PANELFIT

PARTICIPATORY APPROACHES TO A NEW ETHICAL AND LEGAL FRAMEWORK FOR ICT

Guidelines on Data Protection Ethical and Legal Issues in ICT Research and Innovation.

SOCIAL NETWORKS

Brief summary

Iñigo de Miguel Beriain (UPV/EHU)

This document is aimed at helping ICT researchers or **innovators using personal data obtained from social networks.** It is worth mentioning that we will not address here the use of social networks to collect data (such as, for example, by using Google surveys to get data back on a specified set of questions from real people). This is due to a simple reason: in these cases, the data itself does not come from a social network but through a social network. Indeed, social networks only act as a tool to gather those data. Therefore, these data are not so different to any other data collected by a more traditional way (such as a survey in paper) and, thus, they do not deserve special attention here.

If ICT developers consulting these Guidelines are planning to use AI tools to process data obtained from these networks, they should consult the part of the Guidelines devoted to Artificial Intelligence (AI). If they are planning to use them for purposes related to biometrics, Internet of Things or Geospatial location, they should consult the parts of these Guidelines that are devoted to those issues. In order to avoid unnecessary repetitions, we are leaving such issues out of this analysis.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.

DISCLAIMER

This document is an abbreviated version of the part of the Panelfit Guidelines related to Social Networks. For better information, it is advisable to consult the full version of our Guidelines, written by José Antonio Castillo Parrilla and Iñigo de Miguel Beriain (UPV/EHU)

Furthermore, one must always keep in mind that the information provided in our Guidelines does not constitute legal advice; instead, all information, content, and materials provided are for general informational purposes only.

The Guidelines provide general advice around EU data protection law under the GDPR. Accordingly, the reader should be aware that the situation relevant in their specific processing context, as well as in their specific jurisdiction, may deviate from the guidance provided. Indeed, information in our Guidelines may not constitute the most up-to-date legal or other information. The legal situation in relation to data processing in the EU changes regularly. New laws and new interpretations of existing laws relevant to the topics covered by the Guidelines appear frequently and changes may not be reflected in the Guidelines.

In this regard, we would highlight, without any intention of being comprehensive, at the time of writing, the significance of the following draft EU laws to the topics covered in these Guidelines: the ePrivacy Regulation, the AI Act, the Data Governance Act, the Digital Services Act, the Digital Markets Act and the Data Act.

Where relevant at the time of writing, authors may have attempted to highlight provisions of draft laws in relation to the topics covered in these Guidelines. The reader should be aware that drafts may change and that such references may not remain valid over time. Equally, authors' choices to consider certain provisions from certain draft laws should not be taken as indicative of effort to be comprehensive in addressing all relevant provisions from all draft laws.

Readers of the Guidelines should contact their DPOs and DPAs to obtain advice with respect to any particular legal matter. No reader, user, or browser of the Guidelines should act or refrain from acting on the basis of information provided without first seeking legal advice from counsel in the relevant jurisdiction. Only DPOs and DPAs can provide assurances that the information contained herein – and your interpretation of it – is applicable or appropriate to your particular situation. Use of, and access to, the Guidelines do not create any relationship between the reader, user, or browser and the authors, reviewers, validators, or commentors, of the Guidelines.

The views expressed in, or through, our Guidelines are those of the individual authors writing in their individual capacities only – not those of EU Commission, of course. All reference to reviews, validations, or provision of comments or suggestions, refer to the personal opinions of individuals acting in their personal capacities – and do not refer to the opinions of the organisations these individuals represent or to acts of these individuals in their official capacities.

The Guidelines contain links to other third-party websites. Such links are only for the convenience of the reader, user or browser; the authors and the reviewers/validators do not recommend or endorse the contents of the third-party sites.

All liability with respect to actions taken or not taken based on the contents of the PANELFIT Guidelines are hereby expressly disclaimed.

