Research Report

GDPR Data Portability: The Forgotten Right

An in-depth review of the legislative arrangements of the right to data portability under the General Protection Data Regulation Article 20, and an assessment of its implementation in real-world practice.

www.gdpr.dev

ALIAS

Authors

Stéphanie Exposito-Rosso
François-Xavier Cao
Antoine Piquet
Mehdi Medjaoui

Design by www.bernatfont.com

Content

Introduction

Is the most advanced data right generating value for all?

If there is one innovation that the General Data Protection Regulation (GDPR) has enabled, it is the portability of personal data. As the term 'portability' suggests, it is the right for citizens and businesses to request the transfer of their personal data from one service or platform to another for reuse. It was heralded by European policy leaders as a significant new digital right.

We now live in the digital as well as the physical world, and we increasingly work in a digital economy. And in Europe, which operates as a single digital market, we are all stakeholders and participants.

In such a context, data portability is at the core of enabling us to co-create our own value by giving us access to our data and enabling us to determine how it is used.

Data represents our digital footprint – it is the sum of our interactions and can be considered as our digital labour. We generate valuable data through our online and mobile interactions, and whenever we use digital services. And as we maintain and build our digital lives, this data accumulates and can be considered as our digital capital.

Our right to data portability under the GDPR, therefore, becomes our right to share our digital capital with the partners, applications and platforms that we choose.

For the platforms and applications we share our data with, the sum of all of the data from everyone's contributions is more valuable than the data each of us holds. This accumulated data can drive new innovation, enable European businesses to grow, and support local economic development.

The Right to GDPR Data Portability: Ambition, Myths, and Realities

The political origins of the GDPR

To understand the essence of the right to data portability and its implications, it is necessary to briefly recall the context in which the GDPR itself was adopted and applied.

Personal data, and the use of the internet in general, began raising new challenges in managing privacy and individual liberties. The need to establish boundaries and a regulatory context for personal data particularly arose after an in-depth investigation into the privacy policies of Facebook. In the early 2010s, Facebook was collecting a staggering amount of personal data – at the time it processed the data of 700 million users – and it was seen as having a potentially out-of-balance impact on citizens.

European policies, notably Directive 2012/0011 (COD), sought to update the old European framework stemming from Directive 95/46, that had become somewhat obsolete in an era of technological advances and monetisation of individual data.

However, these new policies only introduced a somewhat belated awareness about the need for an internet law to manage how companies benefit from personal data. They led to fairly ineffective sanctions aimed at controlling these emerging tech giants, whose power and influence was often seen as equivalent to small countries. In fact, the development of the new laws even resulted in ambassadors being sent to the US-based headquarters of these companies for negotiations.

In order to regain a certain European sovereignty, in the face of these American titans, the idea of continental regulation over the use of personal data arrived on the agenda: a foundation that was most lacking in the law at the time.

The promise of a uniform European regulation with severe sanctions, and the repositioning of the user at the centre of interests – with a text that sought to enshrine respect for the privacy of individuals – was becoming more and more essential, and finally became a reality.

After starting to draft the regulations in 2016, on 23 May, 2018, the EU Regulation (EU) 2016/679, better known as the General Data Protection Regulation (GDPR), came out of the pen, the typewriter, or more likely the word processing software, of the European legislator.

100 pages of obligations to be complied with, 99 articles of law, up to €20 million in fines or up to 4% of annual global turnover: these are the key figures of the GDPR.

During its development, there was a massive campaign against the drafting of the regulation. Key players in the global data market – notably the lobbyists of the tech giants (11,000 organisations, 80,000 people) – materialised in a massive outpouring of money expended to exert pressure and influence on the policy process. Facebook spent $USD11.5 million, Google $6.6 million, Amazon $3.38 million, and Apple $2.14 million. The level of financial advocacy was so overwhelming that the European Union had to initiate a change in the Parliament's rules of procedure to document any likely influence of lobbyists on parliamentarians.

After those millions were spent to counter the development of these regulations, the major platforms had an epiphany: they realised that the forthcoming legislation could actually benefit them and not restrict their progress nor their ability to collect and use people's data.

Indeed, the GDPR could be leveraged to have the opposite effect of what was originally intended.

Today, after just three years in force, complying with the GDPR is so complex and costs so much money and time that it can quickly become a burden for small operators, compared to the tech behemoths for whom additional compliance costs are not a problem.

Companies such as Facebook and Google have a barrage of lawyers and developers to help them comply, and deep pockets to invest in regulatory operations, unlike smaller players who are also required to adhere to the same levels of data privacy.

As a result, a change of strategy took place, and the giants no longer opposed the GDPR, but supported its implementation. In a magnificent turnaround, Facebook founder Mark Zuckerberg famously touted, 'I think everyone in the world deserves good privacy protection'.

Regardless of the political machinations of global tech platforms, the GDPR may have failed in its mission to contain the big tech oligarchs, it has nonetheless created tools that can serve citizens in a digital economy. Indeed, the GDPR has put the individual back at the centre of its concerns and given European citizens and businesses rights over their data and control over their personal information.

This includes the right to data portability, which if used successfully could become a real tool for competition, enabling the sharing and reuse of data.

What exactly are we talking about?

Throughout this report, we will discuss several concepts in legal terms. One of the core challenges in improving data literacy and fostering greater understanding of digital data rights is that descriptions of rights are often smothered in legal jargon and complex legislative phrasing that make it impenetrable for the average digital service user to understand what they are agreeing to, or what avenues of redress are available to them.

Here are some of the key terms and concepts[1] we will be discussing in this report:

GDPR

The General Data Protection Regulation is a European-wide law that defines European citizen rights to data privacy and access. In Europe, under the Treaty on the Functioning of the European Union (TFEU)[2], it is agreed that certain laws made at the European-level are set as requirements that must be implemented by all Member States at the country level. As part of Europe's commitment to being a Digital Single Market, individual Member States must align their country's data privacy legislation with the GDPR. They can develop additional requirements, or define specific implementation conditions, but they cannot remove rights spelt out under the GDPR.

Data subject

A data subject is the person to which the personal data relates. For example, the user of a platform or digital service.

Data controller

The data controller is the entity that collects and manages the personal data. For example, the platform or digital startup that collects user data in order to deliver their services is the data controller.

Data processor

The data processor is the person, agency, organisational team or other body that is responsible for processing the personal data on behalf of the controller. For example, the credit assessment department of a bank may be the processor for personal data on loan applications. Often, for a company, there may be multiple sub-processors involved at different stages of the data processing chain. For example, marketing may track personal data about responses to advertising, account creation department may be the processor of a new bank account client, and then the credit assessment is the processor of the loan application.

1 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-data-portability/
2 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT

Data protection officer

A platform or service (that is, a data controller) may appoint a data protection officer to guide them on their data responsibilities and ensure they maintain compliance with their regulatory duties. Certain types of organisation are required under the GDPR to appoint a data protection officer, depending on their entity type, their size, the type of data being processed, or other factors.

GDPR Data portability

GDPR data portability is the act of transferring one's personal data from one application or platform to another application or platform.

Data Protection Authorities

These are national regulatory bodies that oversee how the GDPR is implemented in that country. When someone makes a complaint about their rights being breached in contravention of the GDPR, they can lodge a complaint to their data protection authority which will then review the case. There is also a European-wide Data Protection Authority that can hear cases that involve European-wide implementations. The European-wide body, the European Data Protection Board (EDPB), also creates guidelines and issues clarifications to help Member State-level data protection authorities to understand and interpret the law, but these are often non-binding and issued as guidance.

Article 20 GDPR Data Portability Right

European citizens (data subjects) have the right to receive their personal data in a structured, commonly used, and machine-readable format.

Article 20: Article 20 of the GDPR enshrines the right to portability. It states that data must be available to be transferred and spells out:

- The manner in which data must be able to be transferred
- To what extent data should be transferred
- The type of data covered by this right, in terms of content, origin, which basis it has to be processed,
- The way in which data is to be processed, and
- What exclusion situations are permissible where this right does not apply, either directly by defining the precise situations, or indirectly by excluding certain data, based on its content or origin.

In other words, Article 20 provides a framework for data portability processes.

Two transfer possibilities are to be made available to citizens, namely:

- Direct transfer from one data controller to another
- Transfer of the data to the data subject (citizen), who may request to receive the data directly.

What are the benefits of GDPR data portability?

GDPR data portability is the act of transferring one's personal data from one application or platform to another application or platform. If applied as intended, this could create a range of benefits for all stakeholders.

For individuals and business owners

Portability allows us to co-create and generate the value we want

Convenience & choice

Data portability means we do not have to re-enter information or copy our data across to new services as we choose to use them.

Service access & quality

Data portability allows us to share our data in ways that allow us to access more personalised services. This increases the quality of the services as they are adapted to our specific needs.

Economic value & savings

Data portability allows us to sell our data or trade our data for additional features or access special discounts.

Participation & engagement

Data portability allows us to spend less time on rebuilding our digital footprint in new services and instead lets us concentrate on participating and being a part of online communities.

For wider society

Portability allows us to coordinate and distribute value generation

Economy

Data portability can strengthen European-based startups and other businesses that make use of shared data, increasing local employment and business tax contributions.

Equity

Data portability creates a more equal playing field where newer businesses can compete against more traditional stakeholders. Data portability lets people from marginalised populations more effectively pool their data for insights and advocacy.

Engagement

Data portability increases our data literacy and participation in digital economies.

Experimentation

Data portability allows individuals to contribute their data to specific causes such as health research, or as 'crowdfunding' inputs to assist new market players develop new products and services.

Extraction-avoidance

Data portability allows individuals to share their data with local businesses rather than the data being extracted into global tech giants who generate their value elsewhere.

For businesses and industry

Portability allows us to collaborate, complement each other, and compete fairly

Innovation

Data portability allows businesses to build new products and services and increase adoption quickly by limiting the onboarding delays of rebuilding a user's digital footprint.

Fair Competition

Data portability allows disrupters and startups with complementary services to enter established markets and compete quickly against traditional incumbents.

Revenue

Data portability enables businesses to reach viability faster, and to gain benefits from the pooled value of shared data.

Portability: a fundamental digital right

The core of the right to data portability has two key assumptions:

- That personal data has a value, either financial or non-financial
- That the ability to move personal data from one platform or service to another creates a value multiplier effect when the data is introduced into the new system.

Let's look first at different models that estimate the value of personal data. Then we will examine the impact that data portability may have in generating new value when that personal data is moved from one system to another.

The value of personal data

The data generated through our interactions and use of digital services is valuable. All of our clicks, comments and research are data that can be considered as work, provided by the user. In general, this data is collected by platforms, products and services, in exchange for some benefit, usually free access to the platform's main features.

The accumulation of this data over time is digital capital, that is, a mass of value stored by the platform.

Put simply, each user provides their digital work (also referred to as 'digital labour') and delegates part of their digital capital to the platforms they use.

Google

In return for free global search capabilities, Google collects all user search behaviours to estimate a price they can charge advertisers. They then offer up their user base (that is, all Google search users) as a potential audience to the advertisers that pay the estimated advertising cost.

Facebook

In return for access to a global tech platform that allows users to publish and share updates, photographs, comments, links, and opinions, Facebook collects data on their users' social connections and preferences, and sells that to advertisers. They also offer those advertisers a platform on which to display targeted ads to audience segments, categorised by the preference and other user characteristics data collected.

The right to data portability between different platforms can therefore be seen as a reappropriation of the data's economic power and the transfer of its digital capital from one platform to another.

How much is personal data really worth?

If you ask how much someone's personal data is worth, the answer can vary widely. In our interviews with users of platform services, respondents estimated the value of their Facebook data anywhere from under one hundred dollars to tens of thousands. There is a real lack of clear information on the subject because valuing data is not a simple equation.

From a purely economic point of view, there are different ways to value data.

Direct valuation: In this model, data is sold directly from one stakeholder to another. For example, a business may buy data from individuals to enrich email sales leads information, or to understand consumer spending behaviour. The buyer would need to buy from a source where a large group of people's data was pooled, or would need to buy individually from a large number of people.

How do we calculate the direct value of personal data?

The direct value of data is decided by the market. For example, companies that collate personal data and sell it as an aggregated, anonymised dataset will set their own price and adjust the pricing to reach their sales goals. Individuals may share their data with platforms and services in return for direct payment, or for other rewards such as gift cards that have a financial value.

Case study: Models of paying for user data (direct valuation)

Paying users for their individual personal data is a fairly new business model, so there are limited examples as yet of individuals being able to monetise their personal data directly. It is more common that companies are able to aggregate and anonymise all user data on their platform and charge others for access to this data.

cake
The profitable banking app.

The Belgium-based fintech Cake sells aggregated user transaction data to companies looking to understand specific target markets. They have a revenue plan with users who consent to share their anonymised bank account transaction data. Each user receives on average 3.11 EUR per month for their data.[3]

YouGov®

This global, digital market research company invites users to answer questionnaires based on their demographic data and they sell this aggregated data to their clients who have requested specific data products or insights. Users receive points towards gift cards, but need to complete multiple surveys on a regular basis to be eligible for 25 EUR or 50 EUR.[4]

Indirect valuation: In this model, data brings a value through its captive use. For example, social media networks like Facebook, and search engine tools like Google, use data from users on their platforms to improve their machine learning algorithms (such as their advertising targeting algorithms), and they then sell access to these algorithms, by selling targeted advertising placements.

How do we calculate the indirect value of personal data?

Two methods of calculation for indirect valuation are possible:

- Valuation based on revenue generated; and
- Valuation based on potential revenue generation over the user lifecycle.

How to calculate the indirect value of user's digital capital using revenue generated

Where a platform or service sells aggregated data from all users (or a subset of users), the individual user's personal data has an indirect value. It is contributing value to the overall dataset that has monetary value to the platform's customer base.

The following process can be used to calculate the indirect value of a user's data based on current revenue generated:

Calculate the annual revenue generated by a platform service through the internal reuse of its data.

This can be challenging as many platforms and services increasingly use multiple, complex digital business models to operate, and reuse of user-collected data may be only one portion of the overall business model value chain. However, Facebook can be used to illustrate how much revenue a user's personal data is worth to a platform. Facebook derives 98% of its revenue from advertising, so it can be considered a near-perfect model for a data monetisation company. Similar calculations can also be done for Amazon, Netflix, Uber, and Airbnb. However, their calculated value is more complex because it is more difficult to link revenues directly to the personal data collected. Data is often used to improve the platform's products and user experience, which in turn increases activities that occur on their platforms and it is those activities that then generate revenue.

Divide total revenue by total number of users.

It may be necessary to uncover average annual users and compare this with average monthly users, or average daily users. It is assumed that a platform's daily users are of higher value than the occasional users, as they are adding more data points to the datasets that can then be sold to the platform's business customers.

Use ranges to show estimates of the value of data.

Estimate value of personal data as a range. For example, estimations may vary based on whether the calculations used annual, monthly or daily active users. Consider if there may be other demographic variations, such as geographic or age range.

3 https://cake.app/
4 https://yougov.com

Case study: how much is a person's Facebook data worth?

In 2020, Facebook generated more than $USD84 billion[5] from 2.7 billion monthly users which includes 1.6 billion daily users. This is an average revenue of $31 per user per year.

If we examine further, we can see that Facebook generates more revenue in some regions than others.

