrease, rather than an increase in perceived power, as a result of the experience.

2.3 Research gap

There is a need to investigate the claims that data portability can empower users, which we pursue by examining user impressions of this technology. Some prior research has investigated user views of data portability's effectiveness for protecting *privacy* [81], however, the stated benefits of data portability extend to other dimensions of control such as agency and choice. The few empirical studies about user experiences with control and data portability have been small in scale and focused more on experiences with downloaded data than on actually transferring that data to another platform. [11, 49]. Nonetheless, this research has suggested that user experiences with current data portability systems are lackluster [11], which echoes findings about other forms of GDPR-motivated design [33]. Accordingly, there is a need for research to understand what users want and expect from data portability, and how they understand its potential to serve their needs. This is especially important regarding social media, where data portability intersects with sensitive issues of self-presentation and social connection [82, 85]. In this context, some people may be focused on limiting the exposure of their data to other people, a goal that data portability is likely to complicate rather than support. Thus, to understand user perspectives, we presented participants with illustrative examples of data portability to stimulate their awareness and ability to reflect on how data portability could serve aspects of control most important to them in the context of social media.

Lastly, most studies about data portability have been conducted in Europe, where data portability legislation is more advanced than in other regions. However, there is an expansion of similar regulations in other countries, and the GDPR drives design well beyond the European Union, including global implementations of data portability. Thus, there is a need to investigate attitudes toward data portability in other regions, which this study addresses by investigating perspectives among a representative sample of residents of the United States.

3 METHOD

We conducted an online survey of 295 U.S. residents in July 2022. Participants were recruited using *Prolific.co*, an online research recruitment platform, and the survey was delivered using Google Forms. Median completion time for the survey was just under 21 minutes and participants were paid $4.50 USD, following Prolific's guidelines. The purpose of the survey was to understand participants' current attitudes and practices regarding their social media data, their impressions of current data portability systems, and their estimation of how data portability could affect them in the future. After the surveys, we conducted 26 follow-up interviews, which lasted approximately 30 minutes and for which participants were paid $10.00 USD. Interviews served to enhance our understanding of participants' motivations and rationales for their beliefs. Particularly, they yielded in-depth descriptions about current and desired forms of controlling one's social media data.

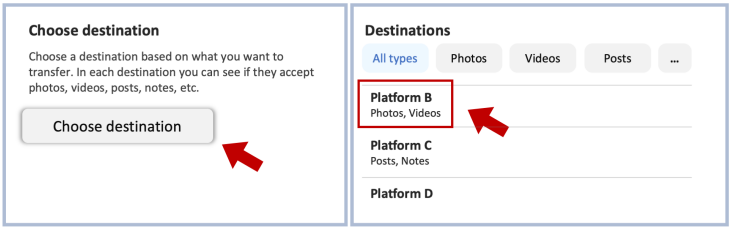
This study was reviewed and approved by our institutional review board. The full survey and interview protocol are included with this paper's supplementary documents.

3.1 Survey design

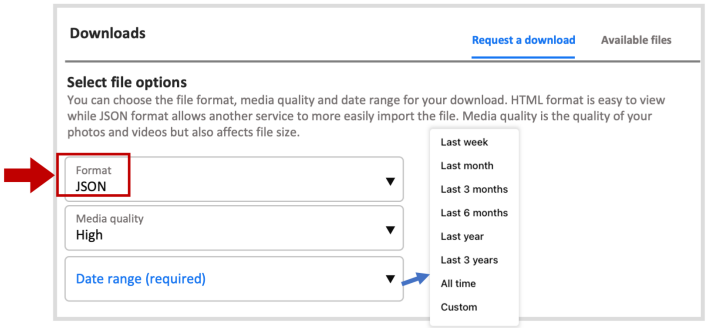
3.1.1 Current social media practices and attitudes. In the first part of the survey, participants were asked to describe their social media use, including which platforms they use, how often they use them, and how important they feel their content is. We also asked about experiences joining and leaving platforms, and about reasons for doing so. When asking about reasons for leaving platforms, we included response options adapted from prior research about people changing technology use out of protest or disagreement with technology providers [54].

3.1.2 Examples of data portability. Next, participants were presented with a description of data portability. This began with a brief overview, which explained "Data portability is the capability to obtain and reuse your personal data for your own purposes across multiple online services or platforms. In other words, it means moving or copying your data from one platform to another." Because data portability is unfamiliar to many people [51], we described simple examples to illustrate how data portability might be used by: (1) someone who wants to stop using a specific social media platform and wishes to back up her photos and videos to a different platform; and (2) someone who decides to join a new platform, which is popular among his new neighbors, and wants to copy some of his content over from a different social media platform.

Participants were then shown two examples of data portability systems, which served as probes to stimulate awareness and reflection in order to solicit participants' impressions. First, we presented a system with which one can copy social media content directly from one platform to another (*direct-transfer* system). Second, we presented a system where a user transfers their data by downloading it from a social media platform to their computer and then uploading it to a second platform (*download-transfer* system). Each system was presented as a series of mocked-up screenshots and instructions for conducting a transfer of photos from a social media platform. Examples of these instructions are presented in Figure 1. We used photos for these examples because they are regarded by most as a particularly valuable type of online content [61], however, the survey instructions



(a) Choosing a destination using the direct-transfer system. Participants were instructed that they would "Click *Choose destination*, then select *Platform B*."



(b) Choosing file options using the download-transfer system. Participants were instructed that they would "Set the format to JSON, which makes it easy to import the data to another service."

Fig. 1. Example screenshots used to illustrate the direct-transfer system (top) and the download-transfer system (bottom).

also noted that "data portability could be used with other types of data, such as status updates, comments on other people's posts, or other content."

Because Facebook offers robust data portability features, the mockup designs and instructions were based on Facebook's "Transfer your information" tool[1] for the direct-transfer system, and Facebook's "Download your information" tool[2] for the download-transfer system. The download-transfer system fulfills the GDPR's baseline criteria of providing users with a machine-readable copy of their data, and the direct-transfer represents the capability to transfer "directly from one controller to another, where technically feasible" [16]. Because we are not aware of widely adopted systems for importing downloaded social media data to a new platform, the second half of the download-transfer instructions presented a simple upload process of our own design. Participants did not actually engage in a transfer of data, but instead these screenshots were used to explain the process. Finally, these mockups did not include any branding, and instead described a transfer from "Platform A" to "Platform B". The absence of branding was to avoid cases where participants may have answered based on their feelings about a particular platform rather than toward data portability itself.

3.1.3 Measuring impressions and intentions toward the data portability systems. After viewing each mockup, participants were asked several questions to understand their impressions. These questions were drawn from the Unified Theory of Acceptance and Use of Technology (UTAUT) [87],

[1]Transfer your information tool: https://www.facebook.com/help/230304858213063
[2]Download your information tool: https://www.facebook.com/help/212802592074644

a model for describing technology adoption that consolidates previous theories such as the theory of planned behavior [1] and the technology acceptance model [18]. It has been employed in studies about technology adoption in a variety of contexts, such as contact tracing apps [45, 46, 93], health technologies [42, 44, 50, 65], self-driving cars [41, 67] and others. UTAUT is useful in the present study because it helps us understand what factors influence participants' positive or negative assessments of data portability.

UTAUT has four main constructs that influence behavioral intention to use a technology: *performance expectancy*–one's belief that a technology will be useful and effective; *effort expectancy*–how much effort an individual believes is required to use the system; *social influence*–the degree to which one believes that other people want them to use a system; and *facilitating conditions*–the degree to which one believes the system to have organizational and infrastructural support. To measure these constructs, we slightly adapted survey questions from Venkatesh [88] to make sense in the context of data portability. For example, while studies about workplace technologies often measure performance expectancy with questions about productivity or professional benefits [e.g., 88], we instead included a question about whether using each data portability system would "help me take control of my data," since this is routinely presented as a merit of data portability for internet users.

Additionally, because data portability involves tasking an online platform to manage a transfer of one's data, we hypothesized that intention may also be affected by *perceived trust* that the system will function reliably and *perceived risk* that something bad might happen as a result of using the system. Questions for measuring these constructs were adapted from prior studies that extended UTAUT to measure perceived trust [2, 3, 30] and perceived risk [3]. All questions were rated on a 5-point ordinal scale.

Lastly, many studies using UTAUT have used varying questions for measuring behavioral intention in order to suit the conditions in which a specific technology may be used [e.g., 3, 88]. We used the following four questions:

(1) If I had a need to transfer photos or other content from one platform to another, I would want to use this kind of system to do so.
(2) It is likely that I will want to transfer photos or other content from one platform to another within the next 12 months.
(3) I would want to use a system like this to transfer my data to a platform provided by Facebook, Google, Microsoft, or Twitter.
(4) I would want to use a system like this to transfer my data to a platform provided by other companies (not listed in the previous question)

We phrased these questions as hypotheticals ("I would want to use" rather than "I intend to use") because the use of a data portability system is contingent upon wanting to transfer social media data in the first place. Additionally, we included question (2), about the likelihood of wanting to transfer data within the next 12 months, to account for that contingency as part of our construct for behavioral intention. Finally, questions (3) and (4) were used to account for potentially different attitudes related to transferring data to currently dominant social media platforms versus newer or smaller platforms.

Using these constructs, factors shaping intentions to use each of the data portability systems are evaluated using the model in Figure 2. Participants answered questions about these measures separately for each data portability system (the *direct-transfer* and *download-transfer* systems), with one exception. We believe that participants can more easily answer about *social influence* toward the act of transferring social media data between platforms than hypothesize about social influence to use a specific system. Therefore, rather than asking about each system separately, we only asked

about social influence once, where the survey referred to whether other people would want the participant to copy or migrate their data to a new platform.

