



PersonalData.io

Geneva, Switzerland
contact@personaldata.io

March 10th, 2022

European Data Protection Board

Dear Sir or Madam,

We are different entities (associations, groups of users of digital services, company, individuals) having in common the defense of individual and collective digital rights and their improvement. We have gathered under the banner PersonalData.io to thank you for your work in advancing the digital rights of Europeans and in particular for your public consultation on the "Guidelines 01/2022 on data subject rights - Right of access" for which you will find our contribution attached.

The NGO PersonalData.io is focused on "making data rights individually actionable and collectively useful" and has been founded in the wake of the Cambridge Analytica scandal. The founder of PersonalData.IO, mathematician Paul-Olivier Dehaye (now CEO of Hestia.ai and co-signatory of this letter) has testified on the right of access in front of the European Parliament (alongside Andrea Jelinek), at the Council of Europe, the UN, the World Trade Organization and the UK Parliament. The association is now chaired by Jessica Pidoux (sociologist and author of a doctoral thesis on dating applications).

Our goal is to use data rights to make digital society more legible. While we are convinced that it is possible and that the key will be the collective exercise of the right of access, we have so far had to focus on the individual exercise of the right, in order to defend its very relevance. Our work is now shifting towards the collective exercise of that right though.

In our choices of very early cases, right after the coming into force of the GDPR, we have tended to focus on cases where we could foresee media interest and a broad shift in understanding of the issues and the interest of that right. We would like to highlight a few examples of our work:

contact@PersonalData.IO

- helping David Carroll, a New York academic, to exercise his data rights towards Cambridge Analytica - this became essentially half of the Netflix documentary *The Great Hack*.
- helping French journalist Judith Duportail to get access to her Tinder data, which led to a widely shared article, a book and a documentary on discrimination in dating apps.
- helping Uber drivers (such as James Farrar in the UK) to obtain their Uber data and understand how the right of access could be used fruitfully; this has eventually led to some jurisprudence on the matter from Amsterdam, amongst other legal actions.
- helping journalist Carole Cadwalladr to obtain her car insurance price comparator data, in order to prove that data from businessman Aarron Banks's insurance company had illegally flowed into the Leave.EU's campaign, for which they were sanctioned by the Information Commissioner's Office.
- thanks to pressure applied through precise SARs, changing Facebook's transparency tools to include some *Custom Audiences* data, and providing material to Senator Blumenthal to directly highlight duplicity of Mark Zuckerberg in his testimony to the US Congress, which has led to broader coverage of the tool to include *Facebook Pixel* tracking data.

As explained earlier, our focus is now shifting towards transforming these individual steps into collective actions. The other co-signatories of this letter are involved in this dynamic:

- members of the data collective *The Eyeballs* are collectively using the right of access to data to reveal the opaque mechanisms behind the monetization of attention, and have for instance helped the *Tracking-Free Ads Coalition*, a group of EU Parliamentarians working with journalist Mark Scott, in uncovering in their Twitter data how Meta was targeting different actors of the Brussels bubble.
- Members of the data collective *Dating Privacy* are using the right to fight against the risks associated with the use of dating apps, to increase the algorithmic transparency of matching systems, and to highlight their discriminatory effects. We are for instance interested in understanding algorithmic effects in situations of cyberviolence, exclusion, or socio-economic discriminations on the basis of our digital identities and behaviors. From the data obtained through SARs, we can understand algorithmic amplification of societal stereotypes as well as their influence on human behaviors (sometimes offline). It is urgent to find solutions to cyberviolence and discrimination problems in a society of algorithmic "personalization" edging ever more towards polarization and fragility.
- Associations of French for-hire drivers such as INV or VTE are using this right to try to rebalance the power between the platforms and workers.
- through a collaborative investigation *#digipower* instigated by the Finnish innovation fund SITRA and involving 15 high level participants in making sense of their personal data together.
- through supervision of various student projects, especially with the University of Geneva.
- through a reader-led investigation with the Swiss newspaper *Le Temps*.