From September 2019 to September 2020, for example, Facebook generated an average of $152 per user in the US and Canada, $58 in Europe, $12 in Asia, and $10 from the rest of the world. That is a factor of 15 between the most and least monetised users on the platform.

If we assume that the value of a Daily Active User (DAU) is higher, and if we value Facebook as the value of its DAU, we can calculate that:

- Revenue per DAU in US/Canada is $256 per year
- Revenue per DAU in Europe is $93 per year
- Revenue per DAU in Asia is $USD19 per year
- Revenue per DAU in the Rest of the World is $16 per year

Worldwide

Q3'19	$7.15	$0.11	$7.26
Q4'19	$8.38	$0.14	$8.52
Q1'20	$6.84	$0.12	$6.95
Q2'20	$6.91	$0.14	$7.05
Q3'20	$7.80	$0.09	$7.89

US & Canada

Q3'19	$33.86	$0.69	$34.55
Q4'19	$40.50	$0.92	$41.41
Q1'20	$33.45	$0.73	$34.18
Q2'20	$35.58	$0.91	$36.49
Q3'20	$39.04	$0.58	$39.63

Europe

Q3'19	$10.51	$0.17	$10.68
Q4'19	$12.99	$0.23	$13.21
Q1'20	$10.43	$0.21	$10.64
Q2'20	$10.81	$0.22	$11.03
Q3'20	$12.28	$0.13	$12.41

Asia-Pacific

Q3'19	$3.22	$0.02	$3.24
Q4'19	$3.55	$0.02	$3.57
Q1'20	$3.04	$0.02	$3.06
Q2'20	$2.96	$0.03	$2.99
Q3'20	$3.64	$0.02	$3.67

Rest of the World

Q3'19	$2.23	$0.01	$2.24
Q4'19	$2.47	$0.01	$2.48
Q1'20	$1.98	$0.01	$1.99
Q2'20	$1.77	$0.02	$1.78
Q3'20	$2.20	$0.02	$2.22

Please see Facebook's most recent quarterly or annual report filed with the SEC for the definition of ARPU.
Revenue by Facebook user geography is geographically apportioned based on our estimation of the geographic location of our users when they perform a revenue-generating activity. This allocation differs from our revenue disaggregated by geography disclosure in our condensed consolidated financial statements where revenue is disaggregated by geography based on the addresses of our costumers.

How to calculate the indirect value of user's digital capital over their lifecycle

An alternative method for calculating indirect value of a user's data is to estimate the contribution to the platform's total valuation, based on the user's lifecycle. In this definition, "lifecycle" refers to the average level of user engagement with the platform, for example, monthly or daily active users. Stock market valuation provides a good approximation of the value of a company's data stock over the lifecycle of its users at today's value, based on future revenues.

The following process can be used to calculate the indirect value of a user's digital capital:

- **Calculate the current valuation of a platform or service**. Valuations for tech companies are often provided by independent analyses or within annual financial reports.
- **Divide total valuation by total number of users.** It may be necessary to compare average monthly users with average daily users.
- **Use ranges to show estimates of the value of data.** Estimate value of personal data as a range. For example, estimations may vary based on whether the calculations used monthly users or daily active users. Consider if there may be other demographic variations, such as geographic or age range.

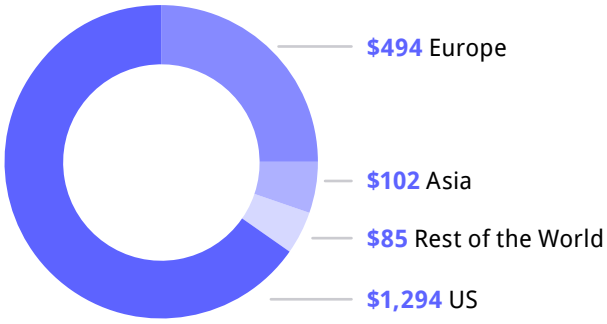
5 Throughout this report, all amounts are calculated in USD for consistency and comparison

Case study: Calculating value of Facebook user's data using the lifecycle method

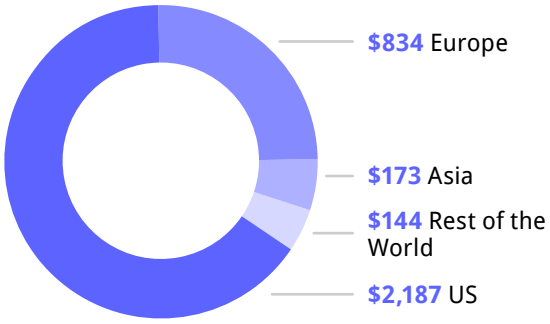
In January 2021, Facebook was valued at $715 billion.

This is an average revenue valuation of $264 per monthly active user or $446 per daily active user.

Average monthly active user lifecycle digital capital value



Average daily active user lifecycle digital capital value



It is important to note that this is the maximum value, estimated by the market, that Facebook can collect from its users. Depending on events and technologies, the market varies the value of this digital capital.

Therefore, we can see that the value of a user's Facebook digital capital over their lifecycle varies from $85 to $2,187 depending on their region and level of activity.

The Facebook user's digital capital, that is, the stock of value that accumulates over time, is kept by Facebook in its data centres. And Facebook collects dividends from this digital capital by improving its algorithms and by collecting more data through having users like, share, comment, fill out forms, watch videos, click, and carry out other digital activities. The user does this work, and contributes a revenue 'dividend' of between $10 and $256 per year depending on their region and level of activity.

Case study: Calculating value of Google user's data using the lifecycle method

This same calculation can be made with other platforms, based on the direct income generated that is linked to user data (such as advertising) or based on the marketable algorithms that the platforms are able to develop with the data collected.

Google, for example, derives 92% of its advertising revenue from $146 billion in advertising revenue. They have 4 billion users, so this works out to about $36 per user per year – around the same as Facebook.

46% of Google's revenue comes from the 246 million users in the US, an average of $487 per year per active user in the US.

Alphabet, Google's parent company, is valued at $1,720 billion. We estimate that 92% of this capitalisation is linked to its advertising revenues – $1,582 billion. 46% of the value of that digital capital is from the US.

For their 246 million US users, the lifecycle digital capital averages at $2,960 per user, or more than twice that of an active monthly US-based Facebook user – and all this is with a free (and certainly very powerful) search engine.

Other models of calculating value of personal data

While the above models discuss ways to calculate the value of personal data in order to determine its value in data portability requests, this is only one lens by which to view the value of data. In practice, data portability demonstrates the underlying value of data that could be exploited in other ways besides monetization.

Data as a social good

Personal data can be used as a contribution to enable greater insights and shared societal benefits. For example, citizens could be willing to share their health data in order to encourage new research into rare diseases, or other health concerns. During the COVID-19 pandemic, there have been isolated examples of people willing to participate in mass experiments. For example, in Barcelona in March 2021, music lovers attended an indoor concert and agreed to share their COVID testing data in the two weeks following the event. In a similar way, if data portability was an available process, people could make similar donations of their data to enable further research as needed. This is similar to the YouGov model described above, but rather than sharing opinions data, citizens could share their platform or service user data. The startups datafunding and Data Fund for example, support users to contribute their data to help digital and data-driven startups design new services and products drawing on anonymised, shared data provided by data donors.

Data as a quality of life enabler

Personal data can be transferred between services and platforms to improve life. If a user wanted to try out a new online service, they could transfer their existing data, preferences, contact details, work history, financial assets, and relationships from one service to another as they wished to avoid retyping or re-entering all of the data needed to make use of the new service or platform.

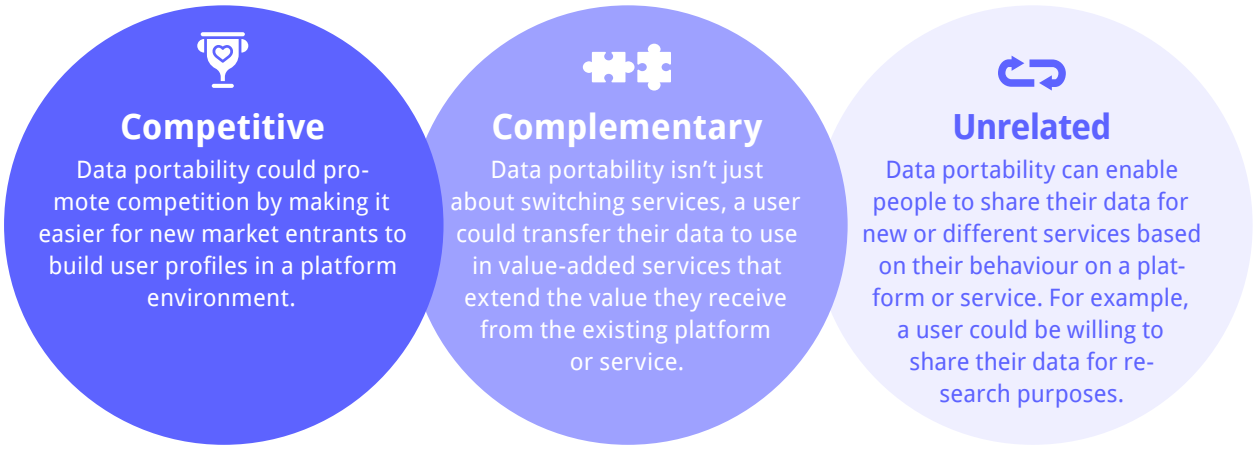
Personal health data could be transferred from one service, platform or device to another service or platform in order to gain new insights on health and wellbeing or to discuss with health professionals.[6]

Personal financial data could be moved from one service to another to allow greater insights into one's financial wellbeing and potential.

In all of these cases, the monetary value of the personal data is not as important as the use value that can be generated by making use of data portability rights.

Case study: The Open Data Institute's analysis of GDPR data portability benefits

When the GDPR was introduced, the UK's Open Data Institute explored potential benefits of exercising one's data portability rights.[7] Here are some of the key opportunities they identified, with examples imagining the impacts of data portability on peer-to-peer accommodation services like Airbnb.



6 https://ftc-workshop-data-to-go.videoshowcase.net/?category=66914
7 https://theodi.org/article/will-gdpr-and-data-portability-support-innovation/

The societal value of enabling data portability[9]

In addition to the value of personal data to the individual, there are also benefits more widely, that models of calculating personal data are unable to reflect. Policy advocates like Ian Brown and Douwe Korff list a range of societal competition benefits that could be realised from data portability:

"Requiring large online platforms (such as Facebook and Google) to enable up-front interoperability [that is, data portability] with other services would give the EU the means to boost competition in digital markets where existing antitrust enforcement has failed to do so.

- "Such enhanced competition would:
- benefit consumers (via increased choice and quality of products and services that better suit their needs);
- stimulate innovation by competitors offering new products and services; and
- bring broader social benefits including:
  - improved social infrastructure (e.g. access for users irrespective of their attractiveness to advertisers, and willingness to sign up to large platforms where increasingly essential communications take place);
  - promoting media pluralism and diversity (e.g. more incentive for news sources to offer quality news rather than seeking to maximise user attention/advertising revenues with disinformation/hate speech);
  - incentives to offer more better privacy (e.g. competing in terms of quality of data privacy/protection safeguards, more data portability);
  - improved moderation of harmful content while protecting freedom of expression (e.g. giving users a choice of moderation regimes);
  - reduced environmental impact of the online economy and "Internet of Things" (e.g. more incentive to offer sustainable products, and to allow users to switch between service providers without buying new hardware);
  - favouring Europe's digital sovereignty (e.g. by allowing new market entrants from Europe to compete successfully)."

The challenge of taking an individual view of personal data

One key concern with a view of data portability being about the data that is "owned by the user" is that often, data is about multiple people, or is valuable only in aggregate terms. For example, DNA data is not just about the single user; it also holds data about family and relatives. Usage data, for example education or energy and utility bills, might not just reflect an individual but an entire household. Some policymakers note that in addition to data portability rights for individuals, there may be a need to create standards and mechanisms that mediate competing rights between people who are represented in the same dataset.[8]

The value of data portability

In the detailed paper, Return on Data[10], business lawyer and researcher Noam Kolt published a very simple and meaningful thesis:

A user agrees to share their data in exchange for a free service. As long as they believe that the value of the service they obtain is of equal to or greater value than the data they allow to be collected about them, they will continue to use the service.

Indeed, it is often when scandals come to light that users decide to switch services, when they reappraise the quantity, use and value of their data that is being collected.

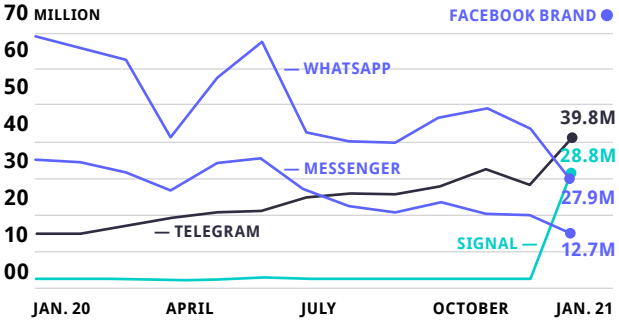
The migration of the messaging apps

In January in 2021, Facebook's messaging apps announced changes to their privacy policies. Following this, cybersecurity expert Zak Dorfman was one of many who shared an analysis[11] of what data is collected by various messaging apps.

As a result, many app users felt that the value of their data being collected from their use of messaging apps was out of balance. Downloads of alternative apps increased, as shown by Fortune and Apptopia:

Monthly messaging app downloads

Number of monthly downloads, globally



8 https://dataportability.projectsbyif.com/summary-and-recommendations/
9 https://www.ianbrown.tech/2020/10/01/interoperability-as-a-tool-for-competition-regulation-2/
10 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3362880
11 https://www.forbes.com/sites/zakdoffman/2021/01/14/3-things-to-know-before-quitting-whatsapp-for-signal-or-telegram-or-apple-imessage-after-backlash/?sh=6e05a7c564f6

According to Noam Kolt's theory, there are two leveraging factors that are a threat to the revenue model of platforms – either the users must be kept in ignorance about the value of their data, or there must be massive and continuous reinvestment in the quality of the service provided.

Current platforms and digital services may add occasional new features, but evolutionary leaps in the quality of services being provided are rare. Therefore, the current revenue model of platforms is based on restricting or downplaying the user's awareness of the value of their data.

Data portability is dependent on a user's knowledge of the value of their data. This value can only be realised if there is a means by which users can exercise their right to portability in order to derive a direct benefit from it.

How much is data portability worth?

Theoretically, data portability could bring enormous value to the market. Using the Facebook example above, we calculated the value of a European user's Facebook digital capital at $494 over the user's lifecycle. Under the GDPR data portability right, this would be the value a user's data could contribute to a new platform or service if it was migrated across to the new platform/service.

Each time a European Facebook user exercises their right to data portability, it yields a maximum value of $494 that is copied and added to another platform or service, without really destroying any value at Facebook: Facebook only loses its monopoly over that data. So, if a user exercises their right to portability of their Facebook data 20 times to 20 different platforms, nearly $10,000 of value is 'created' without Facebook losing their value.

Therefore, in theory, portability is a positive-sum game where the sharing of data does not entail anybody losing value.

With limited evidence of data business models at present, this is still fairly theoretical. In reality, the data capital would not reach its full value as the data necessarily loses value when it is transferred from one platform to another. Why is this?