Some scholars have criticized studies using UTAUT and similar models for taking a narrow view, accepting their constructs as opaque rather than unpacking them, and struggling to produce knowledge about issues beyond individuals' perceptions and adoption intentions [8, 77]. To address these potential pitfalls, we use UTAUT only as a step to guide our analysis. Specifically, the main function of UTAUT in this study is to provide a framework to identify particularly important aspects of users' perspectives, which we then unpack using interviews and other survey items in order to address our research questions.

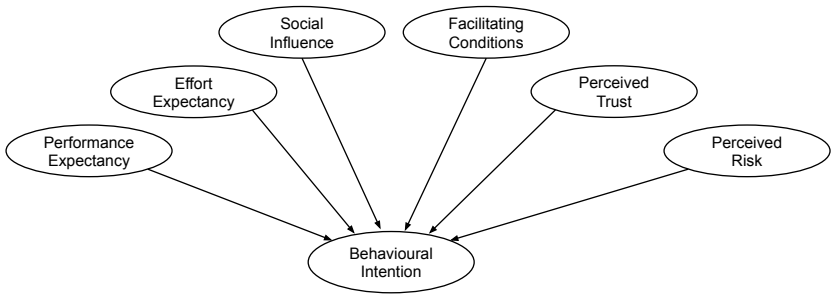


Fig. 2. Extended UTAUT model for evaluating which factors shape intentions to use data portability systems.

3.1.4 Future impacts. After giving their impressions of the two systems, participants answered questions about their potential data portability use in the future. Specifically, they indicated, "if I could use either of the systems described earlier to transfer my social media from platforms I use now to any other platform" whether this would make them more likely (a) to try new platforms, (b) to stop using platforms they currently use, and (c) to feel more in control of their online data. Participants indicated whether they agreed or disagreed with each statement using a 5-point ordinal scale.

3.1.5 Demographic questions. Finally, we asked demographic questions, including participants' race, gender, age, education level, and household income.

3.2 Survey Participants

300 survey participants were recruited using the online research platform, *Prolific.co*. Participants were recruited using Prolific's *representative sample* feature, which balances the participant distribution according to census data about race, age, and gender. To participate, participants needed to be 18 years of age or older and live in the United States.

To prevent bots and multiple submissions, Prolific requires participants to have a unique, non-VOIP phone number and a unique PayPal or Circle account to receive payment [12]. Additionally, we took the following measures to improve data quality: removed one participant who failed an attention check question and removed participants who do not use any social media platforms (n = 4) since these individuals have no data that could be subject to data portability. We considered removing participants who completed the survey particularly quickly, following the procedure described by Malhotra [58]. Only two out of 300 responses exceeded the threshold for completing too fast. Since this was an extreme minority and researchers have mixed opinions about removing unusually fast survey responses [34, 58], we left those responses as-is.

Table 1. Survey participant demographics.

Age	Gender	Race	Household income (USD)
18-24: 35 (11.9%)	Male: 140 (47.5%)	American Indian or Alaskan Native: 1 (0.3%)	<10,000: 19 (6.4%)
25-34: 59 (20.0%)	Male; Non-binary: 2 (0.7%)	Asian: 18 (6.1%)	10,000 - 24,999: 41 (13.9%)
35-44: 54 (18.3%)	Non-binary: 4 (1.4%)	Black or African-American: 35 (11.9%)	25,000 - 49,999: 70 (23.7%)
45-54: 42 (14.2%)	Female: 144 (48.8%)	Hispanic or Latino: 10 (3.4%)	50,000 - 74,999: 55 (18.6%)
55-64: 68 (23.1%)	Prefer not to disclose: 5 (1.7%)	Middle Eastern: 2 (0.7%)	75,000 - 99,999: 43 (14.6%)
65-74: 32 (10.8%)		Multiracial: 20 (6.8%)	100,000 - 149,999: 33 (11.2%)
75+: 5 (1.7%)		White: 201 (68.1%)	>150,000: 28 (9.5%)
		Other: 2 (0.7%)	Don't know: 5 (2.0%)
		Prefer not to disclose: 6 (2.0%)	

After cleaning the data, there were 295 participants, whose demographics are described in Table 1. Participants were regular social media users. The mean number of platforms for which participants had an account was 6.6. Regarding posting frequency, 35.6% of participants responded that they post to at least one social media account on a daily basis, 58.6% post at least weekly, and 77.3% post at least monthly. Visiting social media was more frequent than posting. 95.6% of participants visited one of their social media accounts on a daily basis, and 99.6% (all but one participant) visited one of their accounts at least weekly.

The sample size is more than sufficient for our analysis. Power analysis in PLS-SEM is conducted based on the largest number of paths directed at a single construct. In this study, the largest number of paths is six (the number of constructs related to behavioral intention), therefore, according to Hair [38, p. 26], a sample size of 179 is required to achieve a statistical power of 80% for detecting R^2 values of at least 0.10 (with a 1% probability of error).

3.3 Interview design

Follow-up interviews were used to add nuance to the survey results, such as understanding emotional responses to experiences with social media platforms, reasons for participants' impressions of the data portability systems in the survey, and attitudes about controlling social media data. After a brief introduction, the first interview questions were about past experiences with social media, including how participants would feel if they lost access to content they posted, and the extent to which they feel in control of the content they post online.

We then asked for more details about their impressions of the data portability systems presented in the survey as well as their estimation of future impacts, tailoring our questions to suit their survey responses. During this section, we reminded participants about survey questions referring to controlling their online data, and asked them to define what that sort of control meant to them. We saved this definition question until the second half of the survey so that we could first hear participants' unfiltered responses, and then ask them to reflect and unpack their thoughts about control.

Finally, we asked about participants' feelings toward large social media companies, and about the effect they believed data portability might have on themselves and toward other people. During this final section of the interview, we sought to expand the scope of the conversation in order to engage with broader themes of control and power in relation to social media.

3.4 Interview participants

We used stratified random sampling to select interview participants. To ensure our responses included a variety of perspectives, survey respondents who expressed interest in joining an interview were stratified into four groups based on two dimensions of their survey responses: whether they posted to social media on a daily basis or less frequently, and whether they had high or low

Table 2. Interview participant demographics

Age	Race	Gender	Age	Household income (USD)
25-34: 7 (26.9%)	Asian: 1 (4%)	Female: 10 (38%)	18-30: 3 (12%)	<10K: 1 (3.8%)
35-44: 1 (3.8%)	Black or African-American: 5 (19%)	Male: 15 (58%)	31-40: 4 (15%)	10-24K: 1 (3.8%)
45-54: 3 (11.5%)	Mixed race / other: 3 (12%)	Prefer not to disclose: 1 (4%)	41-50: 2 (8%)	25-49K: 6 (23.1%)
55-64: 11 (42.3%)	White: 15 (58%)		51-60: 10 (38%)	50-74K: 3 (11.5%)
65-74: 3 (11.5%)	Prefer not to disclose: 2 (8%)		61-70: 4 (15%)	75-99K: 7 (26.9%)
75+: 1 (3.8%)			70+: 3 (12%)	>150K: 3 (11.5%)
				Don't know: 1 (3.8%)

intention toward using data portability systems described in the survey. We then randomly invited interviewees from each group. In the first 20 interviews, older participants were over-represented, so we conducted an additional round of interviews focusing on participants in their 20s and 30s. In the end, we conducted 26 interviews. Ten interviewees were daily social media users – five with high behavioral intention and five with low behavioral intention. Ten were non-daily posters with high behavioral intention, and six were non-daily posters with low behavioral intention. Interview participants' demographics are shown in Table 2.

3.5 Data Analysis

RQ1 (current perceptions and practices around controlling social media data) and RQ3 (feelings about the potential of data portability to enhance individuals' control over online data in the future) are exploratory, descriptive questions, so we address them using a combination of descriptive statistics and thematic analysis of the interviews (see Section 3.5.2). We use two stages of analysis to address RQ2 (about the extent to which participants' perceptions of current data portability systems intersect with the topic of controlling their social media data). First, impressions of the *direct-transfer* and *download-transfer* systems from the survey are compared using Wilcoxon matched-pairs signed-rank tests, and then partial least squares structural equation modeling (PLS-SEM) is used to describe a model regarding peoples' intentions toward using each system. Second, we draw on our thematic analysis of the interviews to explain how the structural model results relate to various dimensions and priorities about controlling one's data.

3.5.1 Structural Equation Models. We evaluate two structural models to understand the degree to which the factors described in Section 3.1.3 shaped participants' behavioral intentions toward using the *direct-transfer* and *download-transfer* systems described during the survey. Structural equation modeling is a commonly used technique for fitting path models that measure structural relationships among latent constructs (e.g., *behavioral intention* and *performance expectancy*), while simultaneously assessing the measurement models for those constructs using confirmatory factor analysis [31]. We employed Partial Least Squares Structural Equation Modeling (PLS-SEM), using the *plssem* package for Stata [89]. We selected PLS-SEM (instead of covariance-based SEM) because its relaxed distributional assumptions were suitable for our data, and its higher statistical power makes it useful for examining an area where theory is undeveloped, as is the case with user perspectives of data portability [39]. Additionally, PLS-SEM has been widely used in studies about user attitudes and behavior toward technology [e.g., 2, 42, 57].

