A Longitudinal Analysis of Corporate Data Portability Practices Across Industries

Emmanuel Syrmoudis
Technical University of Munich
Munich, Germany
emmanuel.syrmoudis@tum.de

Stefan A. Mager
Independent Researcher
Munich, Germany
research@stefanmager.org

Jens Grossklags
Technical University of Munich
Munich, Germany
jens.grossklags@in.tum.de

Abstract—**Lock-in practices of online services hinder consumers from switching frictionlessly to a competitor once they are unsatisfied with the company's service offering, privacy practices, or philosophy. The right to data portability (RtDP) is one of the strongest measures introduced by recent privacy regulations to unlock continuously collected user data from centralized silos of market leaders. Introducing the obligation to provide means of data transfers between services, it aims to establish decentralized online markets and to foster competition. In this longitudinal study comprising a unique dataset of 129 online services over three consecutive years, we are the first to provide evidence on the development of the effectiveness of the EU's RtDP. Astonishingly, only 16% of services could provide a compliant data export in all years, with services from the industries *Entertainment* and *Travel* performing worst. Overall, *Finance & Insurance* and *Social Networks & Messaging* include the services with the highest compliance rates. Regarding the usefulness of data portability, our analysis unveils that data export scope and data import options have stagnated between 2020 and 2022. Further, we are able to show that online services with a high presence of third-party trackers are less compliant and ready to export data from their systems. Lastly, our regression analyses show that service popularity significantly increases format compliance, export scope, and import options. This suggests that competitors to incumbents still perceive the regulation more as a bureaucratic burden than a unique opportunity to attract new consumers and their data.**

Index Terms—**Data portability, Longitudinal study, Consumer rights, Competition between online services, Third-party trackers, Privacy regulation, General Data Protection Regulation (GDPR), Online service industry categorization**

1. Introduction

Personal user data has become ever more centralized due to large, market-leading online services locking in their user bases. For instance, services like the social network Instagram or the search engine Google were not designed to give individuals control over the content and data traces they produce in these services. The servers that make up their networks are all owned and operated by these companies. Whereas a few open web advocate groups and early corporate initiatives, e.g., Google Takeout, recognized and intended to oppose user data centralization [1], [2], the private economy did not find effective ways to do so. Thus, to systematically counteract this development and grant individuals more control over their personal data, the European Union (EU), California, Brazil, and China adopted a consumer right to data portability (RtDP) as part of the privacy regulations they enacted between 2018 and 2021 [3].

Data portability is defined as a transfer of personal data from one service provider to another upon request of the user. This concept is designed to shift *control* over personal data from corporations to individual users. By doing so, users would gain more freedom and be less dependent on companies' privacy practices. When combined with the right to erasure, data portability allows individuals to move their data more freely between online services of their choice. A well-designed data portability framework has the potential to become a central pillar of user privacy rights. It can facilitate a transition towards more decentralized and privacy-preserving web architectures, such as the Solid project initiated by Tim Berners-Lee [4]. This framework also enables users to erase their digital footprints from centralized web systems when paired with the right to erasure. Beyond empowering individuals, data portability legislation can promote competition in online service markets [5]. In theory, it allows communities to collectively switch to competing services, thereby reducing the network effects that benefit incumbent service providers and encouraging decentralization.

Taking control of one's own data has been shown to be a central concern of users in the privacy space across many studies over the last decades. For example, in a survey study from 2005, Acquisti and Grossklags found that 90% of participants agreed with a privacy definition centered on ownership and control of personal information [6]. Enhancing control stands in contrast to the largely ineffective, take-it-or-leave-it notion of *notice and consent/choice* [7]. Control approaches include, for example, browser signals like Do-Not-Track or Global Privacy Control [8], [9], or configuration options on the sites of service providers [10] or mobile/IoT devices [11], [12]. Data portability offers a potentially more fundamental alternative vector for users to

The published version of this article is available at DOI: 10.1109/ACSAC63791.2024.00032.

"vote with their feet" and take control of their data.

With the enactment of several RtDPs, academic interest in data portability has increased. As no jurisdiction had ever passed an RtDP before 2018 [13], no experience existed of how markets with data portability consumer rights would develop. In response, game-theoretic models have been devised by economists to predict the strategic behavior of companies in markets with data portability [5], [14] and whether it would facilitate the entry of competitors to concentrated online markets [15]. In computer law and policy articles, scholars further debated about the adequate scope of data exports [16] and discussed which file formats fulfill desired criteria such as machine-readability [17]. Starting with the enactment of the General Data Protection Regulation (GDPR) of the EU, also first conceptual [18], [19] and cross-sectional studies [17], [20] assessed the compliance behavior of companies and the attitudes of consumers to make use of an RtDP [21].

Yet, there is a lack of continuous empirical studies monitoring how corporate data portability practices evolve over time. This paper aims to address this gap by providing a longitudinal analysis of data portability practices. From a regulatory perspective, compliance with the provisions of the RtDP is of central importance, which motivates our first research question:

RQ1: How does corporate compliance with data portability regulation evolve over time?

As a regulation usually places a financial and organizational burden on companies, a vital question is whether the regulation takes effect over time and whether policy goals are reached. In the case of data portability, these goals comprise the increase of data export scope and import options in order to make data transfers more beneficial for consumers [22]. Thus, the research question arises:

RQ2: How does the usefulness of RtDP data transfers change over time for individuals?

In the context of data portability regulation compliance, the privacy-related behaviors of online services are a particularly interesting object of study. Companies that use fewer trackers on their websites often demonstrate a higher regard for user privacy. This behavior may indicate a greater likelihood of adhering to user rights outlined in privacy regulations. Therefore, we pose the following research question:

RQ3: How does the presence of third-party trackers influence the compliance of online services with data portability regulation?

Ultimately, we consider it to be of relevance to investigate whether data portability represents a more meaningful concept for certain online industries and whether there are industries in which services perform better than others when it comes to embracing data portability. Therefore, we pose the following final research question:

RQ4: How does the performance of data portability practices differ between particular online service industries?

To answer our research questions, we gathered a unique dataset on data export mechanisms of 129 corporations and import measures of 137 corporations over three consecutive years, from 2020 to 2022. Among these corporations are the top 100 companies in terms of service popularity according to the Alexa Germany ranking [23] of the respective years. We further enriched this data with company information on location of headquarters, turnover, and number of employees from the ORBIS database [24] and popularity from the Tranco ranking [25]. Furthermore, we pulled information on the number of third-party trackers present on the website of the investigated online services from the WhoTracks.Me project [26]. In contrast to cross-sectional studies, the longitudinal approach allowed us to use a random effects model and, therefore, to control for confounding variables (such as companies' characteristics – size, revenue, and, most importantly, unobservable side factors) more effectively and to reduce biases, e.g., due to convenience sampling (selection bias) and differences in data collection processes.

In order to cluster the companies into industry categories, we gathered data from the EU's NACE framework [27], Similarweb [28], and Wikipedia [29], but needed to recognize that available official industry categorizations of services in online markets are vague, imprecise, and sometimes outdated. Therefore, we conducted a survey in which we set 500 participants for the task of sorting a randomly selected number of companies from our sample into common categories based on perceived service similarity. We used the results of this survey to let the clustering algorithm by Clauset et al. [30] establish 11 industry categories. Finally, we took these categories to be able to assess performance differences in data portability practices among industries. This approach allowed us to group services according to the users' perceptions and needs, providing a more realistic and user-centered overview of the online landscape.

This paper presents several distinct contributions that enhance and build upon existing work on data portability [17], [20]. Firstly, we focus on longitudinal data to investigate whether the adoption of data portability regulation is inherently slow or if compliance by online services remains persistently low over an extended period. Our findings reveal that, since the publication of our previous cross-sectional study [20] in 2021, there has been no significant progress in the effectiveness of Article 20 GDPR. This provides a more robust and reliable input for scholars and policymakers. Additionally, we make improvements to our methodology [20] by incorporating a survey for industry classification and offering a new perspective on the relationship between GDPR compliance and tracking behavior. Thus, through the findings from our panel analysis, we make the following contributions to the literature on data portability:

- Compliance with the EU's RtDP is low and stagnating. Since the GDPR has formulated one of the strictest yet most useful RtDPs for consumers in comparison to the privacy regulations of other jurisdictions, it can be assumed that also other jurisdictions will experience low effectiveness of their respective policies.
- Small and medium-sized enterprises (SMEs) are barely using data portability regulation to attract new consumers (and their data). Thus, popular incumbent services can strengthen their market position while still complying generously with the regulation.

- Surprisingly, online services with a high presence of trackers do not comply significantly better with an RtDP than more privacy-focused companies. This can be attributed to a greater tendency among these companies to disregard privacy regulations.
- Advances in RtDP compliance can be traced back to service popularity but also seem to have their roots in industry affiliation.

In the next sections, this paper unfolds as follows: First, we introduce the reader to the regulatory concepts of the EU's RtDP. Then, we point to recent empirical work on data portability and data access rights since we use some of their results as the foundation for our longitudinal analysis. Next, we describe our methodological procedure for data collection on data portability practices of online services and the industry categorization of online services. In the results section, we then demonstrate the gathered measures in descriptive statistics and present the results of our random effect regressions on compliance, export scope, and import options. Next, we discuss our results in the context of previous literature and provide implications for corporations, policymakers, and consumers. A brief section, which includes our contributions made and states potential limitations, concludes the paper.

2. Related Work on Data Portability

2.1. Theoretical and Legal Studies

The potential of data portability to reduce switching costs between online services, lower entry barriers for startups in online markets, and thus foster decentralization is widely recognized [5], [20], [31]. While the term was originally coined to facilitate user switching between social networks [1], the concept of data portability received more academic attention when the EU started evaluating to make it a consumer right [2], [13]. Four jurisdictions have enacted a right to data portability (RtDP) until today: EU, California, Brazil, and China. In Appendix B, we aim to provide our readers with a brief introduction to the similarities and differences of RtDP designs across jurisdictions.