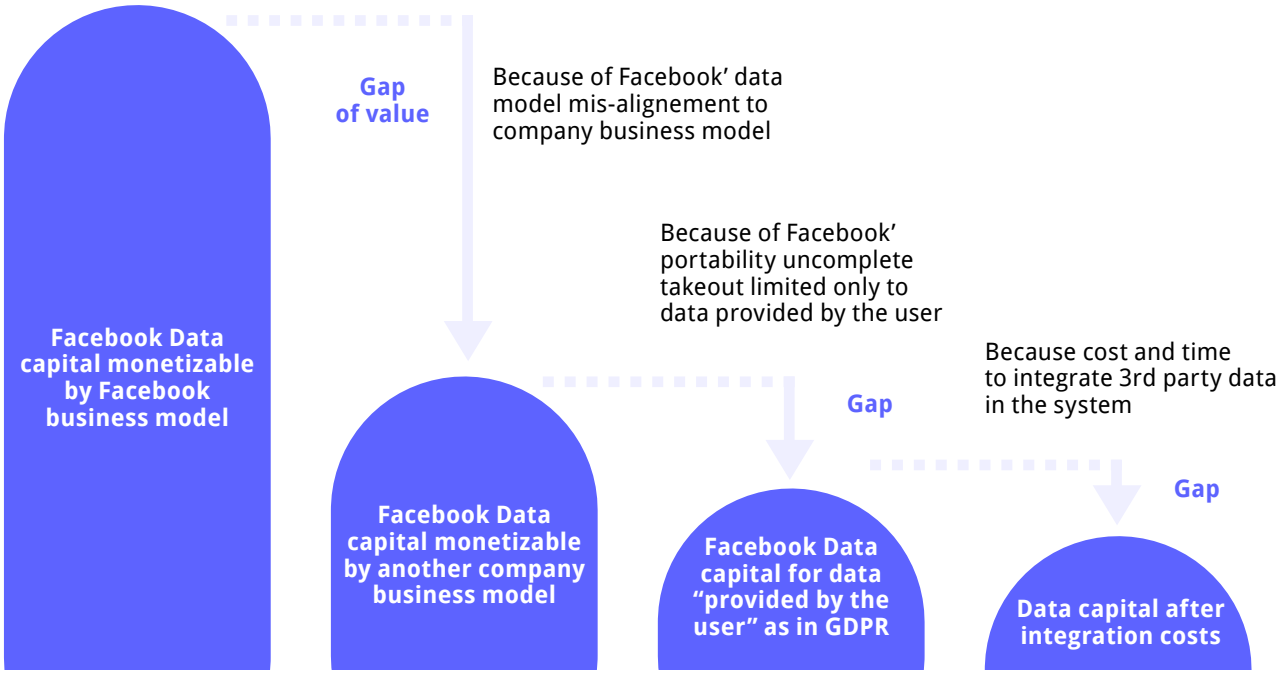
Facebook data is only able to be monetised to the maximum extent possible by Facebook, because it was made for Facebook's revenue model. Data has a model, and the way in which it is thought out and organised at Facebook will not be the same elsewhere. It has a format, and it has contexts of use that are not transferred when a portability request is made, and this makes it lose value. No other platform, even a direct competitor, can exploit Facebook data as efficiently as Facebook does.

Indeed, data cannot be exchanged and used without friction between different economic actors.

Case study: Loss of value during the data portability process

The following example illustrates a study conducted among developers on their perception of the value of portability data from Facebook. On the Facebook Developer platform, millions of developers see value in the data delivered by the Facebook API. They see virtually no value in the data shared via the Facebook portability tool.

Value of portability for Facebook data as a European user in 2020



However, in some situations, companies will have other revenue models and may benefit from Facebook data, uncoupled from the intrinsic value to Facebook of their own data.

A bank, for example, may analyse your social network and verify that you are a trustworthy person, with strong personal connections that can act as unofficial guarantors, and may use this data to better direct credit offers to you that you accept. In this way, it may monetise more than $494 worth of value over the lifecycle.

The accumulation of data by large platforms creates asymmetrical business models

Data captured by large platforms is not easily accessible to newer, innovative platforms and services.

The ability of the tech giants to attract large numbers of users, through free and well-designed services, means that those platforms are able to accumulate a lot of data to support decision and recommendation tools. Eric Ries's philosophy in The Lean Startup[12], begins by stating that the knowledge of the user is more important than monetisation. The platforms use this knowledge to evolve faster to meet user demands. Through this continuous improvement, there is a greater retention of users on the platform and this snowball effect maintains a model where 'the winner takes all'.

The digital capital of personal data is concentrated in a few players, sometimes known as GAMFA (Google, Amazon, Microsoft, Facebook, and Apple). With this concentration of data and power, it becomes more and more difficult for an alternative application to get started and convince users to switch to their service.

The aim of data portability is to enable users to become an economic player in their own digital consumption, and to make their accumulated data stored on existing platforms available to benefit the new platforms they may wish to use.

For businesses, portability allows new market entrants to provide a competing or new service quickly by reducing the friction that occurs when new users have to establish themselves on a new platform. For example, users would need to data enter all of their preferences, their social connections, their work history, and their assets to generate new value from their personal data on any new platform that they wanted to join.

In a world where personal data portability is a reality, anyone should be able to capitalise on the use and value of their data.

Case study: How data portability could work for a job seeker

A freelancer registered on LinkedIn, for example, could transfer all their data (experiences, recommendations, skills, professional and academic background etc.) to job matching services. This could be an online platform (such as Malt) or a more traditional recruitment agency. The agency would receive this data, with their consent, and recommend their profile more easily within their network of clients. The advantage here is the absence of data entry – the freelancer does not need to refill the same information for each agency. The agency also saves time by not having to verify the veracity of what is put forward, because the source of the data allows a reasonable presumption of its reliability. Therefore, the freelancer as well as the agency can multiply the number of opportunities for connections and employment. The freelancer can, in a few clicks, share data with as many agencies as they wish, and the recipient agencies have rapid access to useful information to be able to secure contracts. Everybody wins.

To draw a parallel with economic theory relating to stock markets, portability provides a form of information efficiency. Operators acting in this market can make quicker and better informed decisions drawing on the data ported into their platform or service.

Portability should also increase the capacity for innovation, shifting part of it from the large, established players in a market to the smaller ones, and vice versa.

Under data portability regulation, data, and its value, is therefore destined to continue to grow and could be increasingly pooled. And if this value is something that can be easily captured and accessed, the effective competitive differentiation will no longer be based on the data itself but on how to take advantage of it. In this model, the person and their data are no longer the product, instead the product is the service that should become more and more adapted to the person. We can see a future in which a person, with their capacity for innovation and their creativity in the design of products and services, will be the only real source of value.

12 http://theleanstartup.com/

Data Portability in Theory

The State of GDPR Data Portability: Principles

Article 20 covers the 'Right to portability', and is in Section 3 of Chapter 3 the 'Rights of the data subject', of the GDPR . It defines:

- That the data subject has a right to receive the data;
- The manner in which it must be transferred;
- To what extent it should be transferred;
- The type of data concerned by this right including content and origin, according to which basis it has been processed and the way in which the data is processed.
- It also excludes situations where this right does not apply, either directly by defining the precise situations, or indirectly by excluding certain data, by content or origin.

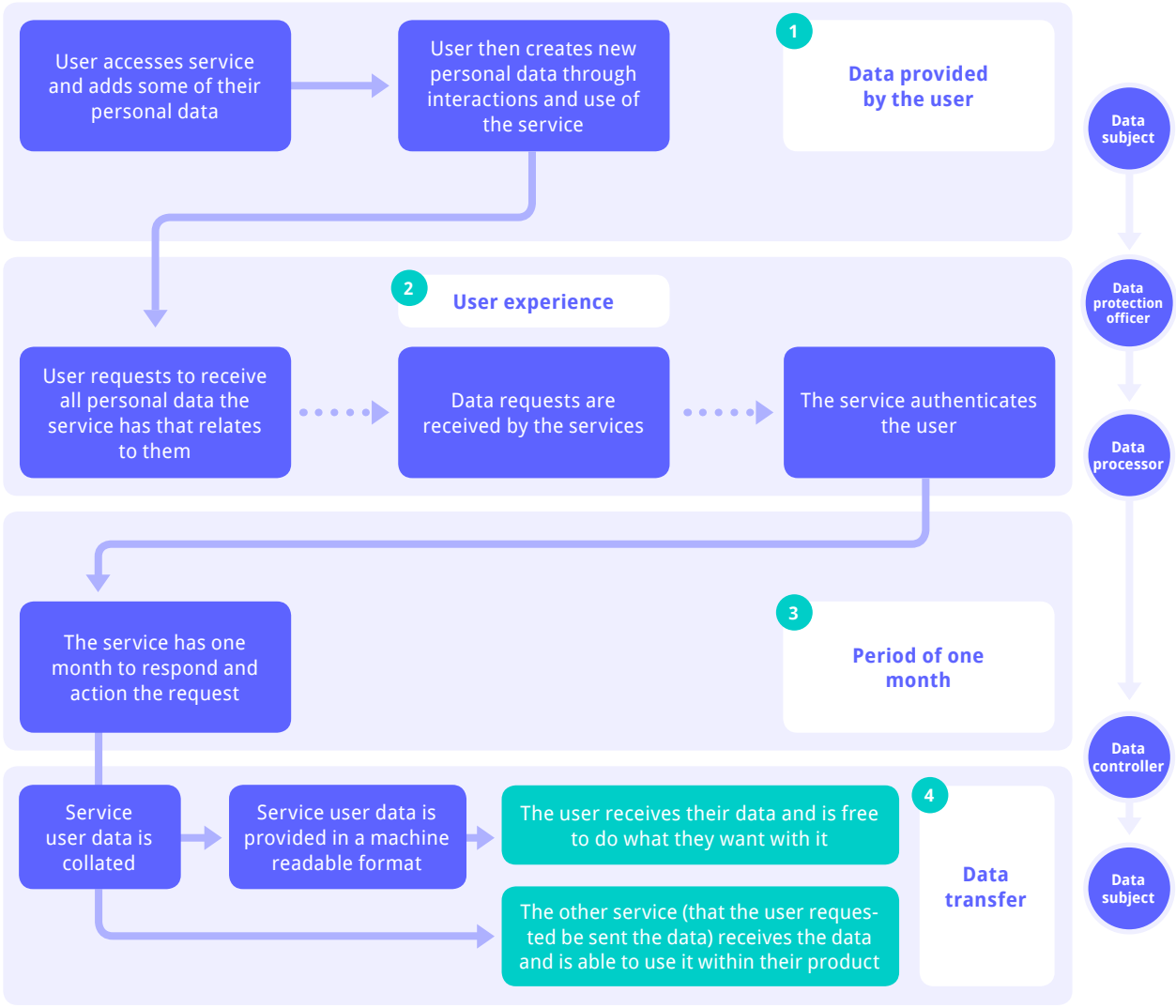
Article 20 defines the right to data portability and it specifically provides that this right does not apply only to personal data provided by the user; nor to data concerning the user; nor to data processed on the legal basis of consent or contract, or where such data is processed by automatic means. That is, the user's data is considered to be both the data the user contributes, plus all of their interactions in the platform or service.

Article 20 also provides a framework for the practice of the right to data portability, that data subjects have the right to receive data in a structured, commonly used and machine-readable format.

Two transfer options are offered:

- Transfer of the data directly to the data subject; and
- Transfer of the data from one controller to another.

The figure below describes the data portability process.



1. Data provided by the user

- A user accesses a service or platform and enters some of their personal data. For example, on LinkedIn, users add their education and employment history. On Airbnb, property owners and experience creators enter details of their offerings. On Facebook, users add their family and friend connections and might indicate some of their personal preferences (hobbies, interests, political affiliations, etc).
- Over time, users also interact within the service and platform which generates more data. On Spotify, users may create playlists of their favourite songs, and even if not, Spotify is collecting data on the number of times they listened to each song, for example. On a rideshare service, data on frequency, length, usual pickup and destination data is all collected. On Fitbit, the number of steps walked, or exercise durations are collected on a daily basis: this is collected automatically by the platform from the user's digital interactions.

2. User experience

- At some point, the user may request access to this data that is collected about them.

Case study: How companies enable users to request data

airbnb

Airbnb makes the data protection officer's email available.

Google

Google has a system in place for users to download a data archive.

Spotify

Spotify provides an online form.

What is personal data?

The GDPR is challenging to implement because it is actually made up of several components, both of which give different levels of clarification around what constitutes personal data:

- **The actual legislation.** This is often vague in its terms. For example, under Article 20 of the GDPR legislation, mentions "personal data [concerning the data subject] and which they have provided". But what is data provided by the user? How do we define what is user-provided and what is not? What are the criteria?
- **EDPB Guidelines.** These are non-binding implementation guidelines provided by the European Data Protection Board and adopted several months after the regulation itself. According to these guidelines, the "data provided by the user" includes data actively provided by the user (for example, data entered by the user when setting up their user account and service profile). The guidelines also state that personal data includes the user's activity such as history, searches, and logs, and it also includes statistical data created by the service or platform from the user's activity. This 'derived' or 'inferred' data, is the data resulting from the analysis of the user's behaviour. The EDPB's guidelines and recommendations often lack legitimacy with EU Member States and also with major digital actors who do not take them into account when responding to data portability requests.

- These requests are then received by the data protection officer or department within a platform that is responsible for data processing. The service may check to verify the identity of the user to confirm that they made the request.

Before the data controller grants the user's request, they have a duty to authenticate the requester, so that no personal data is inadvertently or carelessly transmitted to a malicious third party that could exploit the user's data without consent. Only the user concerned can therefore proceed with a request for data portability.

In order to authenticate themselves, the user must provide the means to certify their identity, such as a customer number. If reasonable doubt remains, the data controller may ask the user to send a copy of an identity document to prove that they are the requester.

3. Period of one month

- Under the GDPR, the service then has one calendar month to collect the data and share it back with the user.

In principle, the user would only have to mention in an email that they want to exercise their right to portability.

A period of one month begins from the date the request is made, and the data controller must comply with the request within this period. If a complex portability request is made, this period may be extended by a maximum of two months and will have to be justified.

4. Data transfer

- The service collects all of the data about the user.
- This data is then put in a machine readable format.

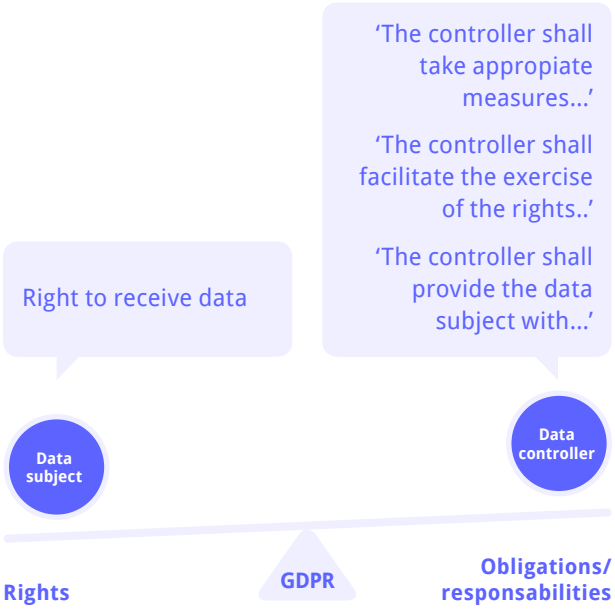
As soon as possible after authentication, the data controller must make available to the user all the data that the user has provided either with their consent or by means of a contract, and must also include all the data resulting from their activity, in a machine-readable format, namely a Json, XML or CSV file.