3.5.2 Qualitative analysis. Interview transcripts were analyzed using thematic analysis [13]. During and immediately after each interview, the interviewer recorded first impressions of important themes. The first author conducted inductive, initial coding using *Taguette* [74]. Following this, the authors reviewed and discussed the thematic coding to validate the results and develop a consensus about how to group individual codes into themes, following a collaborative and iterative process.

3.6 Limitations

Because this study relies on self-report data, responses may be influenced by question phrasing and the order in which questions are asked. For example, when presenting the mockups of two data portability systems, all participants were shown the direct-transfer system before the download-transfer system. Because of this, responses about the download-transfer system may have been biased by comparison to the direct-transfer system.

Additionally, like all survey research, this study's validity is threatened by potential selection bias. For example, participants' decisions to join or not join the study may have been influenced by their level of interest in social media and technology. Our study invitation used simple language and assured that no prior knowledge was necessary, which may have partially alleviated this threat. Similarly, decisions to participate in interviews may be related to participants' schedules and other personal factors. This may partially explain why, even though an effort was made for interview participants to be broadly representative, the average age among interviewees was higher than among survey participants. One consequence of this age bias is that younger people may be under-represented in the results derived from the interviews.

Finally, all participants were residents of the United States. Attitudes toward individualism, corporate power, desired forms of control, and uses of social media vary across cultures, and so attitudes expressed in this study may not be representative of people in other countries. The importance of such cultural differences is one of the motivations for this study, since research about people's perspectives of data portability in the United States is sparse. However, this also gestures to a need for similar research in other regions.

4 FINDINGS

4.1 Current perceptions and practices

To understand the contexts into which data portability systems may be deployed, we asked RQ1: "What are people's current perceptions and practices around controlling their social media data?"

The interviews included several questions about the extent to which participants feel in control of their social media data. In response, participants described a range of dimensions of control, however, the following comment by P15 is a good summation:

> In control of getting it online? Yes. In control of what happens after? No. Absolutely not. Once it's out, it's out. It could go anywhere.[...] It's completely out of my control because it can be remixed, reshared, altered, used for commercial purposes, used against me. Any number of different, potentially bad things. [...] It comes with the territory. But if I thought it was bad, I wouldn't do it. It's the price of admission. [P15]

Nineteen interview participants expressed that they are not in control of how other people see and respond to their content, such as P17 who said it is "out there for everyone to see" and P10 who remarked, "Anyone can comment on it. Obviously, I'm not in control because someone can misconstrue it and write a comment based on that." In light of this, 15 participants asserted that they take control by being selective about what content they make available for others to see, asserting that one should be "careful what you post out there" [P08] and "if I put it on the platform, it means that I want it out there" [P09].

Three participants indicated that they can control who accesses their content by deleting it, but five others remarked they do not feel able to delete something once it is posted online, such as P20, who said, "It just feels so out of my control to actually scrub the internet of anything that I posted." Additionally, P21 identified a specific instance:

> When you like something on Facebook, if that page goes away, you can't unlike it. Or at least that's true for some of them. And so I liked some political cause 12 years ago, and now I would like to unlike it. But they've shut down their page and it won't let me. [P21]

Ultimately, deleting was characterized as a time-limited form of control, e.g., "For the next 5-10 minutes, I might have control over it. [...] then after that it could be anywhere" [P08].

In addition to deciding whether to post at all, several participants described selecting specific platforms or groups when posting different types of content:

> The platforms I use are fairly siloed. As I said, LinkedIn I'll use for my professional reasons. Facebook predominantly for personal reasons [...] And of course, Nextdoor is totally focused on community. Nextdoor is about "I need a good plumber. Can anyone recommend one?" And I would not put that on Facebook because my Facebook people are all around the world. [P10]

Compared to concerns about how *other people* may access or respond to one's content, a smaller portion of interview participants (n = 9) expressed concern that they lack control about how *platforms* use their data (e.g., advertising or surveillance capitalism). For example, P12 said, "I think one of the most dangerous actions would be the selling of data to brokers or other parties because once the data has been sold to a third party you're not sure how they might handle it." Most participants did not describe taking specific actions to control this dimension of privacy, and it was common for participants to characterize giving their data to platforms as just "a part of using social media" [P09].

4.1.1 Mobility between platforms. A core premise of data portability is that it will make it easier for people to be mobile between online platforms. Therefore the survey asked about participants' current mobility practices. 29.5% (n = 87) of participants had started using a new platform within the past 12 months, with the most common being TikTok (10.8%, n = 32), Reddit (7.8%, n = 23), and Instagram (3.7%, n = 11). Among those 87 individuals, the most common reasons for joining a new platform were: "People I know use this platform" (59.8%, n = 52), "It looked fun" (44.8%, n = 39), and "It offers features I wanted to try" (24.1%, n = 21).

Among all participants, 23.1% indicated they had stopped using at least one platform within the past 12 months, with the most common being WeChat (8.5%, n = 25), Snapchat (7.8%, n = 23), and Tumblr (7.8%, n = 23). Additionally, 48.1% indicated they had "significantly reduced" using at least one platform in the past 12 months, led by Facebook (17.6%, n = 52), Instagram (14.2%, n = 42), and Pinterest (13.2%, n = 39). Further, 34.6% indicated they had ever deleted a social media account, with the most popular accounts to delete being Facebook (18.0%, n = 53), Twitter, (9.2%, n = 27), and Instagram (6.4%, n = 19).

Among the 152 participants who described having ever deleted an account, or having stopped using a platform within the past 12 months, the most common reasons were: "It was not useful or fun" (52.0%, n = 79), "Personal reasons, such as self-care and wellness" (36.2%, n = 55), and "I don't trust the company with my privacy, data, or information" (18.4%, n = 28). Notably, although *joining* a platform was largely driven by the presence of other people, only 9.2% (n = 14) explained that they had *stopped* using a platform before "people I know no longer use the platform."

4.1.2 Data loss and backups. Of the 152 respondents who indicated they had either stopped using a platform in the past 12 months or had ever deleted an account, the vast majority (86.8%, n = 132) reported that they did not do anything to back up their data.

In a separate question, 21.4% of all participants (n = 63) indicated that they have ever lost access to data they had uploaded to a social media account. Among these respondents, the most common

reason was forgetting one's password (54.0%, n = 34), followed by a platform being shut down (23.8%, n = 15) and forgetting about the platform altogether (23.8%, n = 15). Participants who had lost data were asked to rate the extent to which they felt "upset, concerned, or regretful about losing that data" from 1 ("Not at all") to 5 ("Extremely"). The mean response was 2.68 (Std. dev = 1.128), indicating that most felt between "a little" and "moderately" upset. Interview participants who lost data by forgetting their password tended not to be upset about it, such as P03:

> Kind of bummed. But then again, I think, if it was really that important to me, maybe I should have sent them a picture of my driver's license [in order to recover my password] or I would have had that data saved on my phone. [...] Is it really that important to me? [P03]

However, those who lost data due to being banned expressed much stronger feelings, such as P11, who described the loss of their data as "horrible. It is the most awful feeling" and P25:

> Hundreds of posts, thousands of comments, a couple of years worth. And they decided that they weren't going to host it anymore and it was gone [...] It pretty much angered me enough that I left the whole platform. [P25]

Both of these participants remarked that they had become more diligent about backing up their content after experiencing these bans.

4.2 Impressions of current data portability systems

This section addresses RQ2: "To what extent do participants' perceptions of current data portability systems intersect with the topic of controlling their social media data?" We address this question by examining participants' impressions and intentions toward using the direct-transfer and download-transfer systems presented in the survey, and then drawing on the interview data to explore how those impressions relate to control.

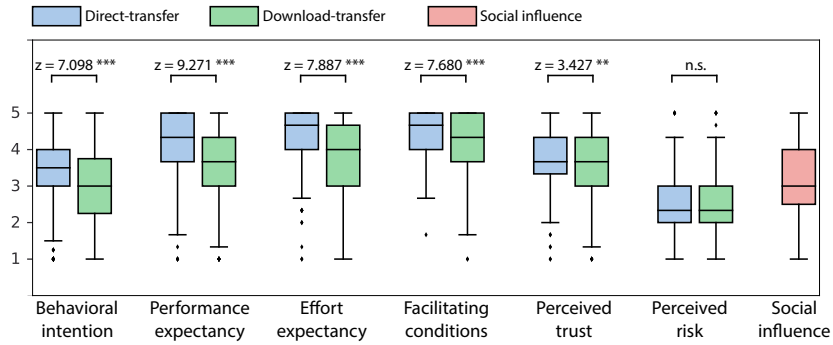


Fig. 3. Boxplots and Wilcoxon matched-pairs signed-rank test comparing impressions and intentions toward using the direct-transfer and download-systems. *** p < .001, ** p < .01, * p < .05.

4.2.1 Comparing two approaches to data portability. Figure 3 summarizes and compares impressions of the direct-transfer and download-transfer systems, using the averaged ratings for each construct, on a scale from 1 to 5. Comparisons were tested using Wilcoxon matched-pairs signed-rank tests with a Bonferroni correction. Participants had stronger intentions toward using the direct-transfer system than the download-transfer system, and rated the direct-transfer system more positively on all metrics except perceived risk, where there was no statistically significant difference. Overall,

Table 3. Measurement models

	Model 1: Direct-transfer			Model 2: Download-transfer		
Latent variable	α	CR	AVE	α	CR	AVE
Behavioural intention	0.799	0.868	0.628	0.816	0.878	0.647
Performance expectancy	0.829	0.898	0.746	0.812	0.889	0.727
Effort expectancy	0.889	0.931	0.817	0.936	0.959	0.887
Facilitating conditions	0.718	0.831	0.624	0.753	0.849	0.655
Social influence	0.871	0.939	0.886	0.871	0.939	0.886
Perceived trust	0.872	0.922	0.797	0.878	0.924	0.802
Perceived risk	0.694	0.815	0.605	0.771	0.867	0.685

α = Chronbach's Alpha; CR = Composite Reliability (DG Rho); AVE = Average Variance Extracted.