One of us has worked in condensing the sum of those experiences as a response to the consultation, which is attached. We hope it will be useful to you, thank you for your consideration, and stand at your disposal for further clarifications.

With our highest regards,

Jessica Pidoux, Director of PersonalData.IO, Leader of the *Dating Privacy* data collective, postdoctoral researcher at Sciences Po Paris, CEE

Brahim Ben Ali, Secrétaire Général du syndicat INV (for hire drivers)

Paul-Olivier Dehaye, Member of PersonalData.IO, CEO of Hestia.ai

Charles Foucault-Dumas, Leader of *The Eyeballs* data collective

Jacob Gursky, Member of PersonalData.IO, and *Dating Privacy* data collective

Judith Herzog, Member of PersonalData.IO

Louis Poncet, member of MyData and CEO of NEEDS services

Jean-Christophe Schwaab, Member of PersonalData.IO, ex-president of the Commission on Judicial Affairs of the Conseil National (Switzerland), dr in Law

Marie-Pierre Vidonne, Member of PersonalData.IO, member of *The Eyeballs* and *Dating Privacy* data collectives

PERSONALDATA.IO

The Eyeballs
We are the targeted audiences

dating
privacy

 **Hestia.ai**
sustainable data solutions



Response to the consultation
by the
European Data Protection Board
on the
Guidelines 01/2022 on data subject rights - Right of
access,
as adopted on 18 January 2022

Written by Judith Herzog, jurist and member of PersonalData.IO, with input from other members of association PersonalData.IO. This was made possible through funding provided by Hestia.ai.

Abstract

There are many welcome clarifications from the EDPB in its draft guidelines to address the practicality of the right of access. Doubts remain, however, in view of the different margins of implementation. We wonder for instance what will be the calibration periods and frequencies of access requests in the case of high frequency processing in ad targeting ecosystems, or how will access be exercised vis-a-vis so-called privacy engineering technologies such as edge machine learning or third party server ad attribution. We also note that the body of case law is likely to change significantly in the near future. We observe a lot of ongoing legislative activity, and particularly on the development of various intermediate actors for the management of data sharing/consent for diverse purposes. Our fear is that this will result in a limiting understanding of individual rights if these schemes are only geared towards sharing to third party legal persons, with an assumption that perhaps goes quickly to associate civic participation with donation, or representation of interests by intermediary bodies. This could harm the concurrent development of channels to re-use the data for oneself, or prevent freeing up space to accommodate explanations coming from experts of our choice in addition to the data controllers. This is why we try to structure our contribution around data mediation capacities and prosocial data activities by individuals.

The guidelines contain promising elements to compensate for this tropism. They consolidate the perception of the access right as a component of a fundamental right attached to individuals, without the prerequisite of justifications towards controllers as to its usefulness for a subsequent cause, and emphasize the cross-cutting issue of helping them to understand the calculations of which they are the subject.

We hope that this contribution will be useful to you and thank you for the many useful clarifications as to the affordance of the exercise of the right of access.

We discuss in order matters concerning:

1. Scope of the right of access
2. Data subject's identity and data subjects' data identification
3. Understanding the Data
4. Involvement of data intermediation/sharing actors
5. Exercising the right of access for pro-social purposes
6. Access rights and freedom of expression

1. Scope of the right of access

Components of access rights

The EDPB could encourage controllers to respond to access requests with a summary that follows the structure set out in the table at §16 — "*article 15 can be broken down into eight different elements*" — with respect to article 12(2) GDPR on information to be provided about processing. The same logic should apply to requests combining article 15 and 20, or other subject information rights.

This would help recipients to understand whether or not the controller has responded to all of these elements, and also obtain justifications when this is not the case ; at the outset rather than through more email exchanges. Indeed, responses to requests are often structured according to the controllers' own activities and vocabularies, thus leaving a risk to mislead people to believe what they received was exhaustive.