It is not intended that the user should just receive an overview or summary of their data. The goal of portability is to be able to reuse personal data immediately (or later) by integrating it onto another platform or service. Therefore, the format in which the user receives their data is paramount. This should be one of the most important conditions respected during the data portability process.

- The service then provides the data in one of two ways, as requested by the service user. Either the user receives the data directly, or if they have asked for the data to be given to another service or platform, this data is then forwarded to this other party.

In the context of the data being transmitted to the requesting user, Article 20 allows the data controller to choose how the user's data is provided. The user is obliged to agree to the data being transmitted by the means chosen by the data controller. So, where a user might prefer to have the data transmitted directly to their phone or to their digital storage account on a service such as Dropbox, they will often have to be satisfied with downloading the data archive from a platform determined by the data controller.

At first glance, then, Article 20 seems clear, concise, precise and leaves little room for interpretation. Article 20 in the GDPR states that individuals have the right to receive their data. This right must be an integral part of the information system and processes of the data controller. As Figure X shows, rights and obligations between data subject (the service user) and data controller (the service provider or platform) are in balance.



Data subject rights constitute an obligation for the data controller, and are linked to Article 12 of the GDPR which includes clauses with the wording like:

- 'The controller shall take appropriate measures...'
- 'The controller shall facilitate the exercise of the rights...' and
- 'The controller shall provide the data subject with...'.

Research into the Current State of Data Portability

Why conduct a study on data portability?

Despite the potential for value generation through data portability, there were several indications that led the authors of this report to doubt the effectiveness of this digital right.

Doubt began to germinate when one of us wanted to use this right with a famous online platform for short term real estate rentals. It was very difficult to get the data in a usable format. In order to consolidate the intuition that this experience would be common, the author then sought to exercise this right with a large professional social network. Yet again, the author was confronted with obstacles and arguments for not granting the data portability request. The barriers described by the platforms did not seem to be in accordance with what can be expected from a reading of Article 20 of the GDPR, which enshrines the right to data portability.

These refusals to make data available for portability inspired the establishment of a team of lawyers and data protection experts. This team, whose main actors are the authors of this report, set itself the task of carrying out a study in order to have a more general idea of the current state of the right to data portability.

What are the study aims?

The aim of the study is:

To measure the maturity of individuals, companies and the regulatory environment with regard to the portability of personal data, as enabled under Article 20 of the European General Data Protection Regulation.

This report is therefore a summary of our observations and findings made during the course of this study.

How was this study conducted?

Our study on the right to portability was carried out in three stages:

- We conducted interviews with participants.
- We assisted those who wished to exercise their data portability rights, and closely monitored the process and user experience.
- We collected and categorised information on response times, the type of responses, and the type of data sent by data controllers.

Contact with participants was made via a post on a social network, in which anyone who wished to take part in a study on the right to portability was invited to participate. Many people responded to this call and contacted us. An interview via videoconference was organised with each person who responded positively.

During these interviews, we included questions about knowledge of the GDPR, portability, whether interviewees had ever exercised their rights and, if so, what their experience had been. These interviews were semi-directive, in order to give participants room to express themselves, and to collect as much feedback as possible about their experiences and aspirations regarding their appetite and interest in exercising their right to portability.

In these interviews we offered to help them exercise their rights to data access and portability as part of the study. Although this study focuses exclusively on the right to portability, the majority of participants were particularly interested in knowing what personal data the companies 'had about them'. As such, we invited those who were enthusiastic about participating to provide us with a list of companies with whom they wished to exercise their rights.

We provided participants with contact information for all the departments and data controllers responsible for handling data-related requests from the companies they had selected, as well as a prepared template for the request to exercise their rights (see Appendix X for copies of these templates).

The participants then sent the email and message templates to the data controllers.

As part of the study process, it was agreed that participants would share all responses sent by the data controllers with us so that we could analyse them and help the participants by providing email templates for their responses. In addition, the participants detailed the type of data they had received, as well as the format of this data so that we could analyse and reference it. We did not get to know the content of this data. We did not seek the information itself but rather the context in which it was provided.

All support was provided free of charge and in accordance with good practice with respect to the protection of personal data and privacy. All the follow-up data on the progress of the requests were anonymised with pseudonyms when necessary.

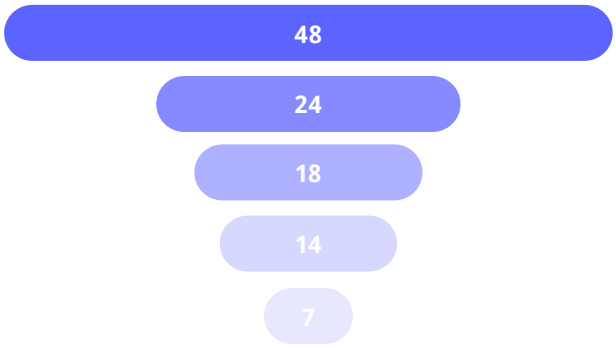
It is important to note that the persons contacted freely gave their contact information because they were interested in the right to portability. Moreover, most of the people contacted to be part of the study already had a strong interest in the digital economy.

We provided email templates and legal support for the people who volunteered to request these two different data rights. In this way we could examine whether the data controllers themselves made a difference or introduced confusion into the processes.

Number of companies to which the participants issued the initial intention to issue a request.

Number of requests actually sent for the study.

Number of responses received to the various requests.



Number of people who pro-actively volunteered.

Number of interviewees actively engaged.

Number of people who signed the engagement contract.

Number of people who actually used their right to portability.

Number of people who actively participated by significant returning data and emails from Data Protection Officers.

Summary: Synthesis and results

Participating actively in the portability requests by accompanying the study participants allowed us to perceive events and redundant comments, and also to identify the global behaviours of the different platforms.

The different types of remarks can be divided into two categories:

- The user experience of portability and
- The data controller's practices.

To start with, people have little interest in portability law. They are not aware of this right. They do not know how to exercise it. They are unaware of the opportunities data portability offers them.

But when individuals are aware of the existence of their data portability rights and are inspired by the potential use they could make of it, the portability process itself ends up discouraging most of them from pursuing access to their personal data.

For example, users do not have any template for sending email requests. There is little outside help available to properly exercise the right to data portability. There is a lack of automated procedures to facilitate data portability.

Furthermore, companies do not cooperate. There is little standardisation in portability processes and the user experience is deliberately obfuscated. The one month time limit is rarely respected, and the manner through which a user can exercise this right is often concealed.

There also seems to be a real confusion for companies between the right of access and the right of portability. These two rights are similar and yet are implemented in completely different ways by data controllers, that is the services and platforms that hold the user's data.

During our study, when requesting data in partnership with data subjects, many mistakes were made by the services and platforms. This was especially the case with the data formats and their structure, as some data controllers sometimes failed to respond correctly in accordance with Article 20 requirements.

The study also demonstrated the excessive complexity of processes related to data portability. The existence of the right to data portability was sometimes only mentioned in privacy policies, and was sometimes hidden within tabs, along with the other rights defined under Chapter 3 ("Rights of the data subject") of the GDPR.

As a result, the user experiences total confusion. Users find themselves alone and somewhat helpless in the face of companies that exceed the legal one month time limit for returning data. They deal with some companies that do not respond, and others that drown the user in useless information. The experience does not make the right to data portability attractive to the user at all.

With regard to the data provided to the user, there is also a major disconnect between what is expected when reading the law, and what is actually provided. There appears to be an attempt by some platforms and services to implement a view of portability that results in locking out any exportable value of the user's data. In some cases, platforms and services only provided data with little added value, or only provided the data that the user had provided throughout their journey on the platform in question.

We were confronted with many examples where data controllers refused to transfer data to another controller on the grounds of technological unfeasibility. This

was often the stated reason (from digital companies who are global leaders in managing user data digitally) without any explanation of why the technical capacity of providing user data was not available.

Despite regular correspondence from the study authors, including a personalised follow-up and the lifting of the legal friction by providing access to standard templates, many of the study participants still lost interest in exercising their right to data portability because the responses from the platforms and services were not satisfactory, and the gaps between responses were too long.

Key finding 1

Exercising the right to data portability is a time-consuming and energy-consuming process that has the effect of discouraging the user from requesting and making use of their personal data.

Key finding 2

Services and platforms (data controllers) often use loopholes in the GDPR law to evade their responsibilities.

Key finding 3

The term data 'provided by the user' is often too vague and does not resonate sufficiently with the practice of data controllers to respond to data portability requests in a way that has value for the user.

Data Portability in Practice: Detailed Findings

While Figure X above outlines the data portability process in theory, our study findings show that in practice, there are multiple blocks and challenges for data users seeking to make use of their data portability rights at each step of the process.

1. Data provided by the user

'Data provided by the user' is not sufficiently defined in the GDPR legislation, and definitional clarity in guidelines from the European Data Protection Board are not always respected at the EU Member State level, or by data controllers.

For a service or platform, the most valuable user data available is the detailed statistical and analytical data of a user's behaviour that is created by the data controller. The more precise this behavioural data is, the more valuable it is to the data controller because it allows the implementation of micro-targeted treatments, such as automated and personalised commercial prospecting, which the data controller can sell to advertisers. Data controllers try not to provide this type of data to the user through the data portability process.

facebook

The "social graph" constitutes a person's knowledge network, or a representation of it, on a social network platform.

On Facebook, a user connects to friends and colleagues and builds their social graph. So this 'friends and family' social graph network is created by the data subject, using Facebook's functionalities to establish connections with other users. We can consider all these connections as data generated by the user.

Therefore, this social graph should be provided by Facebook, and other similar social networks, when a data portability request is made.

In fact, this social graph was available to developers using the Facebook API, before the Cambridge Analytica scandal.

During our study, portability requests made by study participants did not result in the provision of a social graph.

A brief look at various privacy policies shows that the data that the data controllers consider as 'user supplied data' is far less than what the GDPR and the EDPB Guidelines specify. This problem regarding the data provided in requests is not only limited to the definition of data understood as coming from the user, as set out in the privacy policies. The problem also lies between what is written in these privacy policies and the reality of the data that is provided to the user when they exercise their rights.

Certain categories of data will be referred to as 'user-supplied data' in the privacy policy and will not be included in the data provided to the user. This may amount to withholding of data, despite data access being a user right.

We realise that the scope of the data concerned by the right to portability is not clearly defined, at the very least it is subject to interpretation. Even if it is specified by the data controller themselves, in some cases it is not respected when data and portability requests are made. Therefore, on the same type of platform, which processes the same categories of data and makes the same use of them, the user will not receive the same categories of data or the same amount of data.

One solution is to clarify and establish a non-exhaustive list of data categories that is defined as provided by the user. This would avoid this current disparity.

The lack of respect for the right to portability by major tech platforms and services is also due to the imprecise text concerning the data to be transferred. The scope of the data *seems* precise as it is written in Article 20 of the GDPR where it names the data concerned as 'personal data concerning them which they have supplied to a controller'.

In practice, however, the expression 'data supplied by the user' is far too broad and imprecise and still leaves a great deal of freedom to the data controllers, who themselves define what they consider 'data supplied by the user' to transfer to the user.

During our exchanges with data protection officers at platforms and services when supporting the study participants, we noted that they often considered that certain data belonged to them, including data constructed from the activity of the user concerned. In this way, the data controllers disregarded the EDPB guidelines, which specify that data resulting from the user's activity, derived or inferred data, fall within the scope of the right to portability.

If we consider that the data is an extension of the person, how can we accept that companies do not provide it when the user requests? Withholding information could be considered as a violation of certain fundamental rights and freedoms such as the right to bodily integrity or respect for home and communications of Article 7 of the Charter of Fundamental Rights. The retention of certain data could even be considered an infringement of bodily integrity and privacy.

2. User experience

The user experience (UX) is not addressed at all during the entire process of the data portability right and does not easily facilitate the user's access to portability of their data.

There are no official request models known to be available, which leaves the user completely powerless, as they are not aware of the technical and legal possibilities offered by the data portability right.

The request process often changes for each data controller. The user may have to:

- Send an email
- Fill out an online form or
- Use a data retrieval service, where they must choose for themselves which data they want to retrieve.

Each method has its advantages and disadvantages, but the user first has to find out where and how to access the services associated with the data right requests. These are not always easy to find, as they are often only specified in privacy policies.

While Article 20 provides some details of the data concerned and the format of the data, it does not provide any details on how to make it accessible and visible to the user, and unfortunately this is not limited to the right to portability but to all the rights in Chapter 3 of the GDPR.

Current implementations of data portability requests do not reflect the objective of the right to portability, which seeks to enable the free and easy exchange, remuneration and transfer of data.

3. Period of 1 month

A very common experience for all study participants seeking access to their data was the incredible length of time that passed between the date they sent their request and the date they received a response.

Some participants even told us that they had given up on the idea of continuing the application process because the wait was so long.

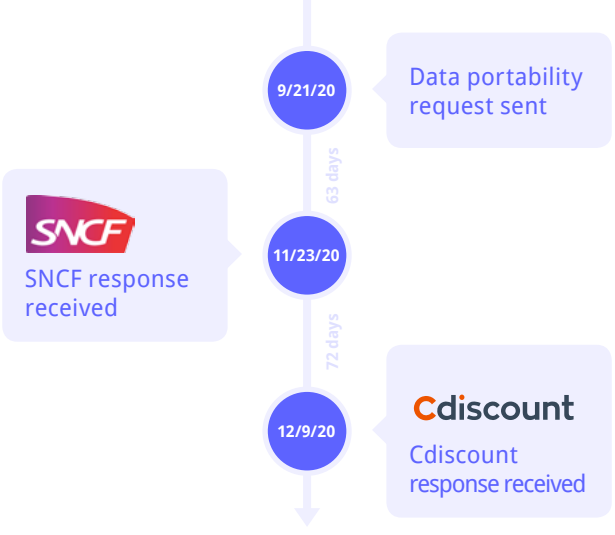
One month may seem short for the data controller. The service or platform has to consider the request, provide their response in the right data format, and collate and send the right amount of data.

But a month is a very long time for a user who would like to recover their data as soon as possible.

"Within one calendar month" is the time requirement stipulated for data controllers under Article 12 of the GDPR.

Normally, any time limit that is exceeded in bad faith or without any justification should be sanctioned. In practice, failure to comply with this deadline is almost standard operating procedure.

Case study: Requests for data from SNCF and Cdiscount



The poor use of this right by users also has repercussions on sanctions. Few requests are submitted to the national data protection authorities, therefore, they are not alerted to the numerous, frequent legal violations that occur during the various portability requests.

This weak control and oversight by the national data protection authorities does not encourage data controllers to comply with the portability right, including implementing the technical means to comply with it, or to respect Article 12 of the GDPR or the EDPB Guidelines, which provide details on the data to be transferred and the fair way to do so.

Article 12 of the GDPR states that the one month period to respond to requests 'may be extended by two months, where necessary, taking into account the complexity and number of requests'. This sentence was treated far too lightly by the legislator who did not suspect that data controllers would use it to their advantage by utilising this extension far too easily.