In contrast to positive arguments for data portability regulation, Engels [32] brought forward that a generalized RtDP for consumers could also increase entry barriers to online markets and harm SMEs since the compliance burden to provide user data on request weighs heavier on them than on bigger corporations. Apart from potential negative effects on competition, Swire and Lagos [13] additionally outlined that an RtDP might reduce consumer welfare, as companies would likely pass on the costs to consumers that they incur in making their software systems compliant.

2.2. Empirical Studies

Apart from theoretical analyses, a couple of empirical studies have been published on rights concerning data portability in recent years. In an early cross-sectional study on data portability exports that covers the first months after the GDPR entered into effect, Wong and Henderson [17] not only thoroughly assessed file formats that fully meet the definition of "structured, commonly used and machine-readable" [22], e.g., CSV, EML, ICS, JSON, MBOX, TEX, VCS, XML. They also examined the file format compliance, response duration time, and authentication process of 230 companies. They find that 74.8% of the online services responded within the required response time.

Taking up the research conversation, in an earlier cross-sectional study [20], we made use of the interpretation of the file format definition by Wong and Henderson [17] when analyzing the data exports of 182 online services from 2020. Beyond that, we investigated the scope of data that the corporations exported by building on the data taxonomy of de Hert et al. [16]. Table 1 gives an overview of the definitions of the data taxonomy elements and related personal data examples [20]. In our baseline study [20], we found that 26.4% of services in the sample met the requirements of the GDPR's data portability right by exporting only received data. 39.6% of services, however, also exported observed data and 6.6% even exported inferred data. Finally, around one quarter of the services (27.4%) did not provide any data export at all upon request until the deadline.

Considering the potential will of consumers to reuse their requested data exports, a data import scope rating was developed [20]. It tracks whether data can be imported on all core functionalities (*full import possibilities*; 6.8% in the baseline study) or only on some of the core functionalities of the online service (*partial import possibilities*; 10.5%). Moreover, it indicates whether data for purposes other than core functionalities of the services can be imported (*minimal import possibilities*; 5.8%) or no data at all (76.8%). Lastly, these measures were regressed on website popularity – with the Alexa ranking as proxy, finding that service popularity exerts a positive effect on data export scope and the data import possibilities offered. Building upon this work, our current study investigates with a refined industry classification whether the results of the baseline study [20] can be confirmed for a longitudinal dataset. Moreover, it evaluates what developments have occurred between 2020 and 2022.

3. Methods

3.1. Data Collection: Data Portability Practices of Online Services

To set up this study, we build on the notion of indirect data portability [20]: Indirect data portability combines the right of Article 20(1) GDPR – which allows individuals to request personal data from any of their online services to save them to their premises – with one of the potential user interests leading to the export of personal data in the first place: to move it to a new competitive service. Article 20(2) GDPR envisions direct data portability for such a user need [22]. As interfaces for direct data portability between

TABLE 1. DATA PORTABILITY EXPORT SCOPE TAXONOMY WITH EXAMPLES

	Received data	**Observed data**	**Inferred data**
Definition (according to de Hert et al. [16])	Raw data that is actively and wittingly provided by the user	Raw (meta)data that is passively and unwittingly provided by the user	Knowledge about past and present behaviors and preferences of a user which is produced by the online service on the basis of received and observed data, plus usually other data sources
Examples (from the 2021 baseline study [20])	Personal information (e.g., name, address, e-mail, ...), Contacts, Calendar entries, Comments, Social media posts and likes, Product reviews	GPS/Location data, Fitness and training data (e.g., heart rate,...), Search engine queries, Usage data (e.g., login times, clicked links, used devices/browsers, ...)	Advertising/Job/Sites interests, Credit rating information, Speech transcripts

most online services are nowadays still missing[1], in fact, an ordinary user would currently typically need to make use of indirect data portability in order to transfer data between two services [34].

Enhancing our prior cross-sectional work [20], we examined the data export practices of 129 services and the import capabilities of 137 services on a yearly basis over a period of three years between 2020 and 2022[2]. As for some online services, we could only collect the data portability practices in one or two of the three years, we have data on 254 exporting and 256 importing online services in total when taking into account all yearly cross-sectional datasets.

3.1.1. Schemes. For labeling the data exports, we built an initial spreadsheet with labels based on prior work. This spreadsheet featured attributes related to *compliance, export scope, duration, authentication, and transfer method*. In particular, the export scope was classified as *received / observed / inferred data* according to the taxonomy defined by de Hert et al. [16] and the EU's official guidelines on Article 20 GDPR [22]. We followed Wong and Henderson [17] in classifying compliant file formats. We further required a complete export within at most one month (as stated in Article 12 GDPR) to classify a service as compliant. According to the provisions of Article 20 GDPR, an export was considered complete if all data that was ever actively provided (i.e., *received data*) by the user to the service was contained in the export.

Moreover, as part of the cross-sectional paper in 2020 [20], we developed a scheme for imports as there was no prior work to base it on. For each service, core functionalities were to be identified (e.g., mail, messaging, video streaming, banking). We then developed import scope labels based on whether online services where providing no import options at all (*none*), import options for non-core functionalities only (*minimal*), import options for some core functionalities (*partial*), and all core functionalities (*full*). During data collection and analysis in 2020, we refined the spreadsheet based on the practical challenges that occurred during the process (e.g., some services unexpectedly confused the RtDP with the Right of Access). In 2021 and 2022, student assistants were provided with the refined spreadsheet to document and label their data exports and imports in a structured way. A detailed overview of the labeling schemes and the instructions given to the student assistants is provided in Appendix A.

3.1.2. Collection Process. Data export requests were made by two of the authors and one student assistant in 2020. To avoid potential biases (it might raise suspicion when the same person is repeatedly making data requests, and thus online services might treat these requests differently), we delegated the request process to 2 student assistants in 2021 and 3 separate student assistants in 2022. To ensure rigor in the data collection process, we thoroughly explained all elements of the spreadsheet to the assistants, gave them detailed instructions, and asked them to read the related work (e.g., [16], [17], [20], [22]).

We aimed at making each year's requests in the first half of the year to make sure that the interval between requests was approximately one year for each service. In 2020, all requests but eight were made between January and April; the remaining ones were made in September. In 2021, the requests were made in March and April. In 2022, the requests were made between February and May.

Regarding service assignment, the authors and assistants involved in the request process compiled a list of all online services (mainly web-based with some app-based services that also provide a web frontend) they were registered with each year. In case of overlaps, only one person was randomly assigned the service. In addition, the top 100 online services of the Alexa Germany ranking [23] were included. Where necessary, new accounts were created and filled with fake data (except for name and address to ensure we could provide identification when requested) in each of its functionalities over the course of several weeks. In total, the share of newly created accounts in the panel dataset is 49.4%. When multiple services were part of a single company, we regarded those that are used with a joint account (e.g., Google and YouTube) as one service and

[1] The Data Transfer Project, as the biggest direct data portability initiative, has launched its version 1.0 in March 2023 with prominent but – as of 2024 – still few companies participating [33].

[2] There is a difference between the export and the import side as 8 services charged expensive usage/subscription fees and could, therefore, only be analyzed regarding their import capabilities.

those that have separate accounts (and, therefore, separate GDPR request procedures – e.g., Facebook and WhatsApp) as separate services. Afterward, requests under Article 20 GDPR were made for both existing and new accounts. From the compiled list of services, we excluded services where no possibility existed to create accounts, where English or German were not among the offered languages, adult services, and services that aimed at business users only (e.g., services targeting institutions or shop creators). To allow enough time for creating and filling accounts where necessary, we gathered the current Alexa rankings around five months before the requests were made, with the first list being the Alexa Top 100 from October 2019.

As a final result, the assistants provided a detailed description of the data export request procedures and results as well as their examinations on import options. For privacy reasons, assistants did not disclose the raw data exports but provided the authors with a detailed list of attributes present in the data exports. Furthermore, the student assistants were asked to classify the data. To ensure consistency over the years, the classifications of the student assistants were not used in the final dataset. Instead, two of the authors classified the data based on the assistants' verbose descriptions.

3.1.3. Reliability. As we jointly discussed, e.g., how to assess data scopes, we could not assess inter-rater reliability in 2020. However, we identified this as a limitation of our approach and, therefore, made sure that the process in 2021 and 2022 was set up in a way that allowed us to assess inter-rater reliability. As a quality control measure, we then computed Kendall's coefficient of concordance [35] (corrected for ties) between the assistants' and the authors' classifications: W = 0.82 for 2021 and W = 0.84 for 2022. An analogous procedure was used to classify imports, with W = 0.88 for 2021 and W = 0.82 for 2022.

3.1.4. Collection of Secondary Data. To analyze services based on their popularity, we chose to use the Tranco ranking [25], a research-focused list of top sites, as a proxy for popularity[3]. After initially using Alexa for service selection, we switched to Tranco for our analyses due to the discontinuation of Alexa in 2022. Furthermore, the switch to Tranco allowed us to use a more robust ranking due to its methodology of aggregating data from multiple sources. To mitigate outliers in the ranking data (e.g., an event leading to a short-term rise of popularity of a website), we made use of Tranco's feature to set custom time frames over which the average ranking is computed and chose January 1 to December 31 as the ranking period for each year.

To answer RQ3, we used the longitudinal dataset on the average number of trackers per online service website from the WhoTracks.Me database [26], [36]. We used snapshots from 2020-03, 2021-03, and 2022-03 in our analysis to match the points in time in which we collected our dataset on data portability practices.

TABLE 2. LOCATION (HEADQUARTERS) OF SERVICES IN THE SAMPLE.

ISO Code	AE	CH	DE	DK	ES	FR	GB	HK	IE
#	1	2	50	1	3	6	4	1	1
ISO Code	JP	KY	LU	NL	NO	SE	US	ZA	
#	2	4	1	5	1	1	53	1	

3.1.5. Sample Characteristics. Among the 137 services in the longitudinal dataset, 53 are based in the USA and 75 in Europe. See Table 2, for a full breakdown by countries. While our service selection did not result in a representative sample, it still includes a well-distributed mixture of highly popular and less popular services, with 33 services ranked between 1 and 100, 70 between 101 and 10000, and 34 above, as measured by the Tranco ranking [25].