Scope in time

Several passages in the text illustrate the ambiguity due to fluctuations over time in the scope of data accessible to data subjects. This is particularly apparent where the controller is asked to update their response to reflect "*the processing operations that are carried out in relation to the specific person making the request.*" (page 3)

If part of the data has been deleted, for example after the retention period has expired, it is still possible to keep track of the types of processing and data that may have been associated with a person. In principle it is possible to have this explained by the controller even if the data itself has been destroyed¹ (cf. §116). In this sense, we suggest to include "operations that are or have been carried out so far". This would be particularly useful where data has been collected from other sources (Art. 14 GDPR) under a legal basis that does not require prior consent.

¹ This would be consistent with the advice to keep a record that the person's identity document has been verified rather than the identity card itself (§78).

Furthermore, it would be interesting to recommend that data controllers provide summaries of the successive versions of their privacy policies with the corresponding periods. This would be particularly interesting in the context where individuals retroactively study past data processing concerning them, e.g, to find out what they are entitled to expect and when to exercise their rights².

In another time direction, we can think of situations where data controllers fail to specify that data collected by the controller, and/or additional information provided by the data subject at one point is not personal, but may become so later. For example, if identifiability gains certainty³. As previously suggested in the consultation on the right to portability, it would be good to have guidance on that⁴.

2. Data subject's identity and data subjects' data identification

It is welcome that the guidelines confirm controllers should not refuse to take additional information provided by a data subject (§64). However, when identifiers are used that are unknown to the public, data subjects are often led, paradoxically, to put themselves in an adversarial position in search for these informations by themselves.

Identifiers unknown to the persons concerned

We understand that it also falls under Art. 12(2) that data controllers be as specific as possible when informing the data subject of the nature of the additional information required to allow identification, e.g. to take into account the situation that most people are not data scientists or pseudo-identifier enthusiasts. These necessary explanations should come at that moment, and not only afterwards along with the files, for those data subjects who have been able to pronounce the right key formulas because they better understand the implicit structures of the internal partitions between the data.

Another subtlety is that legally, data are either personal or not, but that SAR requests are also a matter of strengthening the link between an individual and the personal data. Controllers sometimes use the argument that they cannot be positive one way or the other, or that there is a non-zero possibility that there is data from other people in eligible files. When access refusals are based on this probabilistic view, it is unclear how to evaluate the level of reasonable doubt, or that the controller has sufficiently demonstrated its incapacity (§61).

² Controllers with limited resources could, for example, benefit from open source tools such as the Open Terms Archive, which allow them to quickly visualise successive changes. <https://opentermsarchive.org>

³ By the accumulation of new data (including through the exercise of the data portability right), jurisprudence, new techniques and technologies, or the increased availability of public data (open data, or security breach).

⁴ <https://medium.com/mydata/comments-on-data-portability-guidelines-2102d447f73b#.i5ajdrqgb>



We question this particularly with regards to the variety of remarkable attributes that social platforms are in a position to observe. For example, when Google maps or a dating app notifies us that our profile or review has been viewed 25,000 times, perhaps they are able to tell if this is a unique performance that can distinguish us in a given set. Perhaps sharing these parameters more systematically than through push notifications could help us to consider what additional information can be provided at the time of an access request (cf. §98; §123). Note the irony of the situation if the best way to allow Facebook to link to us our data would be to backtrack through the ad attribution tools, or attempt to target ourselves through their advertising tools⁵.

"Privacy-enhancing technologies"

In line with the above, we question the articulation of the right of access in the light of recent and future implementations of so-called privacy enhancing technologies or privacy by design by controllers.

§67 echoes Art.11 Gdpr in that controllers do not need to maintain identification data for the sole purpose of complying with data subjects' rights, also in light of data minimisation principle. But when intermediate identification schemes are used that are unknown to the users, it blocks their ability to obtain their data while allowing the controller to continue to use it for his own benefit⁶. We would ideally understand that it would not be contrary to the principle of minimisation nor disproportionate to ask the controller the option to be transferred intermediate or temporary identifiers before deleting them, so that one can use them as auxiliary information in a subsequent access request.