We found that data controllers, when dealing with people who are aware of their rights and aware of the data they can retrieve, were still utilising the extension of the response time as if a more complex request for portability even when dealing with a data request for a normal user who is not educated in personal data law.

4. Data transfer

4.1 Technical feasibility of collecting service user data

Paragraph 2 of Article 20 of the GDPR provides a wide margin of discretion to data controllers, and the EDPB Guidelines do not clarify obligations.

This paragraph specifies that if the data subject requests the direct transfer of their data to another controller, the controller to whom the request is made must carry out the request 'where technically feasible'.

WhatsApp

The data controller noted they could not forward the individual messages of the user under the pretext that the data controller did not have access to them.

However, users can link their accounts to Google Drive and store their messages that way when changing device, so it is proven possible that the data controller could create a mechanism that allowed users to collect and move their message data through a portability request.

Data controllers make use of the argument of 'technical feasibility' to deny user data portability rights. In theory, when the user makes the request, the data controller must carry out the transfer using an Application Program Interface (API). However, some APIs do not allow data to be received, even in a structured, machine-readable format, and this leads to a lack of interoperability. The EDPB Guidelines unfortunately state that this should not create an obligation for data controllers to adopt or maintain technically compatible processing systems to send or receive machine-readable data. This is then used as an opportunity by data controllers to prevent, or at least slow down, the transfer of data and to retain the monopoly of the value held by them on the user's data.

The regulation itself creates barriers between the user's request and respecting their rights to their data. The data controllers, who must respect the principles related to data transfers, are concerned about the legal and organisational methods of carrying out this processing and are not always familiar with the platform to which the data is being sent, nor with the country to which the data could be being sent. As a result, some data controllers guard against incidents that may occur during the transfer so that they are not held responsible in the event of a data breach that occurs at the time of the data transfer.

airbnb LinkedIn

During this study, data protection officers at LinkedIn and Airbnb were reluctant to transfer data under the user's portability request, as they said they were concerned that the user's data files could not be received and properly implemented by the receiving party of the data (for example, to another platform or service the user designates).

This raises the question around the responsibility between the data recipient and the data controller that currently holds the data.

Is it the fault of the recipient if they did not set up a tool to receive the data? Or would the responsibility lie with the data controller who acted in bad faith and who does not wish to transfer the data to a platform other than their own and who claims a lack of technical feasibility and interoperability with the recipient?

This whole issue is one of "interoperability": the ability of a product or system, whose interfaces are fully known, to work with other existing or future products or systems, without restriction of access or implementation.

Data controllers only rarely transfer data directly to another platform. If requested by the user, they use the pretext that the information system of the other data controller is not interoperable with their own and that it will therefore be impossible for them to transfer the data and have it be implemented on the platform receiving the data.

Apple

Apple, when addressing the data portability request of one of our study participants, replied that as there is no developed means for data controllers to directly exchange data with each other, they preferred to send the data files directly to the user to 'empower' and allow them to provide the data to a recipient 'in a format that fully meets the requirements of data portability'.

This then obliges the user to receive the data they have requested to be transferred, and to upload the data themselves on the secondary platform or service. This may discourage the user, given that the vast majority of users are non-technical and unfamiliar with file formats and the process of implementing data, which may also need to be adapted, depending on the receiving service's technological choices.

We believe, given the experiences we saw throughout the study, that there is clearly effort at play by data controllers to do everything possible to prevent users, and their data, from moving away from the platform they are on. This partly explains the digital landscape we have today, with tech giants offering suitable services that are dependent on our data, with no real competitors able to offer alternative quality services.

4.2 Machine-readable Format

There is also a clear lack of clarity regarding the format in which the data is sent.

In principle, when a portability right request is made, the data should be sent in a commonly used, structured and machine-readable format.

Article 20 of the GDPR, however, does not specify which data formats are applicable under this definition.

As a consequence, due to lack of knowledge, bad faith, the culture around portability, or a combination of these factors, we observed that data is sometimes sent in the wrong format. Formats used included PDF documents, notes, spreadsheets, and Google documents. None of these are machine-readable data files and they do not allow the transfer of the data to another platform or service.

When the data format is not applied as noted in Article 20, the right to portability itself becomes meaningless. The major objective of the right to data portability is the redistribution of data and the decentralisation of holding data, that is, the capability to allow the user to move their data from one platform to another with ease and with the potential to generate new value from doing so.

In our study, users often received unusable datasets, simply due to non-conformity of the format. This data no longer has any value to the user, and results in simply adding a certain amount of worthless digital storage space for them. In fact, if the data was sent to the user in the right format and the user did keep it in their digital storage space, then they could transfer the data to another service as they needed, creating a data portability from their own data storage.

Spotify

Spotify sent very little data that met portability requirements. Few data formats were interoperable. Their data included file names that were often incomprehensible to the user, such as '1P_Custom_cultural_affinity_Pride'. Compounding this confusion, was the fact that the file itself contained no data!

amazon

Amazon partially respected the data formats by sending .CSV files, a format that is machine-readable and is commonly used. However, like Spotify, many of these .CSV files did not contain any data, making them unusable for data portability.

We also observed situations in which the user only received files in a format that met the right to access data and not the data portability right. The file formats corresponding to the right to data portability are totally different from those sent in response to the right to access data. The data portability files are supposed to be interoperable and machine-readable, because they are intended to enable reuse of the data by the data subject. The data sent in response to the right to access is intended as a consultative purpose so data subject's can see what data is held about them. Sending a .PDF or .XLSX file does not allow the user to reuse this data, and is therefore in violation of their data portability rights.

Cdiscount Spotify

From a number of companies, including Cdiscount and Spotify, most of the data received by the study participants was in file formats that were not machine-readable and were more in line with the right of access than the right of portability.

Synthesis: Is portability being sabotaged by data controllers?

There are several indications that users are being deliberately worn down by data controllers through the data portability process in order to reduce the use of this data right. Evidence to suggest there is some bad faith occurring includes:

- The deadline for responding to data portability requests (one month) is frequently disrespected. Study participants noted this discouraged them from continuing.
- Many of our participants told us that their personal experience with data portability had left them with a sense of incompleteness and weariness from the process overall and from the timelines in particular.
- We found that most data controllers do not bother to respond as soon as they can, and the user often receives a response to their request, only on the day (or day before) the deadline is reached.
- The exchange of various emails initiated through the request process, for example, to confirm authentication by sending IDs (which is part of the portability process), were also convoluted and not conducive to a simple and quick response.
- Users who practiced their right to data portability often found themselves with several mixed files, in differing formats, some of which respond to the right of access (that is, were useful for information only) and others which did enable portability (that is, were machine-readable and interoperable).
- Several data controllers, when responding to the data subject's requests for both access and portability would only mention one of these two request drivers (automatic responses were frequently the first response sent). This often left user's perplexed as to the data controller's understanding of their request, and also raised doubts by our study participants who wondered if they had formulated their request correctly.
- When sent data files under data portability requests that were not in a reusable format (as frequently happened), an average user, with no particular knowledge of data formats, might subsequently want to implement these unusable files on another platform and would have the unpleasant surprise of having the data inadmissible by the receiving platform.

One of the challenges in addressing these potentially bad faith behaviours is that the regulation does not define any obligations regarding the technical implementation of the data portability right.

The history of the formation of the GDPR would suggest that data controllers are hesitant to enable data sharing. The reliance by platforms on business models that make use of user data to sell advertising also suggests a motivation to safeguard these assets and keep the data for themselves. This leads to a reluctance in supporting the establishment of an effective right to portability that would enable the user to emancipate themselves from one service or platform and move to another.

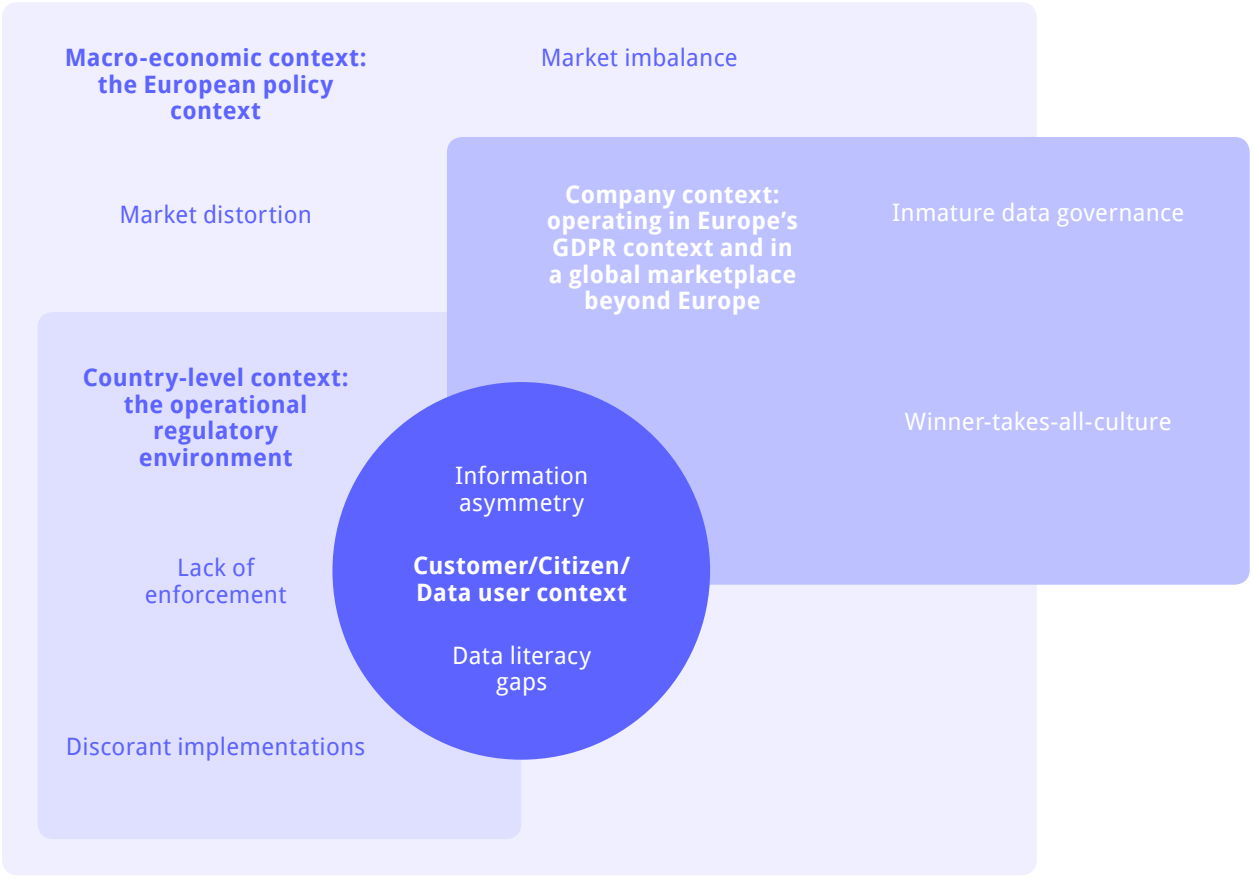
An inherent flaw in the GDPR is the failure to clearly mention that only a copy of the data and not the data itself is transferred to the user. Clarifying this might serve to allay the fears of tech giants and other platforms that are concerned that the right to portability is a threat to their monopoly. Understanding that data portability preserves the data on the existing platform might serve to reduce fears that a portability request is the first step in a user detaching themselves completely from one platform to use another service.

Key obstacles in enabling the right to data portability

The right to portability is still under development

From our vantage point of having studied the GDPR, data portability regulations and guidelines, and having conducted a practical study to test the GDPR data portability processes, we have clearly seen a range of obstacles that prevent the real value of data portability from being realised in the European context.

The current GDPR data portability environment is an obstacle course of blockages to precariously navigate around and tangling nets of confusion and obfuscation to avoid getting caught up in. These traps and challenges have been placed throughout the GDPR context: from the macroeconomic context down to the consumer perspective and every step in between, as shown in Figure X. While many of these challenges feed off each other and reinforce the difficulties of generating value from data portability, we describe each challenge individually.



1. Macro-economic obstacles

1.1 Market imbalance challenges

Allowing all platforms to have an equivalent set of data means giving minor players the opportunity to develop better and offer better services. Data portability also has a very important role in the development and generalisation of "Privacy by design" principles, and ensuring optimal respect for user data. Allowing a greater number of companies to compete in the market with existing digital services forces major platforms to stand out in other ways, such as through innovative features or a more pleasant user experience.

If data portability was available as intended, platforms would stand out through a service that is more respectful of data, and thus change the way users consume the web. Users would no longer come to a platform out of habit, or because no other platform was able to compete with the one commonly used by the majority of users. Users would seek out the platform because it processes data correctly, it processes only the data necessary for the operation of the service, and it does not speculate on the data of its users by selling user data for targeted prospecting, for example.

Therefore, making data portability commonplace would mean allowing all web actors to obtain user data fairly and to focus on improving the services they offer, with high respect for their users and their user's data.

Data portability, therefore, has this role of restoring fair, free and undistorted competition. In fact, when the GDPR was being drafted as a European standard, an urgency to create the regulation was driven by the need to limit the expansion of the big tech giants in the data market, and the potential influence they could have on the protection of personal data, in general.

Unfortunately, by the end of the GDPR drafting phase, the legislation tended to favour these large players at the expense of the small newcomers in digital services, with products that need to establish a place for themselves in this market.

Complying with data portability requirements is costly for all players, and for smaller startups and businesses, their obligations only served to slow down their potential for growth by needing to demonstrate management of user data in a way that enabled data portability.

Complying does not only mean writing the privacy and cookie policy, setting up information mentions or processing only the data necessary for the activity, it also means putting in place the technical and organisational means to frame the transfer of data, to secure the storage of this data, to ensure responses to people's right to their data, and to set up a platform that is interoperable and that can receive the data, and many others. The only ones financially capable of setting up a competent data legal department, then, were the biggest tech platforms.

The economic cost for business was not only about the design of a platform, but also about the constant maintenance of a platform to remain in conformity with legislation. The economic applicability of the GDPR then, appeared too expensive compared to what the market entrants were willing to pay to reach compliance.

1.2 Market distortion challenges

This imbalance on market entry that faces Europe's small and medium businesses seeking to offer innovative digital services continues at a scale that ruptures the equality that works both in their favor and to their disadvantage. Small and medium sized businesses that choose not to make the financial, human and technical efforts to meet their obligations under Article 20 of the GDPR, causes a proportionally unacceptable distortion in terms of competition both with respect to medium-sized companies and large groups, and with respect to those small and medium-sized companies that would have made the effort to make the right to portability effective in the context of their activities, had it been a less cumbersome burden.

By avoiding the regulatory requirements, however, these same companies put themselves in an unfavourable situation by risking their user trust. Users who eventually want to activate their data portability rights may be told that it is not technically feasible to gain access to their data. Thus, these customers will lose trust in the local digital service they have been using and will be driven back to the tech giants, the complete opposite of the intention put forward by the European Commission to justify the European perspective on personal data protection. Yet it is on this trust that the single European data market is intended to be based.