3.2. Industry Categorization of Online Services

As RQ4 aims at online service industries, we needed a way to categorize the services in our sample into industries. We found standard *Classification for Economic Activity* frameworks like NACE (EU), ISIC (UN) or NAICS (US) to be too focused on the old economy and therefore too vague and imprecise for the classification of online services[4] (compare [20]). Furthermore, other (unofficial) sources for industry classification, such as the "type of site" field in Wikipedia articles or Similarweb industries, turned out to be incomplete, inconsistent, or lacked transparency.

We decided to gather a user-generated industry classification by conducting a survey on the platform Prolific Academic. After a pretest with 20 participants, we conducted the survey with a total of 500 participants over the course of 9 days in April 2023. To maximize the likelihood that participants know the services in the sample, we screened for Prolific users who listed German as one of their fluent languages and who resided in Germany, Austria, Switzerland, Liechtenstein, or Luxembourg. 91% of participants were residing in Germany, 6% in Austria, 2% in Switzerland, and 0.4% in Luxembourg. 56% identified as male, while 44% identified as female. Participants were between 18 and 73 years old (mean = 29.8; median = 27.0). The median time spent for completing the survey was 11 minutes.

In light of the goal of Article 20 GDPR to facilitate switching, in particular, between similar services, we asked participants which services they considered similar. For each participant, we randomly ordered 255 services[5] and used six-point Likert scales to ask how well the specific service is known to them (ranging from "do not know at all" to "know very well"). 16 services were shown per survey page, and upon clicking "Next," the subsequent set of 16 services

[3] Available at https://tranco-list.eu/list/52XN (2020), https://tranco-list.eu/list/YGVG (2021), https://tranco-list.eu/list/4KLQX (2022).

[4] To give an example: In the EU's NACE rev. 2 industry classification, Booking.com is listed in the industry of *6312: Webportals*. This is a too broad category for a service that is nowadays serving millions of individuals in the reservation of travel accommodation and means of travel.

[5] As described in Section 3.1, we have data on 256 services in total, but one went out of business, so we included 255 in the survey.

was shown. This process continued until the participant had selected option 3 (“rather do not know”) or a higher option for at least 35 services.

Using a card sort approach [37], [38], participants were then instructed to sort these 35 services into two or more boxes. The number of boxes was at each participant’s discretion; each box had to contain services that the participant perceived as similar. See Figure 1 for a sample screenshot of the task. During the sorting phase, the boxes were labeled “Industry 1,” “Industry 2,” etc. After the sorting phase, the labels were replaced by text inputs and participants were asked to provide a label for each box. These labels were used as quality control measures and taken into account when naming the industries created by the clustering algorithm. Consequently, we checked the respondents’ labels and removed those boxes from the dataset where the label did not meaningfully describe the services (e.g., “random,” “other,” “diverse”).

After conducting the survey, we created a graph with one node per service. For each pair of services (that was in the random selection for at least 3 users), we created an edge with weight $w = \frac{\text{\#samebox}}{\text{\#samebox + \#differentbox}}$, indicating the relative frequency of services being assigned the same box by participants.

We then used the Clauset-Newman-Moore algorithm [30] to generate clusters of similar services[6]. We merged clusters containing 4 or fewer services into an *Other* cluster (that contains 8 services) and gave the remaining clusters names that characterize the services they contain. Using this method, 131 of the 137 services in our panel dataset were assigned a fitting industry. For the remaining 6 services, almost all participants stated that they did not know the respective service. Therefore, we manually assigned these services to the industry categories in which they would fit best, based on information about the primary purpose of the services according to the ORBIS database [24]. Table 3 gives an overview of the industries and the number of services they contain by popularity (Tranco rank in 2022).

3.3. Research Ethics

Our institution does not require an ethics review for questionnaire-based studies (classification task) and data collection. Nevertheless, we aimed to minimize negative implications for the survey participants, the studied services, and the involved employees. We, therefore, made sure to follow the methodology of related studies [17], [20], [39], in particular regarding “undercover” data collection, the choice not to name specific services and not to debrief them to avoid negative consequences for employees. When conducting the survey and analyzing the data, we followed standard practices for ethical research: presenting detailed study procedures, obtaining consent, and not collecting identifiable information or device data.

Given that our data collection started two years after the GDPR became effective, receiving a GDPR request is not uncommon for online services (in 2019, 6%–19% of respondents per EU country reportedly have exercised the RtDP [40]) and, therefore, its fulfillment should not impose an unreasonable cost on the service. For the services, our requests were not distinguishable from the requests of ordinary users. As our focus was on organizational practices rather than on the individual behavior of employees, we limited our interactions with employees to answering questions posed by them during the request process and providing the required data for authentication.

To avoid negative implications, especially for employees of services that turned out to be not compliant with Article 20 GDPR, we decided against debriefing the services and publishing individual results that make services identifiable. We believe that our approach minimizes the harm caused by our data collection and that it is outweighed by the societal benefits of our study.

4. Results

Having executed requests under Article 20 GDPR over the course of 3 years for 129 online services, we show how the data portability practices of services developed from 2020 to 2022. Using the industries clustered from survey data with 500 participants, we compare these practices across industries. The following subsections cover the compliance of online services with Article 20 GDPR (RQ1), the usefulness of data transfers (RQ2), the interplay of third-party tracking and data portability (RQ3), and the comparison of practices across industries (RQ4).

4.1. Compliance with Data Portability Regulation

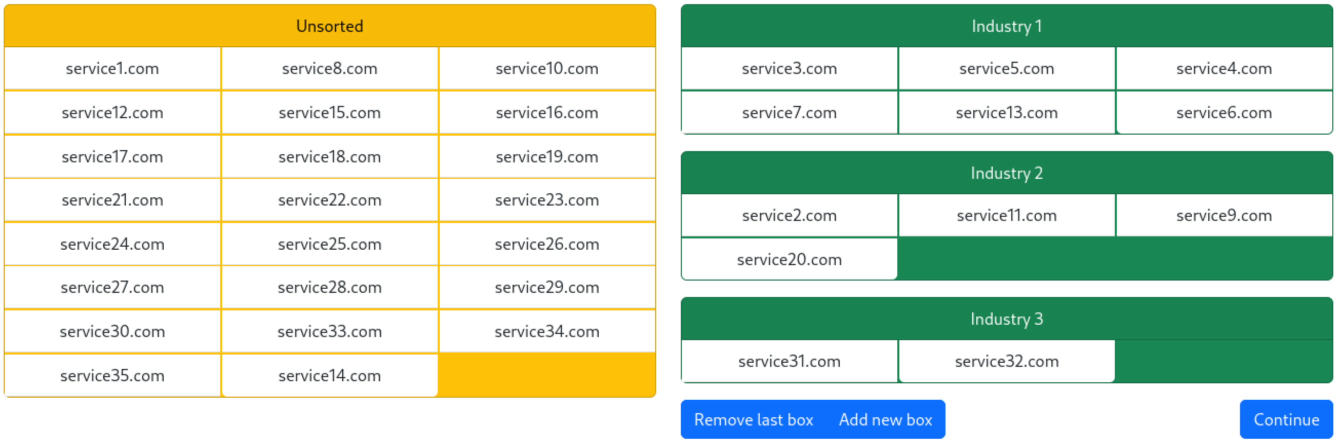
4.1.1. Requesting and Receiving Data Exports. To receive data from the respective online service, users have to make a request under Article 20 GDPR. Table 4 outlines which methods services offer. The number of services that offered dedicated methods (GDPR request form or a button click within the account) was 43% in 2020, 48% in 2021, and 47% in 2022.

In terms of export duration time, we find that the mean duration between requesting and receiving the data export was 9.7 days (2020), 9.5 days (2021), and 8.2 days (2022). The cumulative distribution of durations by year is depicted in Figure 2. While a slight decrease in duration could be observed over the years, the difference is not significant (t = -0.846, p = 0.398).

The means by which services made the data available to us are shown in Table 5. Surprisingly, the number of services that made the data available using a dedicated download portal or e-mail link has decreased from 64 in 2020 to 46 in 2022.

4.1.2. Evaluating GDPR Compliance. After receiving all data exports in the respective year, we analyzed the contents of the exports. We classified the data according to the

[6] The survey raw data as well as the script to generate the clusters can be provided by the authors upon request.



Unsorted		
service1.com	service8.com	service10.com
service12.com	service15.com	service16.com
service17.com	service18.com	service19.com
service21.com	service22.com	service23.com
service24.com	service25.com	service26.com
service27.com	service28.com	service29.com
service30.com	service33.com	service34.com
service35.com	service14.com	

Industry 1		
service3.com	service5.com	service4.com
service7.com	service13.com	service6.com

Industry 2		
service2.com	service11.com	service9.com
service20.com		

Industry 3		
service31.com	service32.com	

Figure 1. Exemplary screenshot of the industry assignment task. Survey participants were asked to sort similar services into the same boxes using drag & drop.

TABLE 3. NUMBER OF SERVICES PER INDUSTRY BY TRANCO RANK (2022).

Rank	1-10	11-100	101-1000	1001-10000	10001-100000	100001-1000000	>1000000	Total
Total	8	25	28	42	26	7	1	137
Retail & E-Commerce	0	3	1	8	5	0	0	17
Entertainment	1	4	3	12	4	0	0	24
Social Networks & Messaging	4	10	7	1	1	0	0	23
Finance & Insurance	0	1	2	0	2	1	0	6
Travel	0	0	4	6	7	4	1	22
Productivity	0	4	3	2	0	0	0	9
Information Retrieval	1	2	4	3	1	0	0	11
Telecommunications, Hosting & E-Mail	1	0	0	4	2	0	0	7
Price Comparison & Marketplace	0	0	0	4	1	0	0	5
Career	0	0	3	0	1	1	0	5
Other	1	1	1	2	2	1	0	8

TABLE 4. OFFERED METHODS FOR REQUESTING DATA EXPORTS UNDER ARTICLE 20 GDPR.