Table 4. Hypothesis testing results

	Model 1: Direct-transfer			Model 2: Download-transfer		
Hypothesis	Std. β	p	Decision	Std. β	p	Decision
Behavioural intention is predicted by:						
Performance expectancy	**0.598**	**.000**	**Supported**	**0.616**	**.000**	**Supported**
Effort expectancy	0.057	.240	Not supported	**0.137**	**.038**	**Supported**
Social influence	**0.160**	**.000**	**Supported**	**0.120**	**.002**	**Supported**
Facilitating conditions	-0.060	.184	Not supported	-0.078	.212	Not supported
Perceived trust	**0.175**	**.001**	**Supported**	0.108	.053	Not supported
Perceived risk	0.005	.902	Not supported	-0.037	.412	Not supported
Average R^2	.68			.63		

participants had moderately positive intentions toward using the direct-transfer system and neutral intentions toward using the download-transfer system. Ratings of performance expectancy, effort expectancy, facilitating conditions, and perceived trust were generally positive. Participants were largely neutral about potential risks, and social influence was also middling. Unlike the other measures, social influence was measured using one set of questions for both systems, so no comparison is presented.

4.2.2 Structural models of attitudes and intention toward data portability systems. Before proceeding to hypothesis testing, we evaluated the validity of the structural equation models. Table 3 presents both measurement models. Across both models, all measures have composite reliability and Cronbach's alpha greater than 0.70, indicating acceptable consistency [38]. Composite reliability for effort expectancy exceeds 0.95, which could indicate redundancy [39]. However, testing alternative models in which effort expectancy was measured by a single indicator did not change the results, so we retained the original models. Average variance extraction (AVE) meets the commonly accepted 0.70 threshold for sufficient convergent validity [38]. Additionally, the AVE for each construct is larger than any squared correlation among other constructs, indicating that all latent variables meet the Fornell-Larcker criterion for discriminant validity [27]. Tables describing the full measurement model (including factor loadings) and the test for discriminant validity are included in this paper's supplementary documents.

Having validated the measurement models, Table 4 shows the results of hypothesis testing to understand what factors influence participants' behavioral intentions. Each hypothesis represents a relationship proposed by the extended UTAUT model in Figure 2. Across both models, the highest variance inflation factor (VIF) was 3.4, indicating that multicollinearity was not a concern. We found

both overlaps and differences regarding the two systems, including statistically significant effects of performance expectancy, effort expectancy, social influence, and perceived trust. Performance expectancy has the largest correlation with behavioral intention, and is statistically significant in both models (direct-transfer β = 0.598, p = .000; download-transfer β = 0.616, p = .000). Social influence is also significant in both models, though has a smaller effect size (direct-transfer β = 0.160, p = .000; download-transfer β = 0.120, p = .002). Effort expectancy was only associated with behavioral intention toward using the download-transfer system (β = 0.137, p = .038), and perceived trust was only associated with behavioral intention toward using the direct-transfer system (β = 0.175, p = .001). We draw on interview results to examine ways in which these factors may be related to control.

4.2.3 Explaining the structural model results. In this section, we draw from our interview data to examine how the results in Table 4 relate to various aspects of controlling one's data.

Among most participants, performance expectancy–the belief that a technology will be useful and effective–was high. Interviews showed that one influence on performance expectancy for data portability was the degree to which one could select which content to transfer or not transfer. Six interview participants expressed this directly, noting that not all content was suited to other platforms, and that one might want to reflect on the suitability of past posts before copying them to a new platform:

> Everything isn't visible to everybody. [...] you might want to say, "I shouldn't have posted that thing, so I don't want to have that over there." So I think it's important to be able to have control over what is seen and what isn't seen as you migrate over to a new platform because it may not be appropriate for that platform. [P19]

As well as providing tangible control over self-presentation on different platforms, some participants expressed that the ability to filter data would contribute to a *feeling* of agency, such as P20: "I just feel like there's such a lack of agency in social media these days to an extent. So anything that could at least let me feel like, 'No, I don't want that' or, 'Yes, I do want that.' Even if I'm saying no on everything, it just feels nice."

However, some interview participants explained that they do not feel a strong need to transfer their content to a new platform. Four specified that they would prefer to start with a blank slate if they join a new platform, rather than transferring any content over: "I feel like a new platform is like a new start of something else because they're all a little different as far as content goes and how interaction goes" [P06]. Another four interviewees explained that their social media use is focused on the present, making the migration of past posts feel unimportant. Two of those individuals specified that what they really want to transfer is the ability to reach their social media contacts even if they switch platforms: "It's not important to me what they said in 2017. It is important to me that I can contact them today" [P05].

This emphasis on connecting in the present, rather than transferring past content, appears tightly bound with the role of social influence. Only one participant specifically stated that their decision about transferring data would depend on social influence: "If everyone else is importing Instagram stuff, I would probably do the same. But if everyone was kind of doing it as a clean slate, I would probably do that as well" [P09]. However, ten interview participants indicated that their motivation for joining a social media platform was primarily shaped by their friends and family. These comments indicated that social influence is most evidently a requirement for wanting to try a new platform in the first place, which is a prerequisite for wanting to transfer one's data: "The real driver for whether or not I'm going to move to a platform is: If I go there, are my friends going to be there? Are other people that do what I want to do there? If they're not, why do I care about getting my data moved over?" [P15].

Regarding content itself, the ability to ensure access to one's past data seemed to be a higher priority for some participants than actually transferring it to somewhere specific. Six participants indicated that data portability would be useful specifically for this purpose, e.g.: "I just like to know that I do have access to things that happened ten years ago. And so I actually really like the idea [of data portability] a lot" [P13]. However, a few participants also remarked that they already feel in control of maintaining a backup or archive through other means, such as P10: "Were Facebook to be the only place that those [photos] existed. I would be worried." Thus, although maintaining backups is an important dimension of control, opinions were mixed about whether data portability was the most appropriate tool for this goal.

Thirteen interview participants made statements regarding perceived trust. Five generally regarded the direct-transfer as trustworthy, such as by noting that it provided less opportunity for user error than the download-transfer. However, another five were critical that the direct-transfer could leak information to third parties or transfer more than one intended, e.g., "Somewhere in that transfer process, my mind worries about who can get access to it [...] There would be a difference. I think the [download-transfer] would be safer." [P07]. The remaining three gave ambiguous responses, such as expressing a general distrust toward platforms. Significant here, is that distrust toward the direct-transfer system was related to its perception as a black box, to be trusted or distrusted blindly.

Some participants described a trade-off regarding the download-transfer system, where it was regarded as providing better performance and trustworthiness at the expense of requiring a greater effort. Specifically, four participants explicitly stated that the download-transfer would give them more "control" or "agency" than the direct-transfer and seven others made statements alluding to that effect. However, out of twelve participants who discussed effort expectancy during their interviews, nine characterized the download-transfer as requiring more effort than the direct-transfer, such as P07: "Are you more in control of it [with the download-transfer]? I don't know. But I would think it's easier [with the direct-transfer]." Participants did not characterize this trade-off as a matter of being capable or incapable of performing the download-transfer, but rather expressed that they preferred the option that involved less work: "I got the skill. It's just effort. Why would I do it if [the platform] could do it?" [p16].

4.3 Anticipated impacts on control in the future

After having viewed and given their impressions of two data portability systems, survey respondents were asked to indicate how they would be affected if they could use either of the systems described in this survey to transfer social media content from platforms they use now to any other platform.

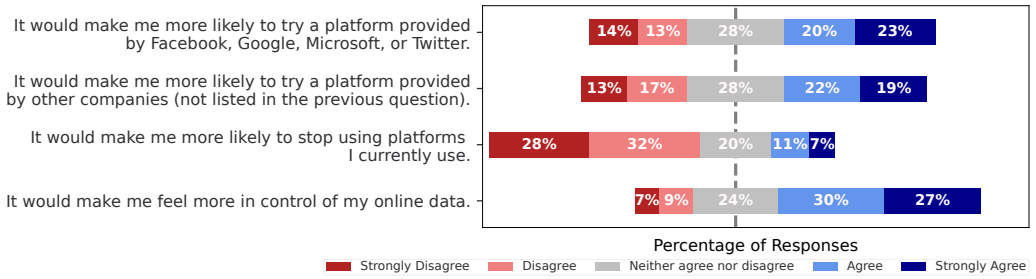


Fig. 4. Participants' responses to the prompt: "In general, if I could use either of the systems described in this survey to transfer from social media content from platforms I use now to any other platform..."

We draw on the responses, summarized in Figure 4, as well as related interview data, to answer RQ3: "How do participants feel about the potential of data portability to enhance individuals' control over online data in the future?"