Similarly, we have serious concerns about restrictions on access rights as regards to processing information that have been placed on the advertising ecosystems (for instance, consider Google's FLOC project, or the Meta/Mozilla partnership relating to ad attribution⁷). It is thus welcome that social networks logs are eligible to the right of access (§ 138), especially if it includes consent logs, to be able to trace exactly what one has consented to (cf. Recital 42, and Art 5 & 19 gdpr). Concerning the processing and customization happening on devices, we hope that it is planned to ask them to provide data subjects with the means to access this information as well under the right of access (cf. § 98). We hope that this is at least taken into account in the parallel work on standardisation and on the Data Act for data exportability.

Controllers could properly explain at the time of enrollment that when one wishes to exercise their rights they will be asked for much more information when exercising their right of access (e.g. linking their online pseudonymous identity with their legal identity), thus helping individuals in anticipating unintended interactions between their different online/offline identities. In that line, it would also be interesting to know more about how to articulate the providing of additional

⁵ See *Unique on Facebook: Formulation and Evidence of (Nano)targeting Individual Users with non-PII Data* <https://arxiv.org/abs/2110.06636>

⁶ Consider scenarios where identifying attributes are split into two still-linked subsets, which might prove harmful and personal data only in their aggregate.

⁷ <https://netzpolitik.org/2022/future-of-online-advertising-adtechs-new-clothes-might-redefine-privacy-more-than-they-reform-profiling-cookies-meta-mozilla-apple-google/>



information (Art 11(2)) with possibilities for providing the confirmation of one's identity by private or public⁸ third parties, taking into account later legislations in development.

3. Understanding the data

We welcome the very useful clarifications provided by the guidelines regarding the agnostic nature of the right of access. In particular the essential requirement to facilitate the understanding of various processings and their consequences as distinct from the purpose of verifying the regularity of the processing. However, it is not a trivial matter to assign the charge of explaining the content of the files and processings to the data controllers themselves.

Adapting to one's audience

Of course it is useful that controllers are asked to take the size and magnitude of their processings as a criterion for making more effort to explain them, and not the other way around (§127). Controllers must of course also take into account the specific needs brought to their attention to facilitate the right of access. On the other hand, if they have not been asked for anything, one can express reservations with the idea of providing different information depending on what they themselves analyse about people. This is true regardless of whether this is about :

- Their intended audience⁹. Note that in this case access requests are likely to come from people who are not users or consumers of their services¹⁰.
- The characteristics of the individual making the request, including specific situations that the data controller would take into account without being notified by the person concerned (§111, §126).
- The average literacy in the general population. To illustrate, in an access request made by one of the members of the PersonalData.IO association, a data controller used this very idea to limit the information provided¹¹. We understand from the guidelines that this should no longer be possible and this is welcomed.

Of course, there are reassuring clarifications that this should not be used to restrict the scope of accessible data (§127, §136 notamment), but it may be a slippery slope to refer to the logic of panel tests, etc.¹² The usefulness of these concepts for the design of the information provided to

⁸ For instance France Connect or similar national provider of authentication service.

⁹ See § 139 on the layered approach and intelligibility as to be understood by *their intended audience*. And also footnote 69, referring to §9 and §16 of the transparency guidelines : the *likely level of understanding of its audience*.

¹⁰ *ibid* §9 : when it says a controller knows the people about whom it collects information and can use that knowledge to determine what that audience would be likely to understand.

¹¹ The employee of the data controller said: *It is of no use to provide the raw data points collected via the driver app, as that would only be an unintelligible sheet of data for a regular person.*
<https://github.com/pdehaye/BigOther/blob/master/uber/letters/letters.txt>

¹² §9 of the transparency guidelines referred to in footnote 69



all visitors to the privacy policy pages can be seen (cf. Art. 13 & 14). However, we do not think that this should apply equally to Art. 15, as stated in the footnotes to §139¹³. To be even more precise, it should perhaps be avoided to think of this subject in relation to concepts geared at consumers¹⁴, in particular in view of the use in some jurisprudence of the concept of the averagely well-informed consumer.