1 Introduction to social networks and data protection issues

Some preliminary advice: **the fact that much of the data contained in a social network is easily apprehensible does not legitimize its processing.** This is a crucial aspect when it comes to the processing of data obtained from social networks: ICT researchers and innovators must carefully ensure that they have a legal basis that allows them access and storage of these data. Once they have already accessed them, they will have to make sure that the same and/or other legal bases allow them further processing of those data. In general, this means that they must have a profound knowledge of the Developer Policies imposed by the social networks

Furthermore, transparency implies that intended research subjects should be informed at some point about the research being performed, what sort of personal data controllers are collecting and how it will be used. Some services make it clear that this must be done before one starts harvesting. In the absence of a specific policy and where researchers/innovators are conducting observational research which the need to obtain consent up front could damage, they should let the individuals concerned know as soon as possible. The ICT researchers/innovators should always remove from their harvesting individuals who do not consent to being included.

2 Preliminary steps: the crucial issues to be considered

Some essential tips regarding these preliminary stages include:

- Ensure that the use of data gathered through social networks does not promote scenarios that are incompatible with the EU fundamental values
- Conduct a proportionality assessment for the use of personal data gathered through social networks
- Ensure that the team members processing personal data have been\are adequately trained on the Developer Policy corresponding to the social network from which data will be extracted, and the key concepts on data protection issues
- Adequate assessment tools on data protection have been implemented from the very beginning of the project
- The roles played by all different agents involved in the data processing are clear through the corresponding agreements and the controller can provide evidence on this.
- Ensure that you have gathered the relevant information on the terms of use of the social network from where you gather the data and abides by them
- Consult the representatives of the key collectives involved in the data processing on the impact of the use of the gathered data and the concrete social network selected and incorporate their inputs to the system

2.1 Make sure that your project is compatible with the fundamental values of the EU

Before considering the use of data gathered from social networks for the project, the developer should have her/his primary objective clear in mind. **A clear idea of the concrete use of data gathered through social networks will help controllers determine in the early stages of development some important legal issues regarding processing**, such as compliance with the Developer Policy of the social network, possible need of international transfers of data, existence of joint-controllers or processors -which need to be carefully selected-, or the security and organizational measures to minimize risks.

2.2 Implement a training programme in ethical and legal issues for ICT developers and other relevant stakeholders

Implementing **basic training programmes** for researchers/innovators. If they are gathering data from a concrete social network, this training should include a careful analysis of its particular Developer Policy. An early involvement of DPOs from the participating institutions is highly advised.

2.3 Define the roles played by all agents involved in the processing

The concepts of controller, joint controller and processor play a crucial role in the application of the GDPR. In the case of utilization of social networks for data processing, it is equally important to properly distinguish the data controller from the data processor, since the responsibilities of each are different. **To dispel any possible doubt,** we must first turn to the list of definitions in the GDPR, interpreted in accordance with the EDPB Guidelines 7/2020 on the concepts of controller and processor in the GDPR and the EDPB Guidelines 8/2020 on the targeting of social media users [1] and the relevant case law of the CJEU[2].

In the most common scenarios, ICT researchers and innovators will play the role of a third party regarding social networks and data subjects. The network will provide them with data that belong to the data subjects. Once these data are already under the control of the researchers/innovators, they become controllers of those data and take the corresponding responsibilities.

2.4 Prepare the contracts with the social network and (in case) with the joint controllers, processors, etc. and document the

Gathering data from social networks often involves entering into some kind of agreement with their representatives. **Written agreements between all agents involved in the development of the tools should be reached and documented, whenever possible (see art. 28 of the GDPR).** These should include clear specifications about the responsibilities taken by all participants. Promoting a

[1] EDPB Guidelines (Guidelines 8/2020 on the targeting of social media users Version 2.0 Adopted on 13 April 2021, at: https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_082020_on_the_targeting_of_social_media_users_en.pdf, p. 11).

[2] The judgments in Wirtschaftsakademie (C-210/16), Jehovah's Witnesses (C-25/17) and Fashion ID (C-40/17) are particularly relevant here.

continuous interaction between all DPOs involved might be an excellent option. Ad-hoc supervisory bodies and tools can be adopted to ensure a smooth oversight of the participants' processing.