	2020	2021	2022
Button click	29%	31%	28%
E-mail	57%	52%	53%
Dedicated request form	15%	17%	19%

TABLE 5. MODE OF TRANSMISSION OF DATA EXPORTS.

	2020	2021	2022
DL from portal	38%	37%	29%
DL link provided via e-mail	26%	24%	19%
E-mail attachment or body	32%	37%	46%
Written letter	4%	3%	4%

taxonomy by de Hert et al. [16] (see Section 2.2) and evaluated whether the data that is classified as *received data* is complete, i.e., contains all personal information that we provided to the service. In addition, we identified whether the data is in a file format compliant with the provisions of Article 20 GDPR (i.e., common, machine-readable, and structured). We used the evaluation of data formats by Wong and Henderson [17] to classify formats as compliant or non-compliant.

TABLE 6. SHARE OF SERVICES THAT COMPLIANTLY EXPORTED DATA IN 1, 2, OR 3 YEARS. (0 = NEVER)

	0	1	2	3
Data exported	6%	12%	25%	57%
Compliant file format	44%	12%	22%	22%
Overall compliant*	53%	17%	15%	16%

**compliant file format & complete*

TABLE 7. DATA EXPORT SCOPE OF SUCCESSFUL EXPORTS PER YEAR.

	2020	2021	2022
Received data	100%	100%	100%
Observed data	64%	64%	56%
Inferred data	14%	23%	16%

We find that in 2020, 53 (41%), in 2021, 53 (41%), and in 2022, 47 (36%) of services managed to successfully export the data and to provide it in a compliant data format. Evaluating the overall compliance of services (successful data export within 30 days, or 90 days when requesting

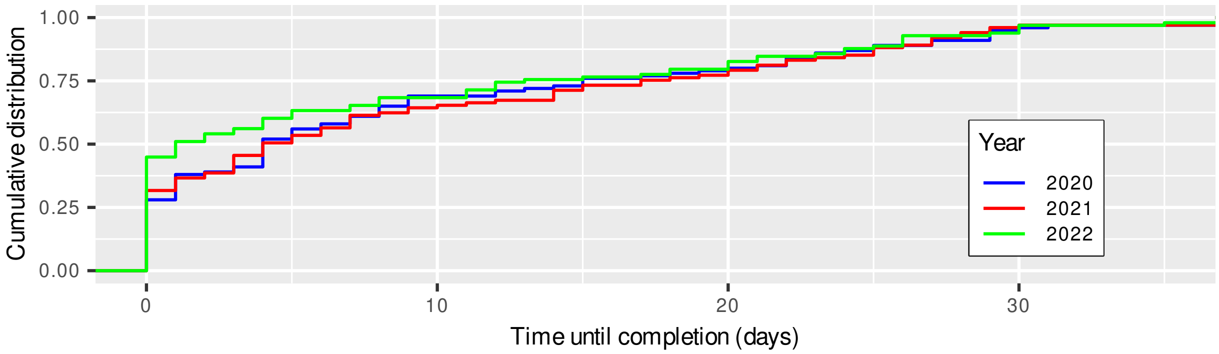


Figure 2. Cumulative distribution of data export duration by year.

an extension, compliant file format, and completeness of data), we find that only 31% (2020), 29% (2021), and 33% (2022) of services are fully compliant with Article 20 GDPR. The difference between years is not significant (t = 0.403, p = 0.687 in an univariate OLS regression of year on compliance). These compliance rates are similar to the 29% of the 182 services in the 2020 cross-sectional dataset [20].

Having panel data from three years allows us to further evaluate in how many years each service has been compliant. This data is shown in Table 6. We find that only 20 services (16%) are compliant in every year we requested the data export. Using a logistic regression model with random effects[7], we find that more popular services significantly more often provide the exports in a compliant format and are significantly more often overall compliant than their less popular competitors. The regression results are shown in columns (1) and (2a) of Table 8.

Moreover, we had to provide one or more means of authentication to most services in order to receive the exports. Two-factor authentication was required by 3 services in 2020, 7 in 2021, and 5 in 2022. Notably, 21 (2020), 30 (2021), and 27 (2022) services did not require any authentication. These services sent us the data to the e-mail address with which we initiated the request without asking for further authentication. Using a linear regression model with random effects, we find that popular services (proxied by their Tranco ranks, where a lower number indicates a higher popularity) require significantly more authentication factors. See column (5) in Table 8 for the regression results.

4.2. Usefulness of Data Transfers

In this paper, we have defined data portability usefulness as the scope of exported data and offered means for importing data.

4.2.1. Scope of Exported Data. Using the taxonomy by de Hert et al. [16], we classified the data exports and verified the presence of *received data*, *observed data*, and *inferred data*. Table 7 shows which types of data were present in successful exports. Note that a data export can (and often does) include more than one type of data. In general, we found that each export that had observed or inferred data also had received data. The number of data types (1: only received, 2: received & observed or received & inferred, 3: received, observed & inferred) does not vary significantly between years (t = -0.689, p = 0.491). Using a linear regression model with random effects, we do, however, find that the export scope (number of provided data types) rises significantly with the services' popularity. The results are shown in column (3a) of Table 8. These findings confirm the significantly positive relationship between export scope and popularity found for the 2020 cross-sectional dataset [20]. Notably, this result is not only robust regarding the time frame but also regarding the popularity ranking, as the baseline study [20] has used the Alexa ranking [23] while this study uses the Tranco ranking [25].

4.2.2. Offered Means for Importing Data. As Article 20 GDPR aims to facilitate data transfer between online services, we investigated whether services make use of the possibility to import data that has been exported at other services. Based on the categorization from the baseline study [20], we assigned each service's import scope the category *None* (no import possible), *Minimal* (only import for minor functionalities available), *Partial* (import for some but not all core functionalities available), or *Full* (import for all core functionalities available).

We find that, in 2020, 96 (70%) services offer no import possibilities at all, 8 (6%) offer minimal import, 22 (16%) offer partial import, and 11 (8%) offer full import. These numbers do not change significantly over the years (t = 0.461, p = 0.645). In fact, only two services that offered no import in 2020 started to offer full import in 2021, and one service that offered no import in 2020 and 2021 started to offer partial import in 2022. In a linear regression model with random effects, we do not find a significant influence of popularity on the offered import scope. See column (4) in Table 8 for the regression results.

[7] As the Hausman test [41] yields insignificant results for all models (p between 0.28 and 0.78), we analyzed our panel data using random effects models.

TABLE 8. Effect of popularity and tracking on GDPR compliance, data scopes, and number of factors required for authentication.

	Dependent variable:						
	Format Comp.	Overall Comp.		Export Scope		Import Scope	Auth. Factors
	panel logistic	*panel logistic*		*panel linear*		*panel linear*	*panel linear*
	(1)	(2a)	(2b)	(3a)	(3b)	(4)	(5)
log(Rank)	-0.237^{**}	-0.202^{*}	-0.222^{*}	-0.087^{***}	-0.126^{***}	-0.042	-0.068^{***}
	(0.111)	(0.118)	(0.133)	(0.021)	(0.026)	(0.027)	(0.016)
Tracking			-0.138^{*}		-0.025^{**}		
			(0.071)		(0.011)		
Ind.: Retail & E-Commerce	-2.586^{*}	-2.670^{*}	-1.566	-0.435^{*}	-0.387	-0.748^{*}	-0.104
	(1.386)	(1.465)	(1.334)	(0.245)	(0.244)	(0.386)	(0.180)
Ind.: Entertainment	-2.791^{**}	-2.574^{*}	-1.525	-0.270	-0.219	-0.530	-0.243
	(1.304)	(1.375)	(1.276)	(0.237)	(0.241)	(0.368)	(0.175)
Ind.: Soc. Netw. & Messaging	-0.325	-0.457	-0.115	-0.112	-0.130	0.038	-0.166
	(1.288)	(1.365)	(1.263)	(0.243)	(0.245)	(0.379)	(0.179)
Ind.: Finance & Insurance	0.401	0.475	0.030	0.051	-0.221	-0.018	0.217
	(1.693)	(1.773)	(2.188)	(0.320)	(0.422)	(0.487)	(0.235)
Ind.: Travel	-2.845^{**}	-2.555^{*}	-1.626	-0.183	-0.391	-0.748^{**}	0.065
	(1.332)	(1.405)	(1.398)	(0.246)	(0.269)	(0.376)	(0.179)
Ind.: Productivity	-1.747	-2.874	-2.098	-0.627^{**}	-0.667^{**}	0.690	-0.418^{*}
	(1.612)	(1.834)	(1.629)	(0.316)	(0.307)	(0.443)	(0.233)
Ind.: Information Retrieval	-1.592	-2.039	-1.663	-0.473^{*}	-0.582^{**}	0.212	-0.239
	(1.395)	(1.529)	(1.430)	(0.285)	(0.286)	(0.421)	(0.211)
Ind.: Telecommunications	-0.556	-2.345	-1.467	-0.649^{**}	-0.571^{*}	0.829^{*}	0.044
	(1.724)	(1.825)	(1.640)	(0.324)	(0.313)	(0.466)	(0.238)
Ind.: Market Place	-1.230	-0.617	0.430	-0.033	0.149	-0.836	0.024
	(1.651)	(1.702)	(1.593)	(0.322)	(0.314)	(0.514)	(0.236)
Ind.: Career	-1.517	-0.867	0.568	0.260	0.457	-0.475	-0.039
	(1.742)	(1.819)	(1.673)	(0.321)	(0.334)	(0.514)	(0.236)
sigma	2.340^{***}	2.459^{***}	2.019^{***}				
	(0.410)	(0.429)	(0.410)				
Constant	2.496^{*}	1.564	1.894	2.622^{***}	2.989^{***}	2.201^{***}	1.356^{***}
	(1.401)	(1.468)	(1.439)	(0.261)	(0.277)	(0.382)	(0.192)
Observations	387	387	313	295	248	411	298
(Pseudo) R^2	0.067	0.053	0.217	0.126	0.209	0.118	0.072
Adjusted R^2	–	–	–	0.092	0.169	0.093	0.036
F Statistic	–	–	–	36.936^{***}	64.445^{***}	53.251^{***}	26.015^{***}

Notes. The table reports the effect of popularity on different characteristics of data transfers under Article 20 GDPR. Columns (2b) and (3b) additionally report the effect of Tracking on compliance and scope. Standard errors are in parentheses below the estimates. *p<0.1; **p<0.05; ***p<0.01
Tranco rank is used as a proxy for popularity (higher rank implies higher popularity, x higher than $y \Leftrightarrow x < y$), logarithmized to account for positive skewness. Tracking indicates the average number of trackers present on the service's website. Dummy variables for the industries are used as control variables.
Format compliance (1) indicates whether format and duration of the data export are compliant with the provisions of the GDPR. Overall compliance (2) further takes completeness of *received data* into account. Export scope (3) takes values from 1 ("no personal info available") to 4 ("received, observed, and inferred data") and indicates how rich the scope of the data export is according to the taxonomy of de Hert et al. [16]. Import scope (4) takes values from 1 ("no import“) to 4 ("full import“) and describes to which extent import possibilities are offered. Authentication factors (5) describes how many factors a person needed to provide to request or access the data export.