Being too specific about the references should not limit the other arguments that can be brought to the attention of the authorities. Indeed, different methods can be taken to assess people's understanding, in addition to panels or else. Additionally research on cognition is an evolving field, just as research on what explanations should be in a world of machine learning. Secondly, we are not sure that it is automatically good to derive standards for exercising a fundamental right from those geared at consumers' or constituents. This is true all the more so in the case of data processing by social platforms, the consequences of which often manifest at the interpersonal relational level¹⁵.

In any case, it is already useful that the guidelines specify (§129) that this reasoning must be documented. Perhaps a compilation of various approaches by EDPB or other bodies would be useful to highlight practices that have shown their success or limitations. From our perspective, it should be possible to choose to involve third parties in these processes to provide explanations in addition to those from the data controller.

Choosing your mediators

We note in this respect, that information provided to the data subject *always must be in a human readable format* (footnote 71). It can however also be argued that someone who is not an expert may still want to get all the raw information and then choose the people they think relevant to help them make sense of it in a second step.

Hence it should be possible to choose all log files and/or the layered approach. For example, it would be interesting if it were part of the framework applicable to data intermediaries that some of them could be certified in order to help make this raw transparency intelligible.

In any case, since not everyone has the same questions in the same order, it seems crucial that individuals are presented with all informational options available, including the more technical ones, insofar as this could help them to return to it at another time to further their investigation. This should be doable without having to re-contact the controller each time, taking into account possible artificial limitations on the frequency with which the controller can be contacted for this.

¹³ *Intelligibility is closely linked to the requirement to use a plain and clear language (WP29 Guidelines on transparency - endorsed by the EDPB, para. 9). What is said about a plain and clear language in para. 12-16 with regards to information referred to in Articles 13 and 14, equally applies to communication under Article 15.*

¹⁴ §12 of the Transparency Guidelines, which refers to Article 5 of the Unfair Contract Terms Directive.

¹⁵ The same kind of remark can be made about information in a "commonly used electronic form" that should be based upon reasonable expectations of data subjects (§146). Of course, it is not ideal either in the other sense to be limited to "what format the controller uses in its daily operations", but sometimes actors have specific uses in mind which may be very common in substance as regards semantics and syntax with the data controller's e.g. in research.

On another front, the lately published draft Data Act states *The user shall not use the data obtained pursuant to a request referred to in paragraph 1 to develop a product that competes with the product from which the data originate* (Art. 4(4)). Let us highlight the problem that with some actors, the processing systems concerned are the instruments for operating in potentially an infinite number of relevant markets. This should not be tantamount to prohibiting competition in all their activities on the basis of these data.

From our particular point of view, we are thinking of actors who deploy services to users, to make sense of their data - processing and consequences - , to help them build projects with it. This includes research and innovation efforts to improve data mediation which are potentially in competition if one considers all the informational activities of intermediation platforms. In any case, it would be particularly damaging if unfair competition could be argued when associations or other non profit actors act in the framework of calls for tender or subsidies to provide these services in partnership with public actors.

4. Involvement of data intermediation/sharing actors

In §11, the Guidelines state that the framework for the right of access is likely to change significantly in the future. In this sense, we note that there is abundant legislative activity, especially on the development of data intermediaries and consent management processes for exports and data re-uses for various purposes.

Sharing to whom and for what

This is all good, but let us suggest that there may be a limiting conception of these facilities if they are only geared towards *sharing*, thereby reducing civic participation to *giving* ; without perhaps sufficiently developing the resources to make use of the data for oneself. To compensate for this tropism, it seems important to add emphasis on the need for mediation capacities and support for prosocial uses of their data by individuals. This is without requiring them to affiliate with entities that may have their own causes and priorities in carrying out these activities, and as mentioned above, it is important that complements to the sole efforts of data controllers are developed.

Let us use this point to draw attention to several elements currently under discussion at European level, which can be expected to raise questions of articulation with the mode of exercise of the right of access described in the guidelines.