4.3.1 Joining a new platform. Responses were mixed regarding whether data portability would make one more likely to try a new platform, whether it was provided by a major social media provider (Facebook, Google, Microsoft, or Twitter) or a different company. Positive responses included that data portability "would make it easier" to move to another platform [P16]. Similarly, P23 asserted that it might encourage one to try new services: "If a good platform comes that's better than Facebook and I'm able to move all my stuff with a click of the button, I could give it a trial more so than if it's going to take some unknown process" [P24]. Similarly, P05 was optimistic that data portability could encourage people to join a new platform, but focused their explanation on the ability to transfer one's contacts:

> When I tried new platforms with some of my friends, one of the things was how do we find each other? How do we know we have the same name on Platform A as Platform B? We have to reach out, send friend requests, you know, sometimes send private messages on another platform, "Hey, what's your name on this platform?" And so the more transferability, sure, the more likely people are to just pick up and go. [P05]

In contrast, negative responses emphasized that other factors were more important for motivating a decision to try a new platform:

> The portability is not the issue there. It's the type of platform. The type of audience who uses that platform. How can you share? What options do they offer? [...] Unless you have another platform out there that meets your needs, what would be the point of moving it? [P23]

4.3.2 Quitting current platforms. Most participants indicated that the availability of data portability would not make them more likely to stop using any platforms they currently use. P06 expressed skepticism that data portability would motivate people to quit because, "I don't know how many people there are that feel like they're forced to use a platform." By contrast, P21 explained, "At least in my demographic, I feel like Facebook has an iron grip on people" and considered that data portability might make it easier to "capture those people," but concluded, "That being said, I just don't see it happening." Nonetheless, even though most participants did not feel that data portability would make them more likely to leave a platform they currently use, they were positive that data portability would be helpful if they had already made that decision, e.g.:

> Knowing that I can easily grab it and post or send it to some other system. That would make me feel more in control if my digital life is going to follow me now that I'm done with Facebook and I go on to X-book or whatever. [P17]

4.3.3 Feeling more in control. The majority of participants agreed that data portability would make them feel more in control of their online data, and only 16% disagreed. Nine interview participants directly expressed that data portability would give them more control. Several noted that data portability would make it easier to do what they want with their data, e.g.: "It would make it easier knowing that I can, and that it's there for me to access it just like that. [...] You can do with it what you want. We're not holding it hostage [...] It gives me more ownership" [P16]. And others emphasized that this would simply make them *feel* better about social media:

> As far as trust and believability and, you know, does this give me a feeling of control? It gives me something of a warm feeling that if I've got a significant collection of media that I want to move over, that's a nice feeling. [P15]

Overall, participants generally did not frame this as allowing entirely new types of control but instead emphasized that it would make things more convenient, such as P12: "In terms of control over where my data goes, it's not much of a change. It just makes it more *convenient* to control my data" [P12] (emphasis added).

In spite of the high scores in response to survey questions about control, eight interview participants expressed that data portability did not address the aspects of control that were most important to them. Six of these interviewees said that rather than copying their content to other platforms, they were more interested in limiting its scope, such as P20: "Personally, for me, I'm not out there looking to join new platforms and give them all my information from the old platforms. It's pretty much the opposite of my mindset." Three of those participants specified that they would feel more in control by increasing their ability to delete their content altogether: "I think the biggest control over online data for me would be being able to tell these companies like, 'Hey, delete my data,' which is rarely possible" [P09].

Related to big-picture changes to social media, three participants suggested that data portability, if well-regulated, could stimulate competition and motivate social media platforms to improve their services for all users, including P24's suggestion that "it could also scare the big social media companies into firming up some of their loose ends when they know that there are options and people can move more relatively at ease." However, another five expressed skepticism that this could result in significant change. These objections were based on a perception that all social media platforms would have the same problems, e.g.:

> You could say, "I don't like Facebook anymore so I want to take my data and I want to move it somewhere else." But the somewhere else is probably not different from Facebook in terms of those things that really matter [...] Two faces same coin. [P11]

And further, three participants remarked that, if anything, this would increase the amount of data to which platforms have access, such as P05 who remarked, "We'd be on five different social media platforms with the same data on all five, and I don't think it would necessarily be a loss to any one social media platform" and P04 who said data portability "empowers the company [...] to a greater extent because they get an influx of new people, new data to work with." Thus, even though participants were generally optimistic about the potential for data portability to increase their control, many were skeptical about whether this could cause a substantial re-balancing of power between platforms and users.

5 DISCUSSION

Regarding their current experiences with social media, participants generally expressed that they have little to no control over their social media data after it is posted. The most visible threat to control among interview participants was managing how *other people* accessed one's data, in order to manage self-presentation and avoid conflicts. The main way in which participants described exercising control over their online data was to be judicious about what they post in the first place. This is consistent with prior research about how users manage self-presentation in social media [43] and with a Pew survey finding that 74% of U.S. adults believe it is important to "keep things about themselves from being searchable online" [5]. Participants were moderately mobile when it came to joining and quitting platforms, and deciding to join a new platform was primarily motivated by the presence of social ties. Upon leaving a platform, most participants did not seem particularly concerned about backing up their social media data. However, those that lost data due to being banned had strong feelings about that loss.

When presented with illustrative examples of direct-transfer and download-transfer-based data portability systems, participants had positive impressions of both, though overall preferred the

direct-transfer system. The largest influence on their behavioral intention to use these systems was performance expectancy. During interviews, participants expressed a range of opinions about whether data portability would be useful for them or not. Some particularly strong themes were that data portability's usefulness would depend on the ability to carefully select what is transferred, and that many participants were more concerned with keeping connected to other people in the present than with moving their data from the past. To that end, it is unsurprising that social influence was also an important factor for both systems, which matches our result that deciding to join a new platform was primarily motivated by the presence of social ties.

To address this study's main query, participants generally agreed that data portability would make them feel more in control of their online data. However, many characterized the scope of this enhanced control as limited, such as making certain behaviors easier but not necessarily motivating new behaviors. A few participants suggested that data portability could give users leverage over platforms, echoing recent research about data leverage [90]. On the other hand, many were skeptical that data portability could create meaningful competitive pressure, consistent with Florez Ramos and Blind's [26] assertion that data portability would fail to drive change in platforms like Facebook and Google that operate without serious competition. In many cases, participants felt data portability would make things easier or more convenient, but did not articulate that it would create important opportunities that were completely out of reach beforehand. And for others, the type of control they wanted was not necessarily served by data portability, such as among those who wanted more ability to delete data rather than to copy it to yet another platform.

In sum, it appears that data portability was generally viewed as a step in the right direction, but alone does not quite fulfill the opportunity described by the European Commission to "re-balance" the relationship between users and online platforms [70, p. 4]. To some extent, this is because participants were concerned with multiple types of control, and had varied opinions about which were most important. For example, they differed in the extent to which they prioritized keeping in touch with contacts, managing self-presentation, preserving access to their data, and protecting their privacy. This poses a challenge for regulators and platforms seeking to implement data portability across large populations, since large scales involve standardization that tends to gloss over heterogeneity [40]. Consequently, some of the most meaningful implementations of data portability may be those that are focused on or can be tailored to specific forms of control.

5.1 Design implications

5.1.1 Being selective about what to transfer. The first set of implications we discuss is related to the ability to choose specific content to transfer or not transfer. The two main uses participants described for data portability–creating an archive of one's data and transferring data to another platform–have different needs. For creating an archive, participants are likely to want as complete a record as possible (accounting for available storage space). However, regarding transferring content to another platform where it will be visible to other users (e.g., transferring data to a social networking site), participants expressed a desire to select what to transfer.

Some platforms allow users to assign different audience settings to specific content (e.g., Facebook's *Public*, *Friends*, and *Only Me* options), and these settings could serve as useful filters for data portability. For example, one might choose to transfer only their *public* posts, if the destination platform is public. However, only a minority of Facebook users change these settings on a post-by-post basis [23]. Even though the use of those audience settings is low, many participants in our study expressed that they were careful about what they posted and considered the audiences that would see it. Some asserted that decisions could be platform-specific, for example by noting that only a portion of the content posted to a platform like Facebook might be appropriate to transfer elsewhere. This is consistent with prior research indicating that internet users project

different self-presentations [76, 85] and engage with different topics [53] on different platforms, and that these differences can be very sensitive [e.g,. 36, 37]. Thus, for facilitating transfers to a new platform, it is vital to consider that decisions about whether some content is appropriate to post to *Platform X* do not automatically extend to *Platform Y*. Accordingly, data portability of social media data may demand more granular visibility controls than most people use within a single platform.

For active social media users, data portability may involve transferring a large archive of content, and manually selecting which data to transfer may be burdensome. Prior research has shown that many users welcome automation to help with the chore of managing their data [91], and so filters based on image-recognition, keyword search, topic-extraction, post geo-location, or similar methods could make it easier for users to identify which content they want to transfer and which they view as inappropriate for the destination platform. Regarding potentially sensitive content, it is possible to go beyond selecting *whether or not* to transfer content by adding tools for obfuscating potentially sensitive information. For example, Li et al. [55] built a system that warns users if their photos contain potentially sensitive content, and provides capabilities for blurring or covering that portion of the image.

In sum, as data portability implementations become more common–particularly *importing* data into a new platform where it may be visible to new audiences–designers should provide ways to carefully review and curate what information is transferred. Failure to provide such options could contribute to harms if users are nudged toward making data more visible than they want to, or could cause hesitance to transfer *any* data if they feel they lack sufficient control over their self-presentation.