4.3. Influence of Tracking on Compliance and Export Scope

Using the longitudinal tracking data from WhoTracks.Me [26], we analyzed whether the prevalence of tracking is associated with a change in compliance with Article 20 GDPR or the scope of the exported data. The medium average number of trackers per service in our sample was 6.088 in 2020, 5.410 in 2021, and 6.585 in 2022. Comparing services with fewer trackers (avg. number of trackers lower than median) to those with more trackers, we find compliance rates of 58% vs. 29% in 2020, 43% vs. 42% in 2021, and 43% vs. 32% in 2022. To verify the relationship between tracking and compliance with Article 20 GDPR, we included *tracking* as an additional independent variable along with rank and popularity in the regression analyses. For overall compliance, we find that services with more trackers are significantly less compliant with the provisions of Article 20 GDPR (column (2b) of Table 8).

Regarding the scope of exported data, we find that a higher tracking prevalence comes with a significantly lower scope of exported data, as shown in column (3b) of Table 8. These findings indicate that while tracking leads to the collection of more personal data that can be classified as *observed data* under the taxonomy of de Hert et al. [16], tracking-intensive online services are not willing to share this data with their users or other online services under the Right to Data Portability.

4.4. Data Portability Practices by Industry

Using the industry classification generated from the survey participants' data, we analyze how compliance, data export scope, and offered import possibilities differ between industries. Table 9 gives an overview of the number of services per industry that have been compliant in 0, 1, 2, or all years of the investigation period. The results from regressions (1) and (2a) in Table 8 imply that services from the industries *Retail & E-Commerce*, *Entertainment*, and *Travel* have significantly lower compliance rates than services from other industries.

TABLE 9. SHARE OF SERVICES THAT WERE COMPLIANT IN 1–3 YEARS PER INDUSTRY. (0 = NEVER)

	0	1	2	3
Retail & E-Commerce	75%	0%	13%	13%
Entertainment	65%	9%	22%	4%
Social Networks & Messaging	22%	26%	26%	26%
Finance & Insurance	40%	0%	20%	40%
Travel	68%	23%	0%	9%
Productivity	67%	17%	0%	17%
Information Retrieval	45%	36%	9%	9%
Telco, Hosting & E-Mail	40%	60%	0%	0%
Price Comparison & Marketplace	40%	0%	60%	0%
Career	40%	20%	20%	20%
Other	50%	0%	0%	50%

Regarding the average export and import scopes per industry and year, in univariate OLS regressions of year on scope per industry, we find a significant rise in export scope between 2020 and 2022 (t = 1.687, p = 0.0968) for *Social Networks & Messaging*. For *Travel*, the export scope sank significantly (t = -2.645, p = 0.0117). For the remaining industries, there were no significant changes in export scope between 2020 and 2022. For the import scope, no industry has seen a significant increase or decrease. A detailed overview of export and import scopes per industry and year can be found in Table 10.

TABLE 10. AVERAGE EXPORT SCOPE (NUMBER OF DATA TYPES) AND IMPORT SCOPE (0: NONE, 1: MINIMAL, 2: PARTIAL, 3: FULL) BY INDUSTRY AND YEAR.

	Export Scope			Import Scope		
	2020	2021	2022	2020	2021	2022
Retail & E-Commerce	1.50	1.53	1.54	0.12	0.12	0.12
Entertainment	1.69	1.88	1.69	0.38	0.38	0.38
Social Netw. & Messaging	1.91	2.24	2.28	0.96	1.09	1.09
Finance & Insurance	2.00	1.80	2.00	0.83	0.83	0.83
Travel	1.87	1.73	1.27	0.05	0.05	0.05
Productivity	1.50	1.80	1.25	1.44	1.78	1.78
Information Retrieval	1.75	1.67	1.60	1.09	1.09	1.28
Telco, Hosting & E-Mail	1.40	1.25	1.25	1.71	1.71	1.71
Price Comp. & Marketplace	2.00	2.00	1.60	0.00	0.00	0.00
Career	2.25	2.20	2.25	0.40	0.40	0.40
Other	2.00	2.00	1.86	0.88	0.88	0.88

Analyzing which industries provide the highest scopes of export and import (regressions (2) and (3) in Table 8), the export scope in the *Retail & E-Commerce*, *Productivity*, *Information Retrieval*, and *Telecommunications, Hosting & E-Mail* industries is significantly lower than in the other industries. For the import scope, being in the *Telecommunications, Hosting & E-Mail* industry is associated with a higher import scope. In comparison, *Retail & E-Commerce* and *Travel* are associated with a lower import scope.

5. Discussion

5.1. On the Stagnated Development Towards Thriving Data Portability in Online Markets

Compared to existing cross-sectional studies on RtDP compliance [17], [20], our longitudinal analysis more reliably confirms the large gap that exists between the policymakers' notion of a thriving and easy-to-use data portability ecosystem and the data transfers that can take place in reality. Many services are not compliant, and data exports are oftentimes severely limited in scope.

One reason for the stagnated development in RtDP compliance might be that Article 20 GDPR seems to be among the least enforced GDPR rights so far: a glance on the *GDPR enforcement tracker* surprisingly demonstrates that six years after the RtDP entered into force only two penalties have been imposed on companies for noncompliance with the RtDP [42]. Another reason could be that companies are not well-informed enough on the RtDP and its practically relevant interpretations [22]. When analyzing the data exports, we found that 14% of services in 2020, 11% in 2021, and 18% in 2022 even confused the RtDP

(Article 20 GDPR) with the Right of Access (Article 15 GDPR), as they explicitly referred to Article 15 in their responses.

5.2. The RtDP as Facilitator to Decentralize Data Assets

Moreover, the RtDP of the EU is equipped with the aspect to support consumers in importing data to competing services and therefore reducing the centralization of data assets in silos among a few incumbents. However, based on our three-year analysis, we do not find evidence that many corporations are taking up this opportunity. We found a very low level of offered import options in 2020 and that only 3 out of 137 services decided to increase their import offerings during the three-year investigation period. We explain this corporate ignorance with a lack of awareness of the chances and opportunities that well-designed data portability regulation can bring to corporate actors. This sheds light on the failed realization of the policymaker's intent to induce decentralization through conceptional regulation.

Furthermore, the current design of the GDPR's RtDP may discourage SMEs from using competitors' data. Long export times, complicated request processes, missing documentation, and a lack of a direct API infrastructure might negatively influence services' willingness to embrace the RtDP. Consistent and reliable data export processes can facilitate the development of data import solutions and lower the barrier for consumers to use the RtDP. As user attention is considered a scarce resource [43], the RtDP could benefit from data exports that are more easily comprehensible.

Given the lack of significant progress in RtDP compliance and acceptance over three years, we propose refining the GDPR's RtDP. Our findings indicate that less popular services or SMEs struggle to implement data portability in a compliant way, raising the question of whether SMEs should face the same regulatory demands as larger companies. Differentiating based on company turnover could reduce market entry barriers for SMEs, aligning with the original goal of data portability regulation. This approach would allow SMEs to focus on offering more import options for consumers and participate in initiatives like the data transfer project to simplify imports. Additionally, RtDP regulation could benefit from decentralization through architectural regulation [34] to ensure easier data importability. We identify four technical directions to achieve this:

- Firstly, policymakers could promote the enhancement of direct API infrastructure platforms [44]. In this regard, corporations could already build on a stable first version of the *Data Transfer Project* platform. This project is maintained by the nonprofit organization *Data Transfer Initiative* [33], which was created by Google, Meta, and Apple, and may therefore reflect the interests of these major market participants.
- Secondly, for-profit companies could be encouraged to build industry-specific direct data portability solutions with the prospect of marketing them to competitors as well as individuals. The French firm *Soundiiz* serves in this context as an illustrative example. It created a direct portability tool for music streaming services, which allows the transfer of data categories such as playlists and favorite marks. Soundiiz sells its solution to competitors, e.g., Tidal, as well as individuals [45].
- Thirdly, policymakers could oblige online services to implement open protocols and service gateways, as pioneered by federated networks. The open protocol *ActivityPub* is probably the most prominent example in this regard. For instance, the online social networks Mastodon and Friendica build on this protocol and enable their users to communicate across service boundaries [46].
- Fourthly, semantic web initiatives such as the project *Solid* [4] could lead the way to unlock user data silos of market-leading incumbents. By providing each individual with a *data pod*, individuals no longer need to store their information on corporate servers but can put their data on their own secure data space to which companies are only granted access upon request. Like the Data Transfer Project, Solid is operational (as of 2024) and users can start engaging with it.