By continuing to tolerate the persistence of a situation of assumed disengagement of small and medium-sized enterprises under the cover of technological unfeasibility, the national authorities responsible for the protection of personal data within the Member States are encouraging them to take the risk of ultimately finding themselves out of the running. The avoidance by these smaller players, due to the implicit form of opting out orchestrated by the unfounded tolerance of the Member States, does not just generate an economic market distortion, but also a legal barrier. In some Member States, such as France, a contract cannot have an illicit object. Therefore, it is unlawful to enter into business contracts, particularly contracts for the transfer of a business or goodwill, where business databases are not processed in accordance with the law. This was already ruled to be the case before the GDPR when, in 2013, the Cour de Cassation (the highest French judicial court) allowed a reduction in the price of a business transfer as the selling business did not store or process its customer data in accordance with the legal requirements in force at the time.

Many companies view the GDPR as a type of hidden tax: an expense item to be assimilated into losses related

to the operation of the company. Without the existence of this Sword of Damocles hanging over the heads of data processors, it would be futile to expect them to be reactive or proactive on data protection issues. Prior to the GDPR, the amount of fines and the likelihood of being sanctioned for failure to comply with the obligations arising from the preceding Directive (Directive 95/46/EC) were too low for any particular attention to be paid to compliance with the legal requirements of the time.

2. Country context obstacles

2.1 Discordant implementations

Article 1.2 of the 95/46/EC Directive recalled that the objective of regulation, which was then taken up in Article 1.3 of the GDPR was to ensure that *"Member States may neither restrict nor prohibit the free movement of personal data between Member States"*.

Before the GDPR, no harmonisation nor standardisation text had been drafted. Each State was responsible for adapting its own domestic data protection law, despite very early legislative positions, as in France with the Data Protection Act of January 6, 1978.

While a first European text brought its semblance of uniformity as early as 1995, and thus well before the real development of data markets and the Internet in general, we can say that the European authorities had taken the lead in a way that predicted the immense influence of the Internet, and all the ensuing legal problems that would evolve from legislating at the European level on avant-garde subjects such as personal data and the protection of individuals on the Internet.

However, in Europe, a Directive only obliges Member States to achieve objectives: it does not commit them to a particular process, nor does it sufficiently sanction them in cases of non-compliance with the provisions contained in the Directive.

And so, between 1995 and the introduction of the GDPR in 2018, there had not really been any evolution in Directive-level policy, despite technical and technological innovations, despite the changing face of the world, increasingly dematerialised, and the growing scale of the Internet and the economic, social and legal stakes these upheavals brought.

To blame the current challenges in interpretation solely on the text that came into force in 2018 would be to partly deny the truth, because the legal gap between conservatism and the applicability of the right to data portability is in fact attributable to the lack of legal maturity on data protection spanning more than 20 years.

Today's current contrast between the legal spirit and the actual applicability of data portability owes its situation to the discordance between the different countries, which do not apply the European regulation in the same way. In France, the CNIL is an independent authority, competent to make decisions and impose sanctions on actors who do not properly comply with the provisions of the GDPR. Other countries do not attach as much importance to the protection of personal data. Some countries do not finance the competent authorities in this area in the same way and with the same intensity, and this is often combined with a lack of the economic means to deploy effective coercive measures.

Even within France, the CNIL received more than 14,000 complaints, an increase of 27%. The total number of staff is around 200, which is relatively little to properly handle all the complaints in order to detect recurring denunciations. As one of the most protective countries asserting the rights of users' data, and one of the strictest in terms of sanctions, the authority in charge of data protection cannot cope with the large number of complaints. This being the case, one can legitimately wonder about the situation for other countries in Europe.

2.2 Lack of enforcement

Apart from the variations in interpretation of GDPR at the Member State level and the resourcing of data protection authorities, the culture and momentum of the national authorities responsible for ensuring respect for the protection of personal data in the Member States also varies. Over the past year, there has been an increase in the frequency and severity of sanctions when sanctions are imposed on data controllers in breach of the provisions of the GDPR, irrespective of the nature and size of the organisations that have been controlled. As part of this study, we were able to discuss data protection challenges for those working as data processors: that is, those responsible for database administration at several major European e-commerce companies.

These key informants confirmed that one of the major industry-wide projects for 2021 will be renewed operational compliance with the GDPR. This follows a recent sanction pronounced by the Commission Nationale Informatique et Libertés, the French authority responsible for ensuring the protection of personal data, against Carrefour. This sanction consisted mainly of a fine of more than 3 million Euros, along with an obligation to comply within a certain period of time. If compliance was unmet, a daily penalty payment would be charged until resolved.

Many data processors and company data protection officers are now being told "You must do what is necessary to ensure that what happened to Carrefour does not happen to us as well".

Despite this new attention to adherence with the GDPR, to date, the study authors have not found any decision by a national data protection authority in any Member State of the European Union that shows that a data controller has been sanctioned because of lack of respect for the right to data portability, or the lack of organisational and technical arrangements to deal with requests to exercise data portability requests.

This inaction is all the more damaging since the European Commission has emphasised that portability plays a key role in its data strategy for the European Union. This silence appears to be due to the current absence of any sector-based standards that national authorities can draw on to assess adherence to data portability, resulting in Member States being tolerant of non-compliance with data portability by companies in their jurisdiction.

The absence of sectoral standards does not constitute an acceptable excuse to allow, at the very least, impunity for non-compliance with the right to data portability. For many years now, there have been technical means of interoperability to which professionals, and moreover the major digital players, are accustomed, to enable the interoperability of information systems. This is the case with APIs (Application Programming Interfaces), or ETLs (Extract-transform-load) and other middleware, which have long been used to meet the interoperability needs of businesses for their operational operating requirements.

3. Company context obstacles

3.1 Reluctance to share data

Data portability consists of the possibility for any individual to simply request a copy of the data concerning them in a format that meets the agreed requirements (such as being machine-readable). However, during the course of this study, it was observed that there was a reluctance on the part of companies to provide this data in accordance with the requirements of Article 20 of the GDPR.

Case studies: 3 examples of data sharing reluctance

facebook

Ruben Verbogh sought to obtain his data from Facebook as part of him exercising his right to data portability. Facebook ended up communicating to him that it would henceforth ignore all his requests regarding the exercise of his rights regarding his personal data, which constitutes an outright admission of non-compliance with the rights enshrined in Chapter 3 of the GDPR. Showing contempt and total lack of consideration for data portability requests that exceed what Facebook considers to be within the scope of the right to data portability, Facebook invited interested parties to go to court to assert their rights.

Source: https://ruben.verborgh.org/facebook/

facebook

Write and civic technologist, Shelby Switzer, sought to download her data from Facebook and transfer to an alternative service. “Facebook imposes very real constraints on the data you can access, from the obfuscation of permissions and data relationships, intentional or not, to limiting access to your friends’ information,” she wrote.

Source: https://www.programmableweb.com/news/i-tried-getting-my-data-out-facebook-quitting-i-even-wrote-code-it-didnt-go-well/analysis/2019/07/02

One study participant who exercised their rights to data portability had to wait four months to get an answer from a major French bank. The response, when it came, was worded as follows:

“Hello XXXXX XXXXX,

Please find attached our answer, your personal data and your right to portability.

With kind regards”

Attached to this e-mail was a PDF document entitled “Right to portability of XXXXX XXXXX.pdf”, which was thus devoid of any machine-readability, and thus unable to be used by any subsequent service or platform. Figure X shows the fields (which amount to personal contact information details) being provided. This is despite European Second Payment Services Directive regulations which obligate banks to provide a means to share customer data with third parties where the customer consents to doing so.

3.2 Winner-takes-all culture

Everyone wants to be a GAFA!

In theory, the right to data portability is an opportunity to enhance competition between players on the Internet. Today, 90% of Internet searches are carried out on Google. Apple and Microsoft have a virtual monopoly on operating systems. About a third of humans have a Facebook account. Amazon is the only market leader in online commercial platforms with a 40% market share in e-commerce worldwide.

All of these companies seek not just to take the lead in their area of online services, they seek to capture all user data worldwide.

This disparity between these internet giants and other digital players who are trying to make a place for themselves is marked by unfair competition. On the one hand Internet giants are companies that have the means to develop innovative services, adapted to the user, with functionalities that can only be found on these platforms; and an abundance of data that makes for endless possibilities. . On the other hand, their competitors try to stand out as well as they can with the little data that they can use strategically to plan their services.

Today the problem remains that responding to data portability requests is seen as a risk to the profitability of the largest actors. Data portability is avoided in order to ensure a monopolistic data market in which only a few players are successful.

3.3 Immature data governance

Our study found that the data provided when responding to data access requests were often the same as those provided by data controllers when responding to requests for the right to portability.

This suggests that data processors have challenges in correctly mapping their personal data media and processing upstream. Yet this is the first step in a plan to comply with the GDPR. There are many reasons for this difficulty, but they can be summarised as relating to immature data governance. Data governance is the policy, processes, and systems in place that enable responsible data management and storage, including ensuring data is of high quality, can be reused, is comparable, and is able to be used as inputs to other processes.

Sometimes, technological and procedural legacy creates difficulties in undertaking an exhaustive mapping of data and managing data processing tasks. Legacy infrastructures that were not built for a data-oriented business are also exacerbated by technical teams that have had to incorporate third-party information systems during mergers or during the absorption of acquired companies by others. The result is an imbroglio of databases and tools that are only interoperable with great difficulty. Errors in this mapping and shortcomings in data governance processes not only imply discrepancies in the compliance with the GDPR of these companies, who are more or less aware of this and hope to slip through the net of the controllers of the national authorities in charge of data protection in the Member States of the European Union, but is also exhibited in the lack of information that is shared with service users who may want access to their data.

Another of the difficulties explaining this incomplete mapping of data, and therefore an equally deficient response to data portability requests, is the delicate communication between the teams made up of employees in charge of complying with the company's GDPR and those working daily to develop and maintain the information system infrastructure. The so-called "technical" teams lack awareness of the GDPR, and even more so of data portability. This has a particularly negative impact on the effectiveness of data portability because it is the developers and those responsible for administering IT systems who are the first to implement this right. It is also a lack of mutual cultural adaptation between these two worlds of legal and IT professionals.

Companies with no internal services dedicated to maintaining their information systems have simply dodged the issue, as if Article 20 simply did not exist. The external service providers to whom they have recourse have in no way grasped the subject, even if this means that they are failing in their duty to provide information and advice to their clients, who are no more accustomed to this change than they are. In the latter case, it is not acceptable to allow these data controllers and subcontractors to avail themselves of the exception based on technical impossibility when the latter do not deign to make the slightest effort to move towards real compliance.

This lack of data governance processes and culture does not only affect lay users, it also affects professionals, who are not always aware of the rules of the European regulation and who cannot afford to pay for the services of a legal professional who would have the necessary knowledge to ensure the compliance of the site or platform.

It is also a question of the benefit drawn by these professionals. Indeed, no guarantee is given as to the gains that data controllers will derive from data portability. It would be quite possible to envisage that those exercising their right to data portability would move precisely towards the global giants of the Internet and totally desert the smallest platforms.

Nor is it certain that people exercising their right to portability from a platform belonging to a GAFA will subsequently transfer their data to a smaller company, although if they did so, the company would need to have put in place the measures to accommodate receiving that data in an interoperable format.

4. Consumer/Citizen/Data subject obstacles

4.1 Information asymmetry

The GDPR, through its protective regime, does not only want to establish an Internet that respects the privacy of users, but it also wants to reverse the unevenness of forces opposing users. At the very least, the regulation hoped to restore a balance between data subjects and data controllers, whose unfettered access had been in the hands of the web's greatest economic players for far too long.

The GDPR has therefore enshrined new prerogatives in favor of natural persons, by devoting an entire chapter (Chapter 3 of the GDPR) to the rights of users over their data and, at the same time, has created more obligations for data controllers.

In this respect, the purpose of the right to portability has a highly consumerist objective in that it gives users back control over their data. It allows users to move more easily between different sites, to diversify digitally, and thus to regain control over their data and transfer of it freely through a machine-readable file system that could (under normal circumstances) be implemented by any platform.

This is somewhat reminiscent of the right to informational self-determination, thought of as the right of all individuals to decide the communication and use of the information concerning themselves.

Moreover, the right to data portability could not be seen as anything other than a right of the consumer, as it is included, for example, in French Consumer Law under Article L224-42-2 and in the legislative code.

However, unlike the GDPR, the provisions of consumer law are real safeguards, protecting the lay consumer from the professional who possesses additional knowledge and resources. Consumer rights law aims to address this balance and give consumers the means to defend themselves, whether by increasing the powers given to consumer associations, or by increasing the obligations on professionals, notably the obligation to provide information that restores a balance of knowledge between the two parties. This recognises that in any normal commercial situation, the consumer enters negotiations as the weaker party in the relationship contract.

The GDPR has tried to do a similar thing by giving the users rights over their data, but it has forgotten this main foundation: that, under normal circumstances, the user is not a professional and does not have the information of which the data controller is fully aware. The right to data portability could have the potential to tip the balance and restore equilibrium between lay users and data professionals, although users must have all the necessary information and have an effective means to seek redress or lodge complaints if their rights are not respected.

4.2 Data literacy gaps

One final concern with our study findings and research into the GDPR data portability experience is somewhat related to this aspect of information asymmetry. In reality, the right to data portability does not reflect the lived experience of most Internet users: most are not aware of their rights, the regulation is not adapted to their needs, and it assumes a degree of data literacy maturity and a dynamic market with users expecting to make use of their rights. Our engagements with study participants further confirmed that this is not the digital society environment facing the majority of users.

As far as data portability is concerned, the large majority of the population do not have the necessary data literacy to deal with the data economy. The applicability of the right to data portability remains legally delicate, as it is difficult to understand for any average person, but also economically too expensive in relation to what the tech giants are willing to pay to achieve compliance.

In addition to the knowledge gap that separates users and data controllers who hold the data, the GDPR does not accurately reflect the reality of the practice of data portability, particularly in the process of exercising this right, which currently requires a great deal of patience on the part of users, in a digital world where everything else is accessible in a matter of minutes or seconds.

Why would a user wait a month to be able to transfer their data to a competitor's service when they can simply create an account on that competitor's service in a few minutes?

The right to portability is also not adapted to reality in the sense that it claims that any data controller will correctly implement the organisational and technical measures for the correct exercise of the right to portability. Unfortunately, this goes against the current mentality of Internet economic companies that want to hoard user data, that have understood for several years that data is a high value resource in the 21st century, and that have had time to grow, to assert themselves and to convince the whole population that their services are the best, that few competitors cam match them, and that have become practically untouchable.

This is the core of the whole legal divide that arises from the right to data portability: the legislator's desire to redistribute data by freeing it from the monopoly of the Internet giants by regaining consumer control over this data, materialized by this right, was not preceded by the text nor was it properly inculcated in the major players, the data controllers.

400 million Euros in GDPR fines: but how much for portability sanctions?

Of the 400 million Euros in fines imposed by the various data protection authorities across Europe, no fines have been imposed for an infringement of the right to data portability.

The problem is not so much that no fine has been imposed for a breach of data portability law. As the study found, what is more concerning is that not only are there no fines, but that this is occurring not because disrespect for data portability is an isolated case, but far from it.