5.1.2 Making data portability social. Given that we are talking about "social" media, it is not surprising that participants overwhelmingly asserted that decisions about joining a new social media platform are largely shaped by their friends, family, and others. And it is not surprising that some participants expressed that being able to transfer their contacts would be much more useful than transferring their content. This resonates with findings from prior research [24] that migration to new platforms is often a community activity, and should be supported as such. In fact, the Electronic Frontier Foundation (EFF) has asserted that Facebook should give users "a way to export the rich contact list that Facebook hosts" in order to facilitate a smooth transfer to a competing service [17]. In response, Facebook has asserted that allowing users to export their social graph would compromise privacy [22]. This is a reasonable position, and in fact, Facebook removed the ability to access unique user identifiers from its API as a response to the Cambridge Analytica scandal [98]. Consequently, third parties had less ability to collect and analyze user data in bulk (a victory for privacy), but this also negatively impacted services that serve user interests by allowing people to transfer their Facebook data to their own personal website [47]. In other words, design decisions made to prioritize users' privacy can constrain data portability, and conversely, decisions that prioritize data portability, particularly in social contexts, may create threats to privacy. Thus, we consider some potential design approaches for transferring one's contact list as part of data portability, while accounting for potential privacy risks.

For an illustrative example, we turn to the migration of some Twitter users to Mastodon after Elon Musk purchased Twitter [94]. Some migrating users implemented ad-hoc solutions to find each other on the new platform, such as listing a Mastodon user ID in their Twitter profile, and then using third-party applications such as Debirdify[3] or Fedifinder[4] to find referenced Mastodon accounts among their Twitter *following* list. While functional, this solution requires users to trust a third-party service, which several participants in our study expressed reservations about. Further,

[3]https://pruvisto.org/debirdify/
[4]https://fedifinder.glitch.me/

even though groups of users are often skilled at coordinating workarounds to migrate to new technologies, this can exacerbate existing inequalities. For example, Bin Morshed et al. [10] found that when Facebook's Free Basics Internet service was banned in Bangladesh for 22 days, many affected users shared information about how to bypass the ban using a VPN. However, knowledge and access to use a VPN were unevenly distributed, which exacerbated a pre-existing digital divide.

To reduce reliance on third-party services or ad-hoc solutions, platforms could implement a standard set of fields for users to list their identities on other platforms, and then provide the ability to export this information (in fact, many platforms include fields like this, though generally do not allow this to be exported). For dealing with non-public profiles, this is complicated by the fact that when making decisions about granting third parties permission to access their data, people have been shown to give less consideration for others' privacy than they do their own [60]. Ultimately, transferring contacts is a case of multi-party privacy, where actions of one party affect the privacy of another [83]. This returns us to the disagreement between Facebook and EFF, that transferring social graphs is a potential privacy risk. But, it is not *only* privacy that is at stake. Restricting access to contact lists due to a privacy concern constrains people's agency to join new social media platforms in a social manner–that is, to keep in touch with their social media contacts even if they change platforms. To that end, a possible solution may be to implement tools for group-level data portability, such as by allowing a circle of peers to coordinate a mutual transfer (or copy), without requiring ad-hoc solutions or removing any individual's agency to determine how others can use their contact information or other data.

6 CONCLUSION

Researchers, regulators, and technologists have identified data portability as an important tool for enhancing internet users' control over their online data. In spite of this potential, there is a lack of knowledge about user perceptions of data portability, which makes it difficult to understand what people want from and expect from data portability, including whether they believe it can contribute to meaningful empowerment. This paper has addressed that gap by presenting the results of an empirical study of USA residents, focusing on the context of social media because it is one of the most vital arenas for personal data rights.

Most participants were in agreement that they lack control over their social media data, but they expressed a diverse range of opinions about what dimensions of control were most important. In light of this situation, most agreed that data portability *would* enhance their control to a degree. However, we observed significant uncertainty and doubt about the scope of this enhancement.

Our interpretation of the results was informed by the observation that, even though data portability is defined in sweeping legislation and the most prominent implementations are in massive, global platforms, what users most wanted from this technology was personal and tied to specific contexts. We presented an argument that meaningfully achieving data portability's promise for empowerment will require designs that are tailored to specific forms of control. Therefore, we identified and discussed opportunities for addressing needs that were particularly important in a social media context, such as filtering data to carefully manage one's self-presentation and coordinating transfers in groups.

Our decision to focus on social media illuminated patterns related to online socializing and group coordination, which may have presented differently, or not at all, had we studied other types of platforms. This also means we have been unable to explore important issues that would arise in other contexts. Thus, future research about data portability in contexts like self-tracking and online collaborative work (e.g., open-source development) may illuminate new tensions and opportunities for future designs.

REFERENCES

[1] Icek Ajzen. 1991. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50, 2 (Dec. 1991), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

[2] Karrar Al-Saedi, Mostafa Al-Emran, T. Ramayah, and Eimad Abusham. 2020. Developing a General Extended UTAUT Model for M-payment Adoption. *Technology in Society* 62 (Aug. 2020), 101293. https://doi.org/10.1016/j.techsoc.2020.101293

[3] Wissal Ben Arfi, Imed Ben Nasr, Galina Kondrateva, and Lubica Hikkerova. 2021. The Role of Trust in Intention to Use the IoT in eHealth: Application of the Modified UTAUT in a Consumer Context. *Technological Forecasting and Social Change* 167 (June 2021), 120688. https://doi.org/10.1016/j.techfore.2021.120688

[4] Sara Atske. 2019. Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. Retrieved {2022-12-05} from https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[5] Brooke Auxier. 2020. Most Americans Support Right to Have Some Personal Info Removed from Online Searches. Retrieved {2023-01-05} from https://www.pewresearch.org/fact-tank/2020/01/27/most-americans-support-right-to-have-some-personal-info-removed-from-online-searches/

[6] Marlene Barth. 2021. A Case Study on Data Portability. *Datenschutz und Datensicherheit - DuD* 45, 3 (March 2021), 190–197. https://doi.org/10.1007/s11623-021-1416-3

[7] Eric P.S. Baumer, Phil Adams, Vera D. Khovanskaya, Tony C. Liao, Madeline E. Smith, Victoria Schwanda Sosik, and Kaiton Williams. 2013. Limiting, Leaving, and (Re)Lapsing: An Exploration of Facebook Non-Use Practices and Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 3257–3266. https://doi.org/10.1145/2470654.2466446

[8] Izak Benbasat and Henri Barki. 2007. Quo Vadis TAM? *Journal of the Association for Information Systems* 8, 4 (April 2007). https://doi.org/10.17705/1jais.00126

[9] Yochai Benkler. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press, New Haven [Conn.].

[10] Mehrab Bin Morshed, Michaelanne Dye, Syed Ishtiaque Ahmed, and Neha Kumar. 2017. When the Internet Goes Down in Bangladesh. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, Portland Oregon USA, 1591–1604. https://doi.org/10.1145/2998181.2998237

[11] Alex Bowyer, Jack Holt, Josephine Go Jefferies, Rob Wilson, David Kirk, and Jan David Smeddinck. 2022. Human-GDPR Interaction: Practical Experiences of Accessing Personal Data. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3491102.3501947

[12] Bradley. 2018. Bots and Data Quality on Crowdsourcing Platforms | Prolific Blog. Retrieved {2021-12-15} from https://blog.prolific.ac/bots-and-data-quality-on-crowdsourcing-platforms/

[13] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[14] Axel Bruns. 2008. *Blogs, Wikipedia, Second Life, and beyond: From Production to Produsage.* Peter Lang.

[15] Bhaskar Chakravorti. 2020. Why It's So Hard for Users to Control Their Data. *Harvard Business Review* (Jan. 2020). https://hbr.org/2020/01/why-companies-make-it-so-hard-for-users-to-control-their-data

[16] European Commission. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance).* Technical Report. http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng

[17] Bennett Cyphers and Danny O'Brien. 2018. Facing Facebook: Data Portability and Interoperability Are Anti-Monopoly Medicine. Retrieved {2023-01-05} from https://www.eff.org/deeplinks/2018/07/facing-facebook-data-portability-and-interoperability-are-anti-monopoly-medicine

[18] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13, 3 (1989), 319–340. https://doi.org/10.2307/249008 arXiv:249008

[19] Deloitte. 2018. *A New Era for Privacy: GDPR Six Months On.* Technical Report. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/risk/deloitte-uk-risk-gdpr-six-months-on.pdf

[20] Edelman. 2018. Trust Barometer Special Report: Brands and Social Media. Retrieved {2022-11-22} from https://www.edelman.com/research/trust-barometer-brands-social-media

[21] Samson Esayas and Angela Daly. 2018. The Proposed Australian Consumer Data Right: A European Comparison. https://papers.ssrn.com/abstract=3236020

[22] Facebook. 2020. *Comments to the Federal Trade Commission on Data Portability.* Technical Report. https://about.fb.com/wp-content/uploads/2020/08/Facebook-Comments-to-FTC-on-Data-Portability.pdf

[23] Casey Fiesler, Michaelanne Dye, Jessica L. Feuston, Chaya Hiruncharoenvate, C.J. Hutto, Shannon Morrison, Parisa Khanipour Roshan, Umashanthi Pavalanathan, Amy S. Bruckman, Munmun De Choudhury, and Eric Gilbert. 2017. What (or Who) Is Public? Privacy Settings and Social Media Content Sharing. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 567–580. https://doi.org/10.1145/2998181.2998223

[24] Casey Fiesler and Brianna Dym. 2020. Moving Across Lands: Online Platform Migration in Fandom Communities. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 42:1–42:25. https://doi.org/10.1145/3392847

[25] Casey Fiesler and Blake Hallinan. 2018. "We Are the Product": Public Reactions to Online Data Sharing and Privacy Controversies in the Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173627

[26] Esmeralda Florez Ramos and Knut Blind. 2020. Data Portability Effects on Data-Driven Innovation of Online Platforms: Analyzing Spotify. *Telecommunications Policy* 44, 9 (Oct. 2020), 102026. https://doi.org/10.1016/j.telpol.2020.102026

[27] Claes Fornell and David F. Larcker. 1981. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research* 18, 1 (1981), 39–50. https://doi.org/10.2307/3151312 arXiv:3151312

[28] Christian Fuchs. 2014. *Social Media: A Critical Introduction.* SAGE, Los Angeles.