5.3. Factors Influencing the Willingness to Comply with Data Portability Regulation

First and foremost, service popularity is one of the key influences on compliance with an RtDP. It motivates companies to provide more data to individuals and use better authentication for RtDP requests. On the other hand, the presence of trackers on a service's website has been shown to negatively influence both RtDP compliance and the scope of exported data. Popular, data-driven services broadly interpret the provisions of the RtDP, share a large scope of data, and, thus, contribute to the usefulness of data portability. We attribute this finding to their popularity rather than their data-driven business model. These services are in the regulator's spotlight and, therefore, are wary of potentially high fines for GDPR violations. In contrast, services with a high presence of trackers but lower popularity seem to interpret the provision of the RtDP more narrowly. As these services collect data through tracking – classified as observed data according to the taxonomy by de Hert et al. [16] – they could export such data upon request under the RtDP, but choose not to do so. They regard their unique datasets as important assets and are unwilling to hand a copy away to users or competitors. We, therefore, attribute the low compliance rates with Article 20 GDPR to a lack of willingness and interest in the regulation that is not challenged by stricter enforcement.

5.4. A Refinement of the RtDP Should Make It Less Generic and More Scenario-Based

Our data on the adoption of data portability among industries implies that for some industries, data portability regulation is more useful than for others. One notable example is the *Social Media & Messaging* industry. Here, we find a comparatively high adoption of both export and import offers. Regulators should take data from consumer studies (e.g., [21], [47]) on switching intentions of online users into account when refining data portability regulation. It becomes evident that users have particular switching scenarios in mind where they would prefer to use data portability, e.g., to transfer photos between social network providers. In other industries, data portability scenarios are less obvious, such as in the *Retail & E-Commerce* industry. The lack of realistic data portability scenarios in some industries brings up the question of whether data portability regulation should affect companies in all industries or rather focus on companies in certain, e.g., only in data-driven industries. Further empirical user studies are necessary to identify feasible data transfer scenarios within and across industries.

6. Limitations and Future Work

As any scientific work, this study is subject to limitations that should inspire future research. First, the empirical analysis of the data exports and imports was conducted with a focus on the data portability regulation of the EU GDPR. As data portability regulations of other jurisdictions differ from the European rules, conducting a comparative longitudinal study on data portability regulation compliance across jurisdictions might be insightful.

Second, as described in Section 3.1, our sample is partly a convenience sample. When starting the data collection process in 2020, we faced a trade-off between having a representative sample and having "real" data. To keep the data exports as authentic as possible, we opted to use our existing accounts and to create new accounts for services in the German Alexa Top 100 ranking, where none of the authors had an account yet. Due to the location of the authors, 50 of 137 services (36.5%) in the sample are based in Germany. Therefore, data portability practices of German companies are overrepresented in our study. For practical reasons (language barrier), services in languages different from English or German, and for ethical reasons (due to the involvement of student assistants), adult services were excluded. While we tried to keep data creation and collection for the newly created accounts as realistic as possible (see Appendix A.1.1), the possibility of data exports from new accounts being systematically different from those from existing accounts cannot be fully eliminated. However, for the cases in our dataset where data in years 1&3 were from existing accounts and in year 2 from a new account or vice versa, we found no occurrence where data scope or compliance in year 2 was fundamentally different from the other years. Future research could replicate our study with a representative sample, e.g., by creating a random sample from the Tranco ranking using a geometric distribution.

Third, due to its established methodology of data collection and the availability of longitudinal data, whotracks.me [26], [36] was used as source for the prevalence of third-party trackers. While this method allows us to draw connections between the prevalence of tracking and data portability practices, it comes with the disadvantage of not being able to explicitly analyze the data collected via tracking. Future research could close this gap by requesting data exports under Article 20 GDPR not only from first parties but also from third parties embedded in the first parties' websites.

Fourth, our study is concerned mainly with indirect data portability (Article 20(1) GDPR), while Article 20(2) also gives the possibility of direct data transfers between online services. Once direct data portability between major online services becomes more widespread technically (e.g., by means of the Data Transfer Project [44]), a replication study on various measures of this paper, such as scope or authentication, will be very interesting. Under such circumstances, services might consider more strongly aspects of competition and innovation based on higher user data availability, and they might show different export behavior once they know that another company and not an individual is the addressee.

7. Concluding Remarks

Over five years ago, in 2018, a jurisdiction (EU) published a consumer right to data portability (RtDP) for the first time, and three jurisdictions (California, Brazil, and China) passed similar consumer rights since then. This dynamic development gives reason to believe that these large jurisdictions perceive an RtDP as a chance to empower consumers to have more control over their personal data, facilitate decentralization, and foster competition in today's often concentrated digital online markets. However, does the outcome of the privacy regulation live up to the expectations of policymakers? And how can RtDPs be refined and further improved in the future?

This empirical study provides the first longitudinal evidence on the effectiveness of data portability regulation. From 2020 to 2022, we monitored the data export and import practices of 129 online services under the EU's GDPR. This allows us to show how online services, on the one hand, have dealt with this legal obligation and, on the other hand, whether they used this intriguing chance to attract consumer data and customers. Our unique dataset shows that compliance with the minimum requirements of the EU's RtDP has been low over the years of observation. Moreover, companies did not significantly increase data export scope or import options over time. This suggests that smaller competitors have not yet leveraged data portability regulation to attract new customers and their data. Beyond that, we can replicate and strengthen the results of our prior cross-sectional study [20] through our panel data that more popular services are more compliant, provide a higher export scope, and use more authentication factors to verify the requester's identity. We find that on average, online services that feature a higher presence of trackers have been less compliant and provided a lower data scope than their more privacy-friendly counterparts. While fewer than 6% of services required two-factor authentication for data exports each year, 16% to 23% sent exports via email without any authentication. We attribute differences in compliance and data scope across services and industries not just to a lack of exportable data but also to an unwillingness to make personal data available.

Acknowledgments

We would like to thank Can Bilgin, Yaren Dalgic, Maximilian J. Frank, Leon Jäkel, Paul Pizzinini, Deniz Yildiz, and Simon Zilker for their valuable research assistance, the participants of our survey for contributing their time, and the anonymous reviewers (including our shepherd) for their helpful feedback. We further are grateful for partial funding support from the Bavarian Research Institute for Digital Transformation (bidt). Responsibility for the contents of this publication rests with the authors.

References

[1] U. Bojars, A. Passant, J. G. Breslin, and S. Decker, "Social Network and Data Portability using Semantic Web Technologies," in *BIS 2008 Workshops Proceedings: Social Aspects of the Web (SAW 2008), Advances in Accessing Deep Web (ADW 2008), E-Learning for Business Needs, Innsbruck, Austria, 6-7 May 2008*, ser. CEUR Workshop Proceedings, D. Flejter, S. Grzonkowski, T. Kaczmarek, M. Kowalkiewicz, T. Nagle, and J. Parkes, Eds., vol. 333. CEUR-WS.org, 2008, pp. 5–19. [Online]. Available: https://ceur-ws.org/Vol-333/saw1.pdf

[2] I. Graef, J. Verschakelen, and P. Valcke, "Putting the Right to Data Portability into a Competition Law Perspective," *Law: The Journal of the Higher School of Economics*, pp. 53–63, 2013. [Online]. Available: https://papers.ssrn.com/abstract=2416537

[3] DLA Piper, *Data Protection Laws of the World*, 2022. [Online]. Available: https://www.dlapiperdataprotection.com/system/modules/za.co.heliosdesign.dla.lotw.data_protection/functions/handbook.pdf?country=all

[4] S. Capadisli, T. Berners-Lee, and K. Kjernsmo, "Solid Protocol," https://solidproject.org/TR/protocol, 2024, Editor's Draft, Version 0.12.0, Accessed: 2024-09-19.

[5] M. Wohlfarth, "Data Portability on the Internet: An Economic Analysis," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 551–574, 2019. [Online]. Available: https://doi.org/10.1007/s12599-019-00580-9

[6] A. Acquisti and J. Grossklags, "Privacy and Rationality in Individual Decision Making," *IEEE Security & Privacy*, vol. 3, no. 1, pp. 26–33, 2005. [Online]. Available: https://doi.org/10.1109/MSP.2005.22

[7] D. J. Solove, "Murky Consent: An Approach to the Fictions of Consent in Privacy Law," *Boston University Law Review*, vol. 104, no. 2, pp. 593–639, 2024. [Online]. Available: https://www.bu.edu/bulawreview/files/2024/04/SOLOVE.pdf

[8] W. Melicher, M. Sharif, J. Tan, L. Bauer, M. Christodorescu, and P. G. Leon, "(Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 2, pp. 135–154, 2016. [Online]. Available: https://doi.org/10.1515/popets-2016-0009

[9] S. Zimmeck, O. Wang, K. Alicki, J. Wang, and S. Eng, "Usability and Enforceability of Global Privacy Control," *Proceedings on Privacy Enhancing Technologies*, vol. 2023, no. 2, pp. 1–17, 2023. [Online]. Available: https://doi.org/10.56553/popets-2023-0052

[10] H. R. Lipford, A. Besmer, and J. Watson, "Understanding Privacy Settings in Facebook With an Audience View," in *Proceedings of the 1st Conference on Usability, Psychology, and Security*, 2008, pp. 1–8. [Online]. Available: https://www.usenix.org/legacy/event/upsec08/tech/full_papers/lipford/lipford_html/

[11] P. Bahirat, Y. He, A. Menon, and B. Knijnenburg, "A Data-Driven Approach to Developing IoT Privacy-Setting Interfaces," in *Proceedings of the 23rd International Conference on Intelligent User Interfaces*, ser. IUI '18. ACM, 2018, pp. 165–176. [Online]. Available: https://doi.org/10.1145/3172944.3172982

[12] P. Wijesekera, A. Baokar, L. Tsai, J. Reardon, S. Egelman, D. Wagner, and K. Beznosov, "The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences," in *2017 IEEE Symposium on Security and Privacy (SP)*, 2017, pp. 1077–1093. [Online]. Available: https://doi.org/10.1109/SP.2017.51

[13] P. Swire and Y. Lagos, "Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique," *Maryland Law Review*, vol. 72, no. 2, pp. 335–380, 2013. [Online]. Available: https://digitalcommons.law.umaryland.edu/mlr/vol72/iss2/1

[14] J. Krämer and N. Stüdlein, "Data Portability, Data Disclosure and Data-induced Switching Costs: Some Unintended Consequences of the General Data Protection Regulation," *Economics Letters*, vol. 181, pp. 99–103, 2019. [Online]. Available: https://doi.org/10.1016/j.econlet.2019.05.015