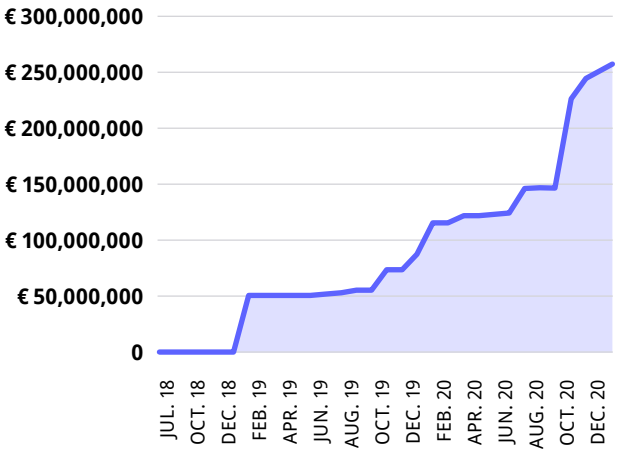
Regardless of the data controller, from the Internet giant to the small associative data controller, few were able to provide user data in a machine-readable and interoperable format. The violation of this right was evident amongst small, more excusable players, who would not have the capacity, the knowledge to set up a correct portability. Concerningly, it was also evident with the large Internet players, that have the logistical, material capacity and the knowledge necessary to develop a data portability system that is functional and of value to the user.

The total absence of fines for non-compliance with the right to data portability is perhaps the best example of how little interest can be shown in this right by both data controllers and users.

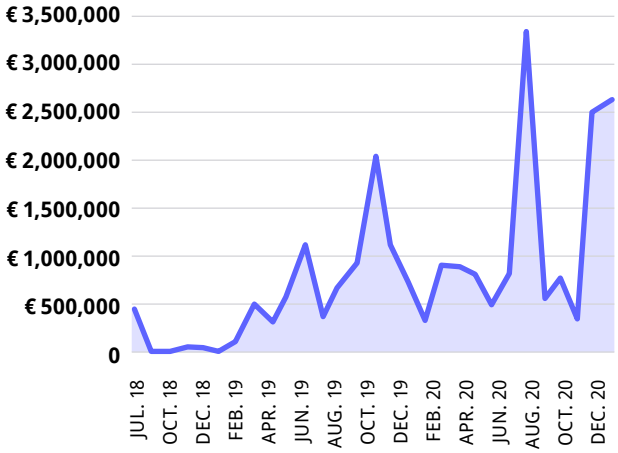
Moreover, sanctioning a non-respect of this right would be quite conceivable when we look at the grievances against the offending companies, in particular the fine imposed on Google, on the main ground of "*lack of transparency and accessibility of information mentions*", or the fine recently imposed on Carrefour, sanctioned in part for the non-respect of various rights (right of access to data and the right to delete data for which the deadline was not respected).

So, in view of the sanctions and grievances upheld, it would be possible to envisage an initial jurisprudence on the non-respect of the right to portability, especially given the importance that European entities are currently attaching to it with the *Data Governance Act*.

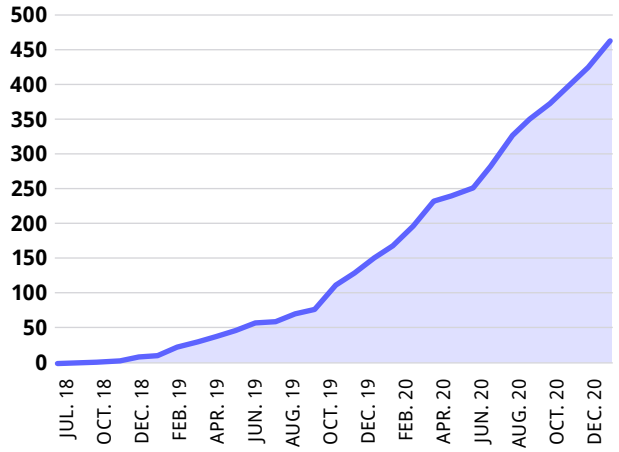
Course of overall sum and number of fines (cumulative)



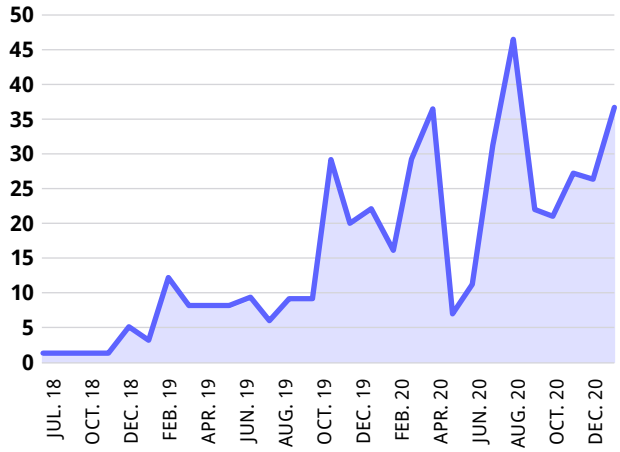
Course of overall sum and number of fines (non-cumulative)



Course of overall number of fines (cumulative)



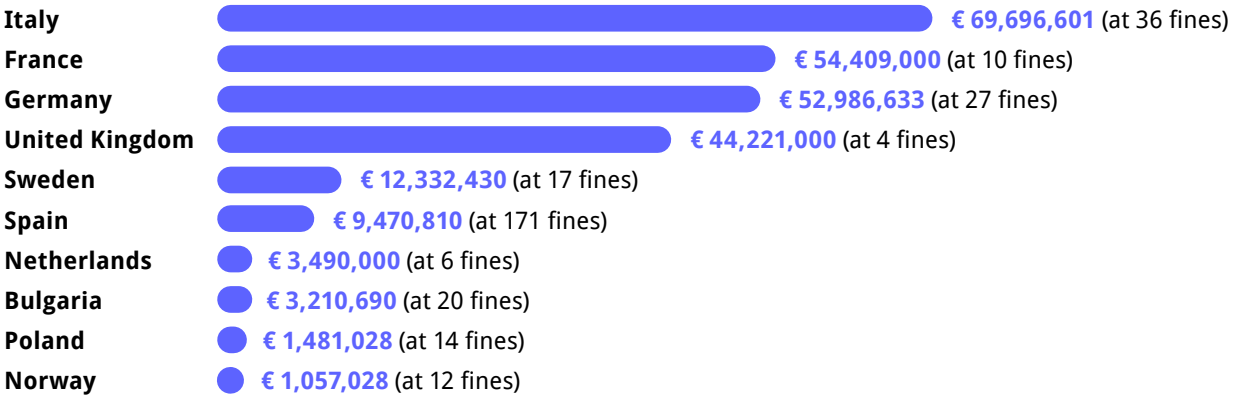
Number of fines per month (non-cumulative)



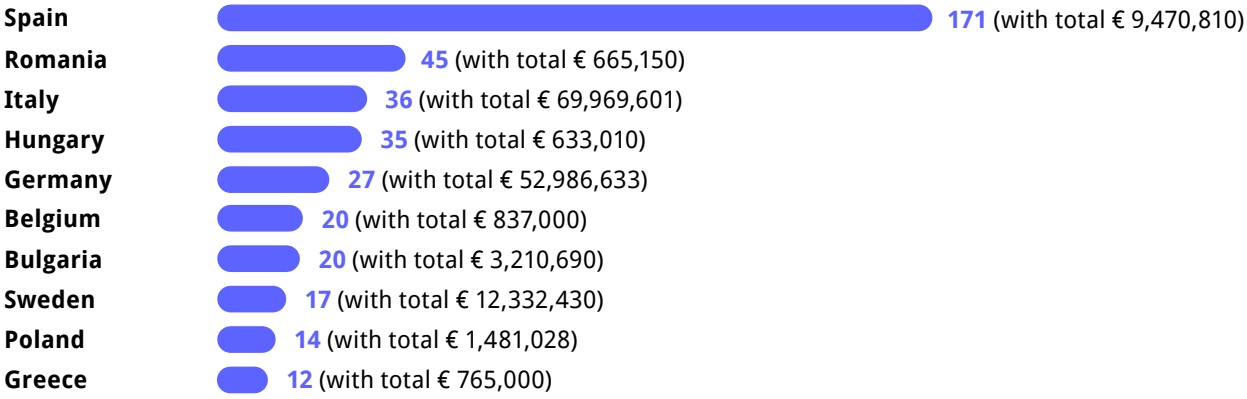
Statistics: Countries with highest fines (Top 10)

The follwing statistics show how many fines and what sum of fines have been imposed per country to date (Only fines with valid information on the amount of the fine are taken into account).

By total sum of fines



By total munber of fines



Findings from other recent studies on data portability

Data portability can encourage market competition, enable innovation and assert the rights of individuals over their data and the data collected about them. Other recent studies and journal articles have examined current practices and sought to measure the impacts of data portability.

The exercisability of the right to data portability in the emerging Internet of Things (IoT) environment

Authors: Sarah Turner, July Galindo Quintero, Simon Turner, Jessica Lis, Leonie Maria Tanczer

Link: https://journals.sagepub.com/doi/full/10.1177/1461444820934033

Summary

This research aimed to deliver the first empirical investigation on Article 20's exercisability in the nascent IoT environment.

As part of the research, two studies were conducted:

- 160 privacy policies of IoT vendors whose products were available for purchase in the United Kingdom were reviewed.
- Four widely available IoT systems were tested.

The goal was to understand the level of information offered to data subjects regarding data portability and to determine whether those actors put procedures in place in order to enable users to exercise their right to data portability.

The analysis of the four IoT systems emphasised on the barriers the data subjects faced when exercising their right to data portability. For example, the format in which the data was sent did not allow for a transfer to a secondary data controller. The criteria of Article 20 were not respected. None of the four IoT system data controllers agreed to transfer the data directly to another data controller. Users had difficulty understanding the nature and extent of the data they received.

Study conclusions

The results were far from positive. Only 63 out of 160 privacy policies (39%) explicitly referenced data portability. Specific issues identified included:

- There was difficulty for users to understand the meaning of the right to data portability through the privacy policies.
- There is an opacity in the language used, creating confusion.
- There is a lack of information on the data transfer process from the original data controller to a secondary controller.
- There is a huge margin for improvement.
- Processes need to mature in order to be effective in enabling data portability, including data transfer mechanisms.
- Technical measures and better guidance are needed from the European Commission and from Member State Data Protection Authorities. This is the same conclusion we came to in our study, as we witnessed the same obstacles to an effective portability.

Alignment with our study

This study reinforces the experiences we saw in our research, but with a deeper focus on Internet of Things environments. Many of the challenges regarding opacity and lack of clear processes is relevant to the entire data portability context.

Data portability among online platforms

Author: Barbara Engels

Link: https://policyreview.info/articles/analysis/data-portability-among-online-platforms

Summary

This study aimed to examine "the effects of the right to data portability on competition, providing policy recommendations for the preservation of innovative, undistorted competitive digital markets." It looked at how data, users, and platform services are connected, and how these relationships change under data portability.

The author notes that competition enforcement needs to occur through a case-by-case assessment. The duty of data controllers and the data portability rights of data subjects as defined in the GDPR need to be distinguished. The main focus of this study was to differentiate between platforms offering complementary products and platforms offering substitutes. As such, it suggests that the GDPR must be "interpreted in a nuanced fashion", in order to take into account the specificities and complexity of the market in order to avoid creating more barriers to the development of new digital business models.

Study conclusion

The study calls for more empirical research on the multi-faceted competition effects of data portability, which are currently lacking.

Alignment with our study

Despite being five years old, the conclusions are still relevant today, as there remains limited evidence of the competition effect of data portability. Our study has discussed the potential value that can be generated by data portability, and the shortcomings of current data portability processes which prevent this value from being realised, just as this previous study had warned.

Dude, where's my data? The GDPR in practice, from a consumer's point of view

Authors: Hanne Sørum, Wanda Presthus

Link: https://www.researchgate.net/publication/342315202_Dude_where%27s_my_data_The_GDPR_in_practice_from_a_consumer%27s_point_of_view

Summary

Researchers calculated the response time and the type of responses received data portability. Data portability and access rights were sent to 15 companies. The researchers noted that the companies concerned did not give detailed explanations when providing the data. In this study, only one data controller failed to send the data in a machine readable format. The most common formats used were HTML, TXT, JSON and CSV.

Study conclusion

More companies contacted were successful in meeting the data portability requirements than then data access requests.

Alignment with our study

While we are surprised at the positive outcomes from data portability requests, which differ from our own, other findings were more similar, with the researchers concluding "it is evident that the companies do not really differentiate between" data portability and data access. The researchers noted the confusion the process raised, and call for standardisation, which aligns with one of our recommendations.

How to attribute the right to data portability in Europe: A comparative analysis of legislations

Authors: Barbara Van der Auwermeulen

Link: https://www.sciencedirect.com/science/article/abs/pii/S0267364916302175

Summary

This analysis states that "the restriction to data portability can be sanctioned by European Competition Law if it qualifies as an abuse of a dominant position as mentioned in article 102 of the Treaty for the Functioning of the European Union (TFEU)." This point resonates with the findings of the above "Data portability among online platforms" article. However, in this research, the author analyses when article 102 of the TFEU can be applied to data portability and to whom instead of analysing the effects of data portability on competition.

It is interesting to note that the author discusses the fact U.S antitrust law could possibly be a source of inspiration for European legislators when it comes to data portability in the context of European Competition Law. Indeed, instead of focusing on privacy legislation, in the United States, the discussion on data portability is addressed through the application of antitrust laws. "Therefore, emerging online service providers could win monopoly claims if they prove that their competitors are violating antitrust law by not providing or supporting data portability tools." The author tries to analyse which law between the European Competition Law and the GDPR could be more effective in order to make the right to data portability more effective.

Study conclusion

The author concludes that Article 102 of the TFEU may apply to some situations, however, it is hard to apply it to online services, as the conditions are difficult to meet.

Alignment with our study

Similar to our findings, this research notes that current legislation is not effective enough, and some changes need to be made.

The Right to Data Portability in Practice: Exploring the Implications of the Technologically Neutral GDPR

Authors: Janis Wong, Tristan Henderson

Link: https://janiswong.org/publication/wong-exploring/

Summary

In this study, the researchers focused on the execution of the right to data portability request by users. The pool of data controllers that is at the core of this study is 230. In order to exercise those requests, a Python program was used. The goal of this study was to analyse the process involved in exercising one's data portability rights and to assess the format of the data received.

The results show that only 172 requests were successful, and not all of them respected the format requirements of Article 20 of the GDPR.

Study conclusion

Some interesting issues were raised in this study. Some data controllers asked for feedback on the data the user received. They wanted to know if the data was satisfactory, what format they wanted the data to be sent in, how was the communication process, and whether the responses were fast enough. Moreover, some mentioned they did not know whether the requested data fell under the jurisdiction of the GDPR. One data breach happened, when a user received the data of another user.

In several cases, it appeared that data controllers were not yet familiar with obligations nor how to process data portability. One data controller confirmed they had never previously received a request. The study highlighted that it may not always be a case of acting in bad faith, but instead that data controllers do not understand what is expected or how to carry out their obligations to meet user's right to portability. The study notes that some data did not meet interoperability requirements. Moreover, the authors noticed that data controllers had to take into account some categories of data, thus that some categories had to be sent in specific format. They were often unsure on what format to send the data in.