[29] Francis Fukuyama. 2021. Making the Internet Safe for Democracy. *Journal of Democracy* 32, 2 (2021), 37–44. https://doi.org/10.1353/jod.2021.0017

[30] Ellen Garbarino and Mark S. Johnson. 1999. The Different Roles of Satisfaction, Trust, and Commitment in Customer Relationships. *Journal of Marketing* 63, 2 (April 1999), 70. https://doi.org/10.2307/1251946 arXiv:1251946

[31] David Gefen, Detmar Straub, and Marie-claude Boudreau. 2000. Structural Equation Modeling Techniques and Regression: Guidelines for Research Practice. *Commun Assoc Inform Syst* 4 (Jan. 2000), 2–76.

[32] Robert W. Gehl. 2015. The Case for Alternative Social Media. *Social Media + Society* 1, 2 (2015), 1–12. https://doi.org/10.1177/2056305115604338

[33] Colin M. Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3411764.3445779

[34] Robert Greszki, Marco Meyer, and Harald Schoen. 2015. Exploring the Effects of Removing "Too Fast" Responses and Respondents from Web Surveys. *The Public Opinion Quarterly* 79, 2 (2015), 471–503. arXiv:24546373 https://www.jstor.org/stable/24546373

[35] Anatoliy Gruzd, Jenna Jacobson, Philip Mai, and Elizabeth Dubois. 2018. *Social Media Privacy in Canada.* SSRN Scholarly Paper ID 3195503. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3195503

[36] Oliver Haimson. 2018. Social Media as Social Transition Machinery. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 63:1–63:21. https://doi.org/10.1145/3274332

[37] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2895–2907. https://doi.org/10.1145/2858036.2858136

[38] Joseph F. Hair (Ed.). 2017. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (second edition ed.). Sage, Los Angeles.

[39] Joseph F. Hair, Jeffrey J. Risher, Marko Sarstedt, and Christian M. Ringle. 2019. When to Use and How to Report the Results of PLS-SEM. *European Business Review* 31, 1 (Jan. 2019), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

[40] Alex Hanna and Tina M. Park. 2020. Against Scale: Provocations and Resistances to Scale Thinking. arXiv:arXiv:2010.08850 http://arxiv.org/abs/2010.08850

[41] Charlie Hewitt, Ioannis Politis, Theocharis Amanatidis, and Advait Sarkar. 2019. Assessing Public Perception of Self-Driving Cars: The Autonomous Vehicle Acceptance Model. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*. Association for Computing Machinery, New York, NY, USA, 518–527. https://doi.org/10.1145/3301275.3302268

[42] Rakibul Hoque and Golam Sorwar. 2017. Understanding Factors Influencing the Adoption of mHealth by the Elderly: An Extension of the UTAUT Model. *International Journal of Medical Informatics* 101 (May 2017), 75–84. https://doi.org/10.1016/j.ijmedinf.2017.02.002

[43] Xiaoyun Huang, Jessica Vitak, and Yla Tausczik. 2020. "You Don't Have To Know My Past": How WeChat Moments Users Manage Their Evolving Self-Presentation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376595

[44] P. Ifinedo. 2012. Technology Acceptance by Health Professionals in Canada: An Analysis with a Modified UTAUT Model. In *2012 45th Hawaii International Conference on System Sciences*. 2937–2946. https://doi.org/10.1109/HICSS.2012.556

[45] Jack Jamieson, Daniel A Epstein, Yunan Chen, and Naomi Yamashita. 29. Unpacking Intention and Behavior: Explaining Contact Tracing App Adoption and Hesitancy in the United States. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. New Orleans, USA, 14. https://doi.org/10.1145/3491102.3501963

[46] Jack Jamieson, Naomi Yamashita, Daniel A Epstein, and Yunan Chen. 2021. Deciding If and How to Use a COVID-19 Contact Tracing App: Influences of Social Factors on Individual Use in Japan. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 481:1–481:30. https://doi.org/10.1145/3479868

[47] Jack Jamieson, Naomi Yamashita, and Rhonda McEwen. 2022. Bridging the Open Web and APIs: Alternative Social Media Alongside the Corporate Web. *Social Media + Society* 8, 1 (Feb. 2022), 13. https://doi.org/10.1177/20563051221077032

[48] Cecilia Kang and Sheera Frenkel. 2018. Facebook Says Cambridge Analytica Harvested Data of up to 87 Million Users. *The New York Times* (April 2018). https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html

[49] Farzaneh Karegar, Tobias Pulls, and Simone Fischer-Hübner. 2016. Visualizing Exports of Personal Data by Exercising the Right of Data Portability in the Data Track - Are People Ready for This? In *Privacy and Identity Management. Facing up to Next Steps: 11th IFIP WG 9.2, 9.5, 9.6/11.7, 11.4, 11.6/SIG 9.2.2 International Summer School, Karlstad, Sweden, August 21-26, 2016, Revised Selected Papers*, Anja Lehmann, Diane Whitehouse, Simone Fischer-Hübner, Lothar Fritsch, and Charles Raab (Eds.). Springer International Publishing, Cham, 164–181. https://doi.org/10.1007/978-3-319-55783-0_12

[50] Boonchai Kijsanayotin, Supasit Pannarunothai, and Stuart M. Speedie. 2009. Factors Influencing Health Information Technology Adoption in Thailand's Community Health Centers: Applying the UTAUT Model. *International Journal of Medical Informatics* 78, 6 (June 2009), 404–416. https://doi.org/10.1016/j.ijmedinf.2008.12.005

[51] Sophie Kuebler-Wachendorff, Robert Luzsa, Johann Kranz, Stefan Mager, Emmanuel Syrmoudis, Susanne Mayr, and Jens Grossklags. 2021. The Right to Data Portability: Conception, Status Quo, and Future Directions. *Informatik Spektrum* 44, 4 (Aug. 2021), 264–272. https://doi.org/10.1007/s00287-021-01372-w

[52] Thomas A Lambert. 2022. Addressing Big Tech's Market Power: A Comparative Institutional Approach. *SMU LAW REVIEW* 75 (2022), 45. https://scholar.smu.edu/smulr/vol75/iss1/8/

[53] Roy Ka-Wei Lee, Tuan-Anh Hoang, and Ee-Peng Lim. 2017. On Analyzing User Topic-Specific Platform Preferences Across Multiple Social Media Sites. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1351–1359. https://doi.org/10.1145/3038912.3052614

[54] Hanlin Li, Nicholas Vincent, Janice Tsai, Jofish Kaye, and Brent Hecht. 2019. How Do People Change Their Technology Use in Protest?: Understanding "Protest Users". *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 87:1–87:22. https://doi.org/10.1145/3359189

[55] Yifang Li and Kelly Caine. 2022. Obfuscation Remedies Harms Arising from Content Flagging of Photos. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–25. https://doi.org/10.1145/3491102.3517520

[56] Siân E. Lindley, Catherine C. Marshall, Richard Banks, Abigail Sellen, and Tim Regan. 2013. Rethinking the Web as a Personal Archive. In *Proceedings of the 22nd International Conference on World Wide Web - WWW '13*. ACM Press, Rio de Janeiro, Brazil, 749–760. https://doi.org/10.1145/2488388.2488454

[57] Dominique Machuletz, Stefan Laube, and Rainer Böhme. 2018. Webcam Covering as Planned Behavior. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3173754

[58] Neil Malhotra. 2008. Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly* 72, 5 (Dec. 2008), 914–934. https://doi.org/10.1093/poq/nfn050

[59] Kate Mannell and Eden T. Smith. 2022. Alternative Social Media and the Complexities of a More Participatory Culture: A View From Scuttlebutt. *Social Media + Society* 8, 3 (July 2022), 20563051221122448. https://doi.org/10.1177/20563051221122448

[60] Maximilian Marsch, Jens Grossklags, and Sameer Patil. 2021. Won't You Think of Others?: Interdependent Privacy in Smartphone App Permissions. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–35. https://doi.org/10.1145/3479581

[61] Catherine C. Marshall and Frank M. Shipman. 2015. Exploring the Ownership and Persistent Value of Facebook Content. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, Vancouver BC Canada, 712–723. https://doi.org/10.1145/2675133.2675203

[62] Tsubasa Morioka, Nicole B. Ellison, and Michael Brown. 2016. Identity Work on Social Media Sites: Disadvantaged Students' College Transition Processes. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative*

Work & Social Computing (CSCW '16). Association for Computing Machinery, New York, NY, USA, 848–859. https://doi.org/10.1145/2818048.2819959

[63] Mika Nakashima. 2022. Comparison of Legal Systems for Data Portability in the EU, the US and Japan and the Direction of Legislation in Japan. In *Human Choice and Digital by Default: Autonomy vs Digital Determination (IFIP Advances in Information and Communication Technology)*, David Kreps, Robert Davison, Taro Komukai, and Kaori Ishii (Eds.). Springer International Publishing, Cham, 159–169. https://doi.org/10.1007/978-3-031-15688-5_14

[64] National Congress of Brazil. 2018. *Brazil General Data Protection Law (LGPD)*. Technical Report. https://iapp.org/media/pdf/resource_center/Brazilian_General_Data_Protection_Law.pdf

[65] P. Ndayizigamiye and M. Maharaj. 2016. Mobile Health Adoption in Burundi: A UTAUT Perspective. In *2016 IEEE Global Humanitarian Technology Conference (GHTC)*. 613–623. https://doi.org/10.1109/GHTC.2016.7857342

[66] Whitney Nixdorf. 2019. Planting in a Walled Garden: Data Portability Policies To Inform Consumers How Much (If Any) of the Harvest Is Their Share. *Transnational Law & Contemporary Problems* 29, 1 (2019), 135–165.