[15] W. M. W. Lam and X. Liu, "Does Data Portability Facilitate Entry?" *International Journal of Industrial Organization*, vol. 69, 2020. [Online]. Available: https://doi.org/10.1016/j.ijindorg.2019.102564

[16] P. De Hert, V. Papakonstantinou, G. Malgieri, L. Beslay, and I. Sanchez, "The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services," *Computer Law & Security Review*, vol. 34, no. 2, pp. 193–203, 2018. [Online]. Available: https://doi.org/10.1016/j.clsr.2017.10.003

[17] J. Wong and T. Henderson, "The Right to Data Portability in Practice: Exploring the Implications of the Technologically Neutral GDPR," *International Data Privacy Law*, vol. 9, no. 3, pp. 173–191, 2019. [Online]. Available: https://doi.org/10.1093/idpl/ipz008

[18] D. Copetti Cravo, "How to Make Data Portability Right More Meaningful for Data Subjects?" *European Data Protection Law Review*, vol. 8, no. 1, pp. 52–60, 2022. [Online]. Available: https://doi.org/10.21552/edpl/2022/1/9

[19] S. Kuebler-Wachendorff, R. Luzsa, J. Kranz, S. Mager, E. Syrmoudis, S. Mayr, and J. Grossklags, "The Right to Data Portability: Conception, Status Quo, and Future Directions," *Informatik Spektrum*, vol. 44, no. 4, pp. 264–272, 2021. [Online]. Available: https://doi.org/10.1007/s00287-021-01372-w

[20] E. Syrmoudis, S. Mager, S. Kuebler-Wachendorff, P. Pizzinini, J. Grossklags, and J. Kranz, "Data Portability between Online Services: An Empirical Analysis on the Effectiveness of GDPR Art. 20," *Proceedings on Privacy Enhancing Technologies*, vol. 2021, no. 3, pp. 351–372, 2021. [Online]. Available: https://doi.org/10.2478/popets-2021-0051

[21] R. Luzsa, S. Mayr, E. Syrmoudis, J. Grossklags, S. Kübler-Wachendorff, and J. Kranz, "Online Service Switching Intentions and Attitudes Towards Data Portability – The Role of Technology-related Attitudes and Privacy," in *Mensch und Computer 2022*. ACM, 2022, pp. 1–13. [Online]. Available: https://doi.org/10.1145/3543758.3543762

[22] Article 29 Data Protection Working Party, "Guidelines on the Right to Data Portability," Guideline WP 242 rev.01, 2017. [Online]. Available: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233

[23] Alexa, "Topsites in Germany – Ranking," https://web.archive.org/web/20201101013937/https://www.alexa.com/topsites/countries/DE, 2020, archived from https://www.alexa.com/topsites/countries/DE.

[24] Bureau van Dijk, "Orbis," https://orbis.bvdinfo.com/, 2023, accessed: 2024-09-19.

[25] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, "TRANCO: A Research-Oriented Top Sites Ranking Hardened Against Manipulation," in *Proceedings of the 26th Annual Network and Distributed System Security Symposium*, ser. NDSS. Internet Society, 2019. [Online]. Available: https://dx.doi.org/10.14722/ndss.2019.23386

[26] A. Karaj, S. Macbeth, R. Berson, and J. M. Pujol, “WhoTracks.Me: Shedding Light on the Opaque World of Online Tracking,” arXiv, Tech. Rep. 1804.08959, 2018. [Online]. Available: https://doi.org/10.48550/arXiv.1804.08959

[27] *NACE Rev. 2 - Statistical Classification of Economic Activities in the European Community*, Eurostat, Luxembourg, 2008. [Online]. Available: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-07-015

[28] Similarweb, “Similarweb – Digital Intelligence Platform,” https://www.similarweb.com, 2023, accessed: 2024-09-19.

[29] Wikipedia contributors, “Wikipedia, The Free Encyclopedia,” https://en.wikipedia.org/wiki/Main_Page, 2023, accessed: 2024-09-19.

[30] A. Clauset, M. E. J. Newman, and C. Moore, “Finding Community Structure in Very Large Networks,” *Physical Review E*, vol. 70, no. 6, 2004. [Online]. Available: https://doi.org/10.1103/PhysRevE.70.066111

[31] J. Krämer, “Personal Data Portability in the Platform Economy: Economic Implications and Policy Recommendations,” *Journal of Competition Law & Economics*, vol. 17, no. 2, pp. 263–308, 2020. [Online]. Available: https://doi.org/10.1093/joclec/nhaa030

[32] B. Engels, “Data Portability Among Online Platforms,” *Internet Policy Review*, vol. 5, no. 2, 2016. [Online]. Available: https://doi.org/10.14763/2016.2.408

[33] C. Riley, “Data Transfer Initiative Launch,” https://dtinit.org/blog/2023/03/28/launch, accessed: 2024-09-19.

[34] J. Kranz, S. Kuebler-Wachendorff, E. Syrmoudis, J. Grossklags, S. Mager, R. Luzsa, and S. Mayr, “Data Portability,” *Business & Information Systems Engineering*, vol. 65, pp. 597–607, 2023. [Online]. Available: https://doi.org/10.1007/s12599-023-00815-w

[35] M. Kendall, *Rank Correlation Methods*. Oxford, England: Griffin, 1948.

[36] WhoTracks.Me, “Github: WhoTracks.Me,” 2024, [Original Date: 2017-10-19]. [Online]. Available: https://github.com/whotracksme/whotracks.me/blob/master/whotracksme/data/Readme.md

[37] C. Righi, J. James, M. Beasley, D. L. Day, J. E. Fox, J. Gieber, C. Howe, and L. Ruby, “Card Sort Analysis Best Practices,” *Journal of Usability Studies*, vol. 8, no. 3, pp. 69–89, 2013. [Online]. Available: https://uxpajournal.org/card-sort-analysis-best-practices-2/

[38] J. R. Wood and L. E. Wood, “Card Sorting: Current Practices and Beyond,” *Journal of Usability Studies*, vol. 4, no. 1, pp. 1–6, 2008. [Online]. Available: https://uxpajournal.org/card-sorting-current-practices-and-beyond/

[39] J. L. Kröger, J. Lindemann, and D. Herrmann, “How Do App Vendors Respond to Subject Access Requests? A Longitudinal Privacy Study on iOS and Android Apps,” in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, ser. ARES ’20. ACM, 2020, pp. 1–10. [Online]. Available: https://doi.org/10.1145/3407023.3407057

[40] European Commission, “Special Eurobarometer 487a,” Report, 2019. [Online]. Available: http://dx.doi.org/10.2838/579882

[41] J. A. Hausman, “Specification Tests in Econometrics,” *Econometrica*, vol. 46, no. 6, pp. 1251–1271, 1978. [Online]. Available: https://www.jstor.org/stable/1913827

[42] CMS.Law, “GDPR Enforcement Tracker – List of GDPR fines,” https://www.enforcementtracker.com, 2023, accessed: 2024-09-19.

[43] R. Böhme and J. Grossklags, “The Security Cost of Cheap User Interaction,” in *Proceedings of the 2011 New Security Paradigms Workshop*, ser. NSPW ’11. New York, NY, USA: Association for Computing Machinery, 2011, p. 67–82. [Online]. Available: https://doi.org/10.1145/2073276.2073284

[44] B. Willard, J. Chavez, G. Fair, K. Levine, A. Lange, and J. Dickerson, “Data Transfer Project: From Theory to Practice,” Tech. Rep., 2018. [Online]. Available: https://services.google.com/fh/files/blogs/data-transfer-project-google-whitepaper-v4.pdf

[45] M. Zeff, “These Two Friends Built a Simple Tool to Transfer Playlists Between Apple Music and Spotify, and It Works Great,” https://techcrunch.com/2024/09/14/these-two-friends-built-a-simple-tool-to-transfer-playlists-between-apple-music-and-spotify-and-it-works-great/, accessed: 2024-09-19.

[46] C. Lemmer-Webber, J. Tallon, E. Shepherd, A. Guy, and E. Prodromou, “ActivityPub – W3C Recommendation 23 January 2018,” https://www.w3.org/TR/activitypub/, accessed: 2024-09-19.

[47] E. Syrmoudis, R. Luzsa, Y. Ehrlich, D. Agidigbi, K. Kirsch, D. Rudolf, D. Schlaeger, J. Weber, and J. Grossklags, “Unlocking Personal Data From Online Services: User Studies on Data Export Experiences and Data Transfer Scenarios,” *Human–Computer Interaction*, pp. 1–25, 2024. [Online]. Available: https://doi.org/10.1080/07370024.2024.2325347

[48] G. Nicholas, “Taking It With You: Platform Barriers to Entry and the Limits of Data Portability,” *Michigan Technology Law Review*, vol. 27, no. 2, pp. 263–298, 2021. [Online]. Available: https://repository.law.umich.edu/mtlr/vol27/iss2/3/

[49] N. Samarin, S. Kothari, Z. Siyed, O. Bjorkman, R. Yuan, P. Wijesekera, N. Alomar, J. Fischer, C. Hoofnagle, and S. Egelman, “Lessons in VCR Repair: Compliance of Android App Developers with the California Consumer Privacy Act (CCPA),” *Proceedings on Privacy Enhancing Technologies*, vol. 2023, no. 3, pp. 103–121, 2023. [Online]. Available: https://doi.org/10.56553/popets-2023-0072

[50] S. Veys, D. Serrano, M. Stamos, M. Herman, N. Reitinger, M. L. Mazurek, and B. Ur, “Pursuing Usable and Useful Data Downloads Under GDPR/CCPA Access Rights via Co-Design,” in *Seventeenth Symposium on Usable Privacy and Security*, ser. SOUPS ’21. USENIX Association, 2021, pp. 217–242. [Online]. Available: https://www.usenix.org/conference/soups2021/presentation/veys

[51] A. Erickson, “Comparative Analysis of the EU’s GDPR and Brazil’s LGPD: Enforcement Challenges with the LGPD,” *Brooklyn Journal of International Law*, vol. 44, no. 2, pp. 859–888, 2019. [Online]. Available: https://brooklynworks.brooklaw.edu/bjil/vol44/iss2/9

[52] L. B. Cano, S. Raza, B. Sekibo, A. Siafaras, Q. Wolf, Z. Yin, W. Xu, and D. Tuncer, “A Comparative Perspective of Data Regulation Frameworks and Their Implications for Connected Vehicles,” in *Proceedings of the ACM SIGCOMM Joint Workshops on Technologies, Applications, and Uses of a Responsible Internet and Building Greener Internet*, ser. TAURIN+BGI ’22. ACM, 2022, pp. 6–11. [Online]. Available: https://doi.org/10.1145/3538395.3545313

[53] I. Calzada, “Citizens & Data Privacy in China: The State of the Art of the Personal Information Protection Law (PIPL),” *Smart Cities*, vol. 5, no. 3, pp. 1129–1150, 2022. [Online]. Available: https://doi.org/10.3390/smartcities5030057

Appendix A.
Labeling Schemes for Data Export Requests and Data Imports

A.1. Data exports

A.1.1. Instructions.