Alignment with our study

Similar to our conclusion, the authors suggest the need to standardise the data portability process. The identify the need to create "new data portability definitions, clarify how data should be made portable, and explain the appropriateness of file formats in relation to how data could be determined, according to their type or industry."

We align with their recommendations that there should be a more technically advanced definition for 'structured, commonly used, and machine-readable' in order to make sure data portability practices are actionable. To do so will require a collaboration between lawyers, policymakers, enforcement bodies, data controllers, and technologists to ensure that data portability is viable in theory and in practice. Like us, the authors also note that technological solutions may be able to make the process for data portability requests easier for data subjects and data controllers.

The right to data portability in the GDPR: Towards user-centric interoperability of digital services

Authors: Paul De Hert, Vagelis Papakonstantinou, Gianclaudio Malgieri, Laurent Beslay, IgnacioSanchez

Link: https://www.sciencedirect.com/science/article/pii/S0267364917303333#!

Summary

This paper describes how open to interpretation the GDPR is and how this could lead to additional challenges in implementation. As such, the goal of this article is to propose an interpretation of the data portability right, by suggesting a pragmatic approach while taking into account the state of the digital market and the fundamental rights of users.

To commence, the authors focus on Article 20 itself. They suggest that it may have been written in a vague manner on purpose in order to anticipate future technological developments. They note that expressions like "not adversely affecting rights and freedoms of others" and "without prejudice" allow judges to adjust solutions on a case-by-case basis.

Study conclusion

The authors offer two interpretations of the expression "provided by the user". It could be defined as referring to only the personal data that the data subject has explicitly provided or it could be defined as all personal data that the data controller has collected upon consent or according to a contract. This second interpretation includes the data that was "observed" by the data controller.

The authors note that the European Data Protection Board recommends that, in order to be effective, the right to data portability should have "a wide scope of application, and not only be applied to the processing operations that use data provided by the data subject". What is for sure is that it should be a case-by-case analysis by the judge, but those elements help determine what interpretation is more fit.

Interestingly, the authors also refer to Recital 68 of the GDPR, which refers to "his or her own data". This emphasizes the relation of the user to their data, and the importance in ensuring the data user's right to keep a certain level of control over their data.

Alignment with our study

Similar to our research, the authors recommend applying the more extensive definition to data "provided by the user" when interpreting and acting on the right to data portability.

Recommendations

Educate

Education of users on their right to GDPR data portability is essential.

The European Data Protection Board should be responsible for creating a series of educational resources to explain to citizens and businesses their rights to data portability and to explain the process of how to request data portability.

Any service or platform collecting data about users should have site resources that explain portability. At a minimum, they should share the educational resources suggested above, created by European Data Protection Board.

Educational resources should include:

- Plain language explanation of data portability rights, as defined under Article 20 of the GDPR
- Simple workflow diagram that shows the ideal process from request to receipt of data files and key dates and time frames in which citizens can expect action to progress
- Video/animation resource with subtitles available in all European languages
- All resources should be created in line with W3C Web Accessibility Guidelines

Investment in an awareness campaign should also be conducted, perhaps in conjunction with the establishment of the European data spaces, for the annual International Data Protection Day (in January each year). Data innovation hubs may also be a natural education partner, as businesses participate to support developments in which everyone can benefit from citizens and data users sharing their data from one service and platform to another.

Cell phone portability allows a consumer to change operator while keeping his or her fixed phone number. This is done by the new operator. **Everyone understand their right to phone number portability**. The process is very simple and clear, the customer simply has to ask their new operator if they can keep their phone number. When this was introduced, governments and competition authorities promoted the phone number portability on TV ads and in newspapers. We do not see yet the same effort being made for data portability. In Europe where fair competition is promoted, fair digital competition should have the same treatment.

Simplify

The process of GDPR data portability is so complex that it is itself a barrier to accessing and using one's own data.

Focusing on the user experience and creating simple consent flows could improve use of GDPR data portability rights. A data portability assessment guide or certification could also be a solution, with recommendation on design, and user flows about how to handle data portability requests.

Service users should be able to request information through a variety of mechanisms:

- **By mail/email/online contact form:** Platforms and services could make a template or online form available to more easily exercise their data portability rights. At present, services often just give the contact email for the delegated Data Protection Officer, making a simple portability request a daunting and intimidating experience.
- **Within user accounts:** Platforms and services could create a simple button within a user's account dashboard to allow seamless transfer of the user's data. In cases where the user seeks to share data directly with another service or platform, the interoperability capabilities of this functionality will be essential.

Apply a takeout system: A small number of platforms and services are using an integrated takeout system, and some are making this the only mechanism by which users can request access to their data for portability. Google even allows for personal data to be exported directly to competitor services, such as data storage providers. However, there are far too few examples of access being enabled in this way.

The Data Transfer Project, a collaboration between Google, Microsoft and other big platforms have demonstrated that it is possible to make it easy to include data portability features directly in the user interface. (However, it must be noted that with all of the resources of the big tech giants, they have not been able to progress beyond a proof-of-concept since 2018).

Standardise

Formats for sharing GDPR data portability results should be standardised.

The current incompatibility of data formats with which data can be provided should not be a valid argument for refusing to respect the right to GDPR data portability. Under the European Data Protection Supervisor's Guidelines, platforms and services can select an appropriate format. While this is understandable to avoid additional technical burdens on operators, the digital economy has matured sufficiently that standards should be expected for sharing data. Under Europe's Digital Single Market goal, interoperability is given prominence. Standards help achieve interoperability.

Establishing an API standard would make it possible to overcome the difficulties linked to the incompatibility of computer systems when companies make a citizen's data available for portability.

The banking sector, under the PSD2 Directive, is a useful example of what standardisation can offer. By opening up information systems using APIs, external services such as rideshare services can offer payment functionality directly from their application. This allows banks to offer a better user experience for customers.

4 Develop alternative models

By truly enabling GDPR data portability, a new market ecosystem can be fostered which would create a range of different roles for data institutions.

As shown in the following diagram from Mozilla, new types of data organisations can develop including data stewards (companies that help transfer personal data between services), data unions and cooperatives (where people can pool their data for beneficial use, either for individual benefits such as rewards or for common social goods like use in new health research), data trusts and fiduciaries (where data custodians can make financial decisions around how to 'invest' someone's data), and other models.

Forthcoming legislation from the European Union on a Data Governance Act is expected to introduce some of these new institutional forms. It appears to suggest that large platforms will need to separate the data collection and stewardship functions from the reuse of their app user's data for commercial benefit.

New models of data intermediaries are also suggested in the legislation. These entities would act as nonprofit, independent platforms responsible for facilitating the exchange of citizen and business data between agreed parties. The proposed Act also recognises the value of data altruism, where individuals may donate their data to research or for social good purposes. In each of these endeavours, it will be necessary for these new forms of data institutions to play an intermediary role in facilitating data portability.

As this new ecosystem grows and models are tested, the ability of intermediaries to enable timely, appropriate data portability at the user's request will be an essential indicator of whether these new institutions will be able to function effectively.

SOURCE: MODELED ON HTTPS://FOUNDATION.MOZILLA.ORG/EN/INITIATIVES/DATA-FUTURES/DATA-FOR-EMPOWERMENT/

Facilitate and build the transition

New tools and startups are emerging to build the next generation of data portability tooling.

DAPSI, an innovation incubator for technological solutions and services that ease GDPR data portability, is funded by the European Union as part of the NExt Generation Internet (NGI). In the initial phase, operating until February 2021, 11 projects were chosen to develop a proof-of-concept to aid data portability. A second phase, operating March to June 2021, will support go-to-market readiness for shortlisted projects.

Examples of emerging data portability technologies aimed at supporting data portability

- ALIAS: Enables the next generation of applications to happen, by automating GDPR portability for applications developers.
- Checkpipe Charlie: A tool for describing and validating data.
- DIP: Human-centric Vaccination & Immunization Management using Verifiable Credentials.
- Domi: SSI-based digital passport to facilitate data portability in the housing rental sector.
- DPella: Data analyses with privacy in mind
- IDADEV-P2P: Blockchain Based Data Portability System.
- OpenPKG: A decentralised data provenance system for improved governance and portability of personal data.
- OpenXPort: Open export of data across different systems and providers.
- ORATORIO: Energy data exchange platform.
- Prov4ITData: Provenance-aware querying and generation for interoperable and transparent data transfer.
- UI-Transfer: A complete solution for the "user initiated inter-controller and continuous data transfer" pattern.

An emerging suite of tools are also becoming available to support data portability. These include:

- Udaptor (a Chrome extension that assists with data recovery).
- The European eSSIF Lab, a non-governmental collaborative project that supports the creation of technological tools to promote interoperability between companies).

A full range of initiatives funded by NGI incubators are available at www.ngi.eu/

Mozilla also curates a list of projects building on transition to GDPR data portability accessibility at www.foundation.mozilla.org/en/initiatives/data-futures/who-is-trying/

Join efforts as a community

Individuals, companies and agencies can support groups that are committed to improving GDPR data portability.

These groups work all year on the work of making data portability a reality according to regulations, and also to improve the regulation at the same time.

These groups include:

- MyData.org
- Mozilla Foundation
- Privacy international
- Radical Exchange Institute
- None of Our Business
- Digital Commoners

Individuals and businesses can sign on to a Declaration of MyData Principles. MyData aims "to empower individuals with their personal data, thus helping them and their communities develop knowledge, make informed decisions, and interact more consciously and efficiently with each other as well as with organisations". Under data portability right principles in the Declaration, MyData note:

The portability of personal data, that allows individuals to obtain and reuse their personal data for their own purposes and across different services, is the key to make the shift from data in closed silos to data which become reusable resources. Data portability should not be merely a legal right, but combined with practical means.

7 Mandate APIs

Currently, companies are required to put in place technical means to enable portability, however, no further direction is given. The industry-led Data Transfer Project was established to create an open source inter-service portability platform to facilitate data transfers between services. However, the main members of the project are Facebook, Twitter, Apple, Google, Microsoft. While it may be a good initiative, it reinforces an already existing monopoly, and it could be argued that it is an attempt to show "doing something" rather than actually implementing solutions (despite the global wealth and resources of the collaborators, there has been little progress since 2018 in creating solutions as part of the project). Portability should enable innovation by smaller companies, and they should be involved in these cross-industry portability standards initiatives.

To make the right to portability effective, companies must implement automated tools to extract the relevant data. To facilitate data transfer between platforms, automated systems such as application programming interfaces (APIs) can be used. This principle is already provided for in the current GDPR, but APIs are not stipulated as the mechanism for automation, which has led to confusion and lack of implementation. Other governments around the globe are facing similar issues. For instance in 2019, bipartisan US senators proposed a bill enabling users to get their data back from data platforms with APIs.

To date, European Commission legislation and directives have been reluctant to articulate the role APIs can play to exchange information. This has limited effective action and has led to ongoing fragmentation. APIs are a general purpose technology that should be specifically referenced as the preferred solution within policy documents. This would still allow a breadth of implementation decisions to be made based on current technological developments (for example, there are a range of API protocols and architecture designs that each offer specific advantages and limitations).

Stating that APIs are the technology to be used for interoperability and data portability would reduce the risk of individual solutions being created for one-off use cases. Other European Commission strategy documents have noted this risk in recent years. The evaluation of the former public information directive, for example, noted that because "via API" was not stated, there had been little progress in creating standard means of exposing public sector information for reuse. The revised Open data and Public Sector Information Directive has sought to overcome this obstacle by stating that high value datasets should be made available as dynamic data using APIs.

In a similar way, guidelines from the European Data Protection Board could insist that services and platforms provide APIs as the mechanism for automated data portability.

8 Create the case for GDPR data portability fines

Out of the 500+ fines issued since May 2018 (as described in the GDPR enforcement tracker), European data authorities have NEVER penalised any company or institution for lacking portability engagement. Article 20 has never been mentioned in any rulings.

To strengthen recognition of portability as a regulatory requirement to be managed appropriately by platform and service user Data Protection Officers, at least one data regulation authority must set a precedent and review data portability concerns and set appropriate penalties.

This will send a strong signal to companies and their Data Protection Officers that portability is a fundamental right under the GDPR and must be followed, and that not respecting it is a serious compliance threat risk to a company's regulatory obligations.

Community advocacy organisations could support individuals to create test cases that could be submitted to various data regulation authorities. For example, the Norwegian Consumer Council has conducted reviews in partnership with a range of digital rights organisations to review unconsented data flows in dating apps. This has resulted in substantial fines for non-compliance. Similar test cases are needed to ensure regulatory enforcement of GDPR data portability rights. Advocacy bodies such as the European Consumer Organisation (BEUC) have also been involved in filing complaints against digital rights breaches.

Disincentivize data retention with a digital VAT on data

In 2013, a report on fiscality for the digital world by french fiscal administration reporters Collin and Collin proposed to apply digital Value Added Taxes to companies that don't give back data to their users.

Following the model of value-added taxes, in which a company does not pay taxes on production processes as long as they are adding value to a product and selling it to someone else. It is the end user, the final consumer of the product, that does not add value to the product and therefore pays the Tax on Added Value, the VAT. This principle could be applied to the digital economy. As long as a company doesn't return the full data to their users, they could pay a tax on the value of the data they keep, that is, on its indirect valuation. The amount of taxes paid can be calculated on the average revenue per user, a percentage of the capital value of a user in a certain region, or a percentage of revenue made in the country.

Impose API neutrality for platform monopolies

Platforms give access to their data and their users' data via APIs according to specific terms of services. In these terms of services, they often give themselves the right to revoke access for any reasons, according to their business judgements. For instance, if they consider a user is running a business model in direct competition with them, or if the user is extracting too much data form their platform, or that a user re-uses the data in a way they don't like, they can decide to unilaterally cut off the access to data on the platform. There are many known cases of hard-cutting API access by companies like Google, Twitter, Netflix, Linkedin , Facebook and so many others. In a 400+ page report and the US Antitrust Commission accused Facebook of using APIs access as an anti-competition practice.

This can be solved by obliging monopolistic platforms to give access to their API in a neutral way. As Professor of Law Jonathan Zittrain explained in his book in 2006, *The Future of the Internet and How to Stop It*, API neutrality would apply net neutrality principles to APIs: regardless of whether companies were competitors or not, open APIs would be provided to ensure full data portability when requested by the user. In this way, portability would be guaranteed with the same level of quality, with the minimum user experience fatigue for users and the maximum efficiency for competition.

In a minimalist portability approach, this neutrality of APIs could be made available only for users themselves to have API access to their data for portability, and linked directly to data storage services or managed by data stewards to it is not used as a backdoor to unbalanced market competition.

Conclusion: Where next for data portability?

Emerging policy drivers suggest there can be a renewed focus on improving data portability rights. Under new Data Governance and Digital Services Markets legislation in Europe, there is the opportunity to address many of the gaps and obstacles identified in this study.

The right to data portability is conceived as a value-generating opportunity that could allow local, minor players to enter markets and expand their user footprint. For citizens, it represents an opportunity to enter the data economy and engage with the value of the digital capital in new ways, and to move between platforms and services as they wish. For society, it could help wrestle control for an ever-shrinking pool of Internet giants who restrict and exploit user data for their own advantage. Supporting the development of practical processes that enable the right to data portability could result in data circulating more freely, where user data is no longer held in the hands of just a few data controllers.

www.gdpr.dev

ALIAS