We note that the EDPB and EDPS have issued a joint opinion about the the Data Governance Act in which they express reservations about in relation to data intermediaries and data sharing in the public interest and through bodies pursuing *altruistic* purposes¹⁶. To concur with this, we draw your attention to several downstream consequences for the right of access that should be avoided :

¹⁶ https://edpb.europa.eu/system/files/2021-05/edpb_statementondga_19052021_en_0.pdf

- De facto dependence on bodies to fully exercise one's rights, with no real choice as to which legal entity one can join. For example, with regard to "sharing for altruistic purposes", there is a risk of conceiving of the "vetting" of eligible bodies in a way that reduces altruism to advocacy and representation in negotiating bodies, or to be pushed to include entities representing specific subpopulations e.g. people with or without digital literacy difficulties.
- Differences in treatment as to the scope of information obtained, between individuals collectively requesting their data, and those who exercise it alone. Think in particular of partnerships between platform operators and vetted NGOs or academic institutions, especially in setups subject to confidentiality clauses.
- Widespread limitation of the right of access by opportunistic recourse to non consent legal bases for data reuse. Clarifications would be welcome on the distinctions between the notion of public interest pursued by altruistic activities (vetted organisations); and the notion of public interest (legal basis for processing)¹⁷. This is all the more so as these different legal bases can be combined together, particularly with research (as EDPB and EPDS noted)¹⁸. If the use of codes of conduct develops, it would be interesting to know, from the perspective of the right of access, from whom to obtain the elements showing that it is indeed scientific activity and not pseudo-science for example.
- Ultimately not knowing who to turn to in order to exercise one's requests among the various actors involved in data spaces or other pooling facilities¹⁹. In the case of data spaces and other multi-actor facilities, it would be good to know who is in charge of checking the reality of the public interest, and from whom to obtain confirmation that the processing operations are really compatible with each other, vs. legal basis free riding by third parties. Even more so if we are dealing with national restrictions and the multi-actor cooperation is transnational, or if data is transiting between data spaces.

Data mediation facilities

To compensate for the focus on sharing facilities for the benefit of third parties, we believe it is essential to concurrently develop mediation facilities for data subjects who wish to use their data for themselves. In the case of social platforms, it is often necessary to go back and forth between the representations of society generated for public action and the concepts deployed by private controllers in order to situate these questions.

Public data services are ideally positioned to develop these bridges between the right of access and the resources developed for access to administrative documents and open data. Depending

¹⁷ We have the same questions as to the respective meaning of *legitimate interest* in the sense of the GDPR and "*legitimate policy objectives*" referred to several times in the Parliament's resolution on the Data Governance Act https://www.europarl.europa.eu/doceo/document/A-9-2021-0248_EN.html

¹⁸ EDPS: preliminary opinion on data protection and scientific research, 6 January 2020, p2

¹⁹ We note that the Commission's work programme on standardisation includes the development of *smart contracts for data sharing*. And that the same programme also refers to the work of the Trade and Technology Council (a forum on trade relations between the EU and the US). This raises questions about the articulation of these contracts with the right of access in the case of data transfers to other jurisdictions. <https://ec.europa.eu/docsroom/documents/48601>

on the questions one may have about personalised processing, one can learn about how a company describes the data subject in the light of the categories by which public actors describe it, and gather around that.

This is unsolicited advice, as the guidelines indicate that this topic is not addressed. However, we identify a relevant link to the guidelines in that :

- Official statistics bodies continuously collaborate with other administrations and civil organizations. Some of them have developed statistical matching services to facilitate cooperation between third parties²⁰. Many administrations have data-driven collaborations with private bodies, including sometimes on the basis of their data obtained via their data subjects rights, for instance on-demand drivers²¹.
- Often the departments in charge of the internal contact point for Freedom of Information responses also take on a DPO role. This is particularly interesting when they are doing so in contact with different ministerial statistical services.
- Open data curation and dissemination services chronically ask themselves the question of "discoverability". We can say that one of the fruitful curation policies could be to start from the questions that citizens ask themselves as data subjects.