1. (Only if account is newly created:) Use all functions of service, create received & observed data wherever possible.
2. Request data export, try in this order: Use button or request form to trigger export, consult service’s privacy policy on GDPR Art. 20 requests, send mail/message to data protection officer (“Datenschutzbeauftragte*r”), send mail/message to support.
3. Answer any questions / requests from service.

4. Receive data export, analyze it according to Section 3.2 (measures) of [20].

IMPORTANT: If you observe something that seems unusual or whenever things are ambiguous (can be interpreted one way or another), make a note.

A.1.2. Attributes.

- Service
  - Format: String
  - Description: Name of service
- Status
  - Format: Integer (0-2)
  - Description:
    - 0: data export successful
    - 1: service did respond but export not successful (e.g., not within 1(+2) months)
    - 2: service never responded
- Request date
  - Format: YYYY-MM-DD
  - Description: Date the export under Art. 20 was requested
- Received date
  - Format: YYYY-MM-DD
  - Description: Date the data export was received
- Duration
  - Format: Integer
  - Description: Received date - Request date in days (0-90)
- Completeness
  - Format: Boolean
  - Description:
    - TRUE: data contains all "received data"
    - FALSE: some or all "received data" is missing
- Missing
  - Format: String
  - Description: Comma-separated list of missing data categories
- Scope
  - Format: Integer
  - Description: Data scope according to taxonomy of De Hert et al. (2018)
    - 0: no personal info available
    - 1: received data
    - 2: received & observed data
    - 3: received, observed & inferred data
    - -1: none of these categories apply (e.g., there is also predicted data or received & inferred but no observed data, provide description in parentheses)
- Scope Info
  - Format: String
  - Description: Comma-separated list of data categories contained in export
- Art. 15
  - Format: Boolean
  - Description:
    - TRUE: service confused request with Art. 15 (only when Art. 15 is explicitly mentioned)
    - FALSE: no confusion
- Format
  - Format: String
  - Description: File format of the data export, if more than one: format of the metadata
- Additional formats
  - Format: String
  - Description: Other file formats contained in the export
- Request procedure
  - Format: String ("button click", "email", or "request form")
  - Description: How the data export was requested.
    - "button click": export was requested by clicking on a dedicated button provided, e.g., on the user interface
    - "email": an email was sent to the service, e.g., to its data protection officer
    - "request form": a dedicated request form, provided by the service, was filled out
- Authentication
  - Format: String
  - Description: Verbose description of authentication procedure
- Login
  - Format: Boolean
  - Description:
    - TRUE: login was required to request / access data export
    - FALSE: no login required
- Mail confirmation
  - Format: Boolean
  - Description:
    - TRUE: an email address, which is linked to the account, needed to be confirmed to request / access the export, e.g., by clicking a link, entering a code, or answering a mail from this address
    - FALSE: no email confirmation required
- Phone confirmation
  - Format: Boolean
  - Description:
    - TRUE: a phone number, which is linked to the account, needed to be confirmed to request / access the export, e.g. by clicking a link, entering a code received via SMS, or answering a phone call
    - FALSE: no phone confirmation required
- Letter confirmation
  - Format: Boolean
  - Description:

    - TRUE: the user's address needed to be confirmed to request / access the data export, e.g., by entering a code received in a letter, or by receiving a registered letter ("Einschreiben")
    - FALSE: no letter confirmation needed
- Personal information
  - Format: Boolean
  - Description:
    - TRUE: user needed to provide additional personal information, e.g., date of birth, to request / access data export
    - FALSE: no additional personal information needed to be provided
- ID
  - Format: Boolean
  - Description:
    - TRUE: identification needed to be provided to request / access the data export, e.g., by sending a copy of the ID card, by showing it in a video call, or by showing it in a post office branch ("POSTIDENT")
    - FALSE: no ID needed to by provided
- Transmission way
  - Format: String ("email", "download link via email", "download from portal", or "letter")
  - Description: How the data export was received / could be obtained
    - "email": data was sent as an email attachment or within the email body
    - "download link via email": data could be downloaded from a link that was sent via email
    - "download from portal": data could be downloaded from a portal, e.g. by clicking a link on the services user interface
    - "letter": data was sent as a letter
- Additional info
  - Format: String
  - Description: Other notable observations

A.2. Data imports

A.2.1. Instructions.

1. Identify 1-5 core functionalities of the service.
2. Find ways to import data / migrate data.
   1. Inspect the service's user interface for ways to import data / migrate data from other services.
   2. Inspect the service's documentation / help files for mentions of data import / migration.
   3. Use a search engine and search for the service's name in combination with import / upload / migration / data portability and check the first five results, respectively.
3. List all providers that can be used to import data via OAuth (usually via a "Login with . . . " button, sometimes there is a "Connect to . . . " in the settings).

IMPORTANT: If you observe something that seems unusual or whenever things are ambiguous (can be interpreted one way or another), make a note.

A.2.2. Attributes.

- Service
  - Format: String
  - Description: Name of service
- Core functionalities
  - Format: String
  - Description: Comma-separated list of the service's most important functionalities (e.g., mail, messaging, video streaming, banking, . . .)
- General import
  - Format: Integer (0-4)
  - Description: Import scope
    - 0: no data import possible
    - 1: there are data import possibilities but not for core functionalities
    - 2: there are data import possibilities for some but not all core functionalities
    - 3: there are data import possibilities for all core functionalities
- General import scope
  - Format: String
  - Description: Comma-separated list of all types of data that can be imported
- Additional import info
  - Format: String
  - Description: Detailed verbose description of how which data can be imported from where
- OAuth count
  - Format: Integer
  - Description: Number of providers from which data can be imported via OAauth (usually using a "Login with xyz" button)
- OAuth services
  - Format: String
  - Description: Comma-separated list of supported OAuth data providers
- Additional info
  - Format: String
  - Description: Other notable observations

Appendix B. Overview of Established Rights to Data Portability (RtDP) in Privacy Regulations

As the first jurisdiction, the European Union enacted a RtDP with Article 20 GDPR in May 2018. It consists of two parts: First, Article 20(1) grants consumers the right to request a set of their personal data from any online service and receive it at their premises within a maximum of 3 months. Second, Article 20(2) further enables them to request a

TABLE 11. DATA PORTABILITY RIGHTS (RTDP) IN PRIVACY REGULATIONS

	European Union	California (USA)	Brazil	China
Name of Regulation	General Data Protection Regulation (GDPR)	California Consumer Privacy Act (CCPA)	Lei Geral de Proteção de Dados (LGPD)	Personal Information Protection Law (PIPL)
Article/Section of:				
(1) Data Portability legislation	Article 20	Section 1798.100	Article 18	Article 45
(2) Data Access legislation	Article 15	Section 1798.100	Article 18	Article 45
Date of Enactment	25.05.2018	01.01.2020	16.08.2020	01.11.2021
Regulation covers RtDP aspect of:				
(1) transfer to own premises?	yes, in Article 20.1	yes, in Section 1798.100	no	yes, in Article 45
(2) direct transfer to another service?	yes, in Article 20.2	no	yes, in Article 18 (V)	yes, in Article 45
Export needs to be machine-readable?	yes	no, but is recommended	no	no
RtDP response time	1 month (extension of 2 months possible)	45 days (extension of 45 days possible)	15 days	no response time set
Export request limitation	no	yes, twice per year	no	no

direct transfer to another online service of their choice [22]. Thus, while the EU foremost claimed to concentrate on data portability's potential to create higher consumer well-being through more user control over one's own personal data, it also provided a right addressing data portability's potential to impact competition in digital markets with Article 20(2).

Since then, more jurisdictions enacted modern, comprehensive privacy regulations that include either one or two of the aspects of the RtDP (see Table 11):

Section 1798.100 (4) of the *California Consumer Privacy Act (CCPA)* allows citizens of this US state to exercise a right similar to GDPR 20(1), in which they can request personal data from any corporation to their premises twice a year starting from January 2020. Although recommended, it does not need to be in machine-readable format – as it is coupled with the Right of Access (RoA) – and the competition promoting direct transfer to another online service is not granted in the law [48]–[50].

Since 2020, Article 18 (V) of the Brazilian *Lei Geral de Proteção de Dados (LGPD)* explicitly establishes a RtDP for direct transfer to another provider, but not for an export to one's own premises. Whereas the mandated response time of 15 days for data controllers is more strict than in the GDPR and CCPA, there are no requirements made on the data format aspect of machine-readability [51].

Article 45 of the Chinese *Personal Information Protection Law (PIPL)* includes both aspects of the GDPR's RtDP: the export of personal data to one's own premises as well as the transfer of them to another online service upon the consumer's request. As the most recently enacted RtDP, it lacks information and an interpretation of the duration of corporate response time [52] and on the question of whether machine-readability is demanded as a form factor [53].

Considering the similarities in designs of the RtDPs, we see that the results of this first empirical longitudinal analysis on a right to data portability across industries are of high informative value for stakeholders beyond the EU as well.