[67] Sina Nordhoff, Victor Malmsten, Bart van Arem, Peng Liu, and Riender Happee. 2021. A Structural Equation Modeling Approach for the Acceptance of Driverless Automated Shuttles Based on Constructs from the Unified Theory of Acceptance and Use of Technology and the Diffusion of Innovation Theory. *Transportation Research Part F: Traffic Psychology and Behaviour* 78 (April 2021), 58–73. https://doi.org/10.1016/j.trf.2021.01.001

[68] Government of Canada. 2022. Canada's Digital Charter: Trust in a Digital World. Retrieved {2022-11-14} from https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter-trust-digital-world

[69] Tim O'Reilly. 2005. What Is Web 2.0. Retrieved {2015-06-09} from http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1

[70] Article 29 Data Protection Working Party. 2017. *Guidelines on the Right to Data Portability*. Technical Report 16/EN WP 242 rev.01. European Commission. https://ec.europa.eu/newsroom/article29/items/611233

[71] Andrew Perrin. 2018. Americans Are Changing Their Relationship with Facebook. Retrieved {2023-01-05} from https://www.pewresearch.org/fact-tank/2018/09/05/americans-are-changing-their-relationship-with-facebook/

[72] Wanda Presthus and Hanne Sørum. 2019. Consumer Perspectives on Information Privacy Following the Implementation of the GDPR. *International Journal of Information Systems and Project Management* 7, 3 (2019), 19–34. https://doi.org/10.12821/ijispm070302

[73] Data Transfer Project. 2018. Data Transfer Project Overview and Fundamentals. Retrieved {2023-01-06} from https://datatransferproject.dev/dtp-overview.pdf

[74] Rémi Rampin and Vicky Rampin. 2021. Taguette: Open-Source Qualitative Data Analysis. *Journal of Open Source Software* 6, 68 (Dec. 2021), 3522. https://doi.org/10.21105/joss.03522

[75] Richard Rogers. 2020. Deplatforming: Following Extreme Internet Celebrities to Telegram and Alternative Social Media. *European Journal of Communication* 35, 3 (June 2020), 213–229. https://doi.org/10.1177/0267323120922066

[76] Leah Scolere, Urszula Pruchniewska, and Brooke Erin Duffy. 2018. Constructing the Platform-Specific Self-Brand: The Labor of Social Media Promotion. *Social Media + Society* 4, 3 (July 2018), 2056305118784768. https://doi.org/10.1177/2056305118784768

[77] Aviv Shachak, Craig Kuziemsky, and Carolyn Petersen. 2019. Beyond TAM and UTAUT: Future Directions for HIT Implementation Research. *Journal of Biomedical Informatics* 100 (Dec. 2019), 103315. https://doi.org/10.1016/j.jbi.2019.103315

[78] Frank M. Shipman and Catherine C. Marshall. 2020. Ownership, Privacy, and Control in the Wake of Cambridge Analytica: The Relationship between Attitudes and Awareness. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376662

[79] Clay Shirky. 2008. *Here Comes Everybody: The Power of Organizing without Organizations*. Penguin, New York.

[80] Nick Srnicek. 2017. *Platform Capitalism*. Polity Press, Cambridge, UK.

[81] J. Strycharz, J. Ausloos, and N. Helberger. 2020. Data Protection or Data Frustration? Individual Perceptions and Attitudes Towards the GDPR. *European Data Protection Law Review* 6, 3 (2020), 407–421. https://doi.org/10.21552/edpl/2020/3/10

[82] Frederic Stutzman and Woodrow Hartzog. 2012. Boundary Regulation in Social Media. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 769–778. https://doi.org/10.1145/2145204.2145320

[83] Jose M. Such and Natalia Criado. 2018. Multiparty Privacy in Social Media. *Commun. ACM* 61, 8 (July 2018), 74–81. https://doi.org/10.1145/3208039

[84] Emmanuel Syrmoudis, Stefan Mager, Sophie Kuebler-Wachendorff, Paul Pizzinini, Jens Grossklags, and Johann Kranz. 2021. Data Portability between Online Services: An Empirical Analysis on the Effectiveness of GDPR Art. 20. *Proceedings on Privacy Enhancing Technologies* 2021, 3 (July 2021), 351–372. https://doi.org/10.2478/popets-2021-0051

[85] Edson C Tandoc, Jr., Chen Lou, and Velyn Lee Hui Min. 2019. Platform-Swinging in a Poly-Social-Media Context: How and Why Users Navigate Multiple Social Media Platforms. *Journal of Computer-Mediated Communication* 24, 1 (Jan. 2019), 21–35. https://doi.org/10.1093/jcmc/zmy022

[86] Sarah Turner, July Galindo Quintero, Simon Turner, Jessica Lis, and Leonie Maria Tanczer. 2021. The Exercisability of the Right to Data Portability in the Emerging Internet of Things (IoT) Environment. *New Media & Society* 23, 10 (Oct. 2021), 2861–2881. https://doi.org/10.1177/1461444820934033

[87] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27, 3 (2003), 425–478.

[88] Viswanath Venkatesh and Xiaojun Zhang. 2010. Unified Theory of Acceptance and Use of Technology: U.S. Vs. China. *Journal of Global Information Technology Management* 13, 1 (Jan. 2010), 5–27. https://doi.org/10.1080/1097198X.2010.10856507

[89] Sergio Venturini and Mehmet Mehmetoglu. 2019. **Plssem** : A *Stata* Package for Structural Equation Modeling with Partial Least Squares. *Journal of Statistical Software* 88, 8 (2019). https://doi.org/10.18637/jss.v088.i08

[90] Nicholas Vincent and Brent Hecht. 2021. Can "Conscious Data Contribution" Help Users to Exert "Data Leverage" Against Technology Companies? *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 103:1–103:23. https://doi.org/10.1145/3449177

[91] Francesco Vitale, William Odom, and Joanna McGrenere. 2019. Keeping and Discarding Personal Data: Exploring a Design Space. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. ACM, San Diego CA USA, 1463–1477. https://doi.org/10.1145/3322276.3322300

[92] Emily A. Vogels. 2021. 56% of Americans Support More Regulation of Major Technology Companies. Retrieved {2022-12-05} from https://www.pewresearch.org/fact-tank/2021/07/20/56-of-americans-support-more-regulation-of-major-technology-companies/

[93] Michel Walrave, Cato Waeterloos, and Koen Ponnet. 2020. Ready or Not for Contact Tracing? Investigating the Adoption Intention of COVID-19 Contact-Tracing Technology Using an Extended Unified Theory of Acceptance and Use of Technology Model. *Cyberpsychology, Behavior, and Social Networking* (Oct. 2020). https://doi.org/10.1089/cyber.2020.0483

[94] Kyle Wiggers. 2022. Boosted by Twitter Drama, Mastodon Reaches 1 Million Active Monthly Users. Retrieved {2023-01-09} from https://techcrunch.com/2022/11/07/boosted-by-twitter-drama-mastodon-reaches-1-million-active-monthly-users/

[95] Janis Wong and Tristan Henderson. 2018. How Portable Is Portable?: Exercising the GDPR's Right to Data Portability. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, Singapore Singapore, 911–920. https://doi.org/10.1145/3267305.3274152

[96] Savvas Zannettou, Barry Bradlyn, Emiliano De Cristofaro, Haewoon Kwak, Michael Sirivianos, Gianluca Stringini, and Jeremy Blackburn. 2018. What Is Gab: A Bastion of Free Speech or an Alt-Right Echo Chamber. In *Companion Proceedings of the The Web Conference 2018 (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1007–1014. https://doi.org/10.1145/3184558.3191531

[97] Ben Zefeng Zhang, Tianxiao Liu, Shanley Corvite, Nazanin Andalibi, and Oliver L. Haimson. 2022. Separate Online Networks During Life Transitions: Support, Identity, and Challenges in Social Media and Online Communities. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 458:1–458:30. https://doi.org/10.1145/3555559

[98] Mark Zuckerberg. 2018. Hearing Before the United States House of Representatives Committee on Energy and Commerce. https://docs.house.gov/meetings/IF/IF02/20180508/108260/HHRG-115-IF02-Wstate-BarrettG-20180508.pdf

[99] Diana Zulli, Miao Liu, and Robert Gehl. 2020. Rethinking the "Social" in "Social Media": Insights into Topology, Abstraction, and Scale on the Mastodon Social Network:. *New Media & Society* (July 2020). https://doi.org/10.1177/1461444820912533

[100] Zoe Zwiebelmann and Tristan Henderson. 2021. Data Portability as a Tool for Audit. 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*. ACM, Virtual USA, 276–280. https://doi.org/10.1145/3460418.3479343

Received January 2023; revised April 2023; accepted May 2023