The public actors who are able to do so could therefore take advantage of this state of affairs to support the steps taken by laypeople who have shown an interest in data processing by contacting them via access requests. For instance by orienting people to existing resources to understand more about the treatments that are done by the private sector in the light of the expert resources of official statistics. E.g. How are, I or my group, perceived in data by public instances compared to private controllers? Or participate in action research projects, citizen science, or even help set up their own projects using public matching services.

It may be particularly appropriate for private controllers to support these channels (see §106, "education")²², and perhaps even more so on the part of public controllers, as compensation for the limitations of their right of access to processing flows carried out in the public interest. It can be argued that the bodies whose mission it is to gather, produce and disseminate statistical knowledge are particularly well placed to shed light on the challenges of artificial intelligence, for instance to help explain in a meaningful way the objective functions assigned to trained models referred to in the European AI Regulation.

5. Exercising the right of access for pro-social purposes

Where the data controller restricts the scope of access in the name of protecting the personal data of third parties, it should be possible to provide additional information as to the consent of these persons. In this respect, we did not find details concerning grouped requests in the guidelines.

²⁰ French Center for remote secured access to data, "CASD"
https://www.cnis.fr/wp-content/uploads/2022/01/Programe-journ%C3%A9e-appariement_28-janvier-2022_versi-on-du-18012022.pdf ; <https://idan.network> European Network for research

²¹ See <https://www.peren.gouv.fr/projets/>

²² This is all the more so as bridges are already being built between regulators, researchers and platforms with the Digital Services Act (*vettred researchers*).



For example, we have been confronted with the case of a platform that uses the argument of passenger data protection to unilaterally restrict the scope of data retrievable by drivers²³, without consideration for technical means to achieve respective goals. This should be different if the wish of the passengers concerned is instead to share them with a particular driver²⁴. Data controllers should take into account authorisations provided by passengers to give access to certain parts of the data concerning them. (cf. §103). It could even be argued that in such situations the platform is forced to ask the opinion of the user on the other side of the data transfer.

A similar logic leads us to cultivate relationships with and between actors present on the different sides of intermediation, e.g. customers and workers of on demand services, or to cross-reference feedback between residents of different member states on the differences in data obtained pursuing requests and the corresponding national limitations (e.g. article 23) or via data transfers vehicles.

We more broadly can see data collectives as a field of opportunity to introduce playability of social distances by the individuals concerned. An illustration is the instrument created by MIT²⁵ to act on the filter bubbles linked. We expect opportunities to go beyond content curation, e.g. data pooling as an effective approach for grouping data subjects who don't "pool" very often in real life. As described by the NGO "IT for Change": *communities are generated through specific acts of data processing and individuals may not know the potentially innumerable communities they are part of; or even be in a position to identify other individuals in these communities*²⁶.

People could thus gather using the findings about polarities that manifest in (or because) of the treatments performed by social platforms as an opportunity to work in the opposite direction with their data. This could presumably assist in thinking about desirable heterogeneities in the composition of communities with "altruistic" goals.

Guidelines for individuals

Social platforms tend to take advantage of the household exemption²⁷ to have individuals be vectors for collecting data on their relatives and other contacts²⁸. This favorable regime does not detract from the fact that these processes place an ethical responsibility on people to impact on others²⁹, thereby potentially damaging their relationships³⁰. Moreover, it is our understanding that

²³ These refusals are quite remarkable when at the same time links are established between platforms that forgot to ask our opinion to reuse information about us from context to context.

²⁴ See for example the dynamics of cooperation between a developer - videographer - Uber client and a union leader of VTC drivers. First <https://www.youtube.com/watch?v=mqU1QWmvjho> ; then <https://twitter.com/Sqli69/status/1450057484861906951> .

²⁵ <https://ilp.mit.edu/node/43998> A browser extension that enables Twitter users to replace their own feed with that of another real Twitter user.

²⁶ (p10) Anita Gurumurthy and Nandini Chami - Governing the Resource of Data: To What End and for Whom? January 2022 <https://datagovernance.org/files/research/1641877030.pdf>

²⁷ Article 2§2 c) ; Recital 18

²⁸ Think for example of contact lists processed by Facebook or more recently Clubhouse.

²⁹ It also raises the bar for difficulty in domestic surveillance situations.

³⁰ The case of the suggested contacts made by Facebook between a lawyer and a psychologist are illustrative of these issues.



self-employed professionals (freelancers, etc.) do not benefit from the household exemption when acting in a professional capacity, but it just so happens that a recurring trend in the social platform economy is to facilitate links between various social spheres.

This is why we believe that there is a growing need for guidance to accompany natural persons processing data. It is part of a context that goes beyond the sole framework of the right of access but still has a significant impact on it. It is particularly apparent at §104, concerning data subjects who become data controllers of third-party data as a result of access requests³¹. From our own experience, we think mainly of groups that use their own files to inform their collaborations on issues of common interest, like local transport policies. We can also think of students wanting to understand the traces generated by the use of digital education services. Here, examples and flowcharts geared at professionals often provided by the EDPB could be declined from the perspective of individuals acting in a personal capacity. In this particular case, for the follow-up of their access requests, and further processing involving third parties under various relevant regimes (household, independent research, arts, etc.).

6. Access rights and freedom of expression

Individuals engaging in the process of reverse engineering data flows often want to **spread their comments about the data processing they are subject to, including by posting glimpses of them in a public space**. Benefiting from exchanges with others is indeed decisive for the understanding of treatments by those who do not do this for a living³². In doing so they could benefit from guidelines for balancing their freedom of expression with their responsibility to protect the information of others, and appropriate mitigation measures. Particularly if we consider groups of individuals who act as joint-controllers of each other's data, with regard to the information they disseminate among themselves or to third parties.

Apart from the case of data obtained via SAR, a very concrete example of this is the development of collaborative open source investigation activities by social network users ("OSINT"). Here, risks of unintentionally revealing personal information about others have been realised several times on the occasion of events such as the attacks on the US Capitol³³. But at the same time, regulators

<https://gizmodo.com/people-you-may-know-a-controversial-facebook-features-1827981959> We can add the recreational use of genetic analysis services. "23andMe sold all of the genetic information it had, without users' consent, to the British pharmaceutical company GlaxoSmithKline".

<https://www.usinenouvelle.com/article/23andme-vend-l-integralite-des-donnees-genetiques-de-ses-clients-au-laboratoire-gsk-et-cree-la-polemique.N729654>

³¹ §104. *If the data subject uses the obtained record which includes personal data of the interlocutor for other purposes by, for instance, publishing the record, **the data subject will become a controller** for this processing of personal data relating to the other person whose voice was recorded.*

³²Even (especially) for specialists, public social spaces are also an important vector of information since the field is constantly evolving.

³³ Cases of misidentification and erroneous accusations were also reported in this event

<https://www.vice.com/en/article/4ad5k3/how-normal-people-deployed-facial-recognition-on-capitol-hill-protests>



need the active contribution of civil society to contribute to the effectiveness of regulations, in this case data protection³⁴.

Finally, questions arise concerning the rights of third parties other than privacy. For example, some companies are trying to install restrictions on the dissemination of information obtained through dispute resolution forums³⁵. In this case, the other persons concerned are not at an aligned level of information. This may need to be clarified by the authorities, for example to support the possibility of publishing the results of its exchanges as many already do on github or social networks.

It would therefore be judicious to acknowledge the evolution of these practices, providing guidance to individuals to encourage them to follow through in a responsible manner and to help each other. We believe that taking these considerations into account would help to disseminate literacy more widely, to the benefit of people's agency in line with European objectives.

³⁴ Social platforms themselves rely on a partial delegation of control to users, and this mechanism is about to be upheld in the Digital Services Act (e.g. trusted flaggers).

³⁵ Example of Coursera, under Safe Harbour, a regime previously in force for data transfers to the US <https://pdehaye.medium.com/the-fine-print-on-the-coursera-partner-conference-2016-4832fc83cd6c> There are numerous new instances of this since.