2.5 Promote end-users engagement

Since ICT involves the use of personal data from different types of data subjects, it is highly recommendable to hear the voices of the representatives of the collectives involved so as to ensure that the Data Protection by Design policies are in line with their interest, rights and freedoms. Organizing some **preliminary discussions** with those representatives ensures the implementation of a bottom-up framework that could be very helpful to this purpose.

3 Gaining access to data. Some essential tips

In the case of processing data from social networks, it is essential to underline that **ICT researchers or innovators must be aware that they will certainly need different legal bases for data processing at the moment of accessing the data and at the moment they performs their research or innovation based on those data.**

In the first case, what is needed is a legal basis to obtain the data from the social network. In the second case, it is a matter of finding a basis that allows the data, already legitimately acquired, to be used for research purposes. **It is essential to note that the mere fact that the data subjects have published their data in online public spaces does not allow for their processing.** These are still personal data, even if the data is publicly available. The publication might serve to avoid the ban included in Article 9.1 of the GDPR, if we are talking about data of special categories, but does not serve as a legal basis for processing. As such, companies may not freely re-use the data, and may not further process it without the individuals' knowledge and without an adequate basis for lawful processing.

These are some essential tips provided by the Ethics information for Linguistics and English Language[3] that you must follow if you are planning to gain access to data from a social network:

- If the data are in the public domain, you must abide by any requirements stated by the corpus provider, including with respect to anonymity, or any other conditions on use.
- Some corpora may require ethical approval, especially corpora that include physical or mental health data, or corpora that contain data that could be used to de-anonymise individuals (e.g. when free-text responses are allowed).
- If the data are not in the public domain, you must ensure that your use of the data conforms to any requirements stated by the corpus provider. For example, the data must not be shared in any unauthorized manner (e.g., posted online).
- In either case, if there is reason to suspect that the people who initially provided the data were not aware that it would be used for research purposes, you should

[3] https://resource.ppls.ed.ac.uk/lelethics/index.php/frequently-asked-questions/corpus-research/

carefully consider the ethical implications of your research, including whether you should obtain informed consent.

All these tips can be concreted in the following steps:

- First, always keep in mind the **reasonable expectations of the data subjects about the use of their data (Recital 47, GDPR).** This is essential in most social networks Developers Policies. For instance, Twitter Developer Policy states that "we prohibit the use of Twitter data in any way that would be inconsistent with people's reasonable expectations of privacy. By building on the Twitter API or accessing Twitter Content, you have a special role to play in safeguarding this commitment, most importantly by respecting people's privacy and providing them with transparency and control over how their data is used."[4]

- Second, obtaining approval to access the APIs and the Contents of a social network is never enough to ensure a lawful data processing. It is just the first step. Most of the social networks have developed detailed **Platform Usage Guidelines that researchers must strictly follow to ensure policy compliance for their planned use of the platforms and compliance with data protection ethical and legal requirements.**

- Third, most of the social networks have developed tools that **provide support to researchers** willing to use their Application Programming Interface (APIs). It is always recommendable that researchers use these services in case of doubt on data processing.

- Fourth, however, researchers and innovators should never forget that, as a controller, you are responsible for ensuring that the data protection framework is adequately respected. Thus, you should check whether the statements on the legitimacy of the data processing carried out by the social networks correspond to reality. Reviewing their data collection policies to check the soundness of the consents granted from a GDPR perspective seems a necessary or, at least, prudent requirement.

- Fifth, researchers/innovators shall keep in mind that social networks **might change their policies** from time to time without notice. Since they usually introduce this caution in their own policies, researchers take responsibility for keeping themselves informed about these possible changes. Thus, periodic reviews of such policies are highly recommended.

- Sixth, since researchers will be processing data that have not been obtained from the data subject, they shall provide the data subject with the information requested by article 14 unless any of the circumstances quoted in its point 5 apply.

- Finally, in case of doubt, always consult your Data Protection Officer and, if necessary, the corresponding Data Protection Authority.

Researchers and innovators who use data obtained from social networks are responsible for complying with all policies settled by those networks. Thus, it is essential that they review and understand these policies before they access the social networks' APIs and contents. The time spent reviewing their policies may

[4] https://developer.twitter.com/en/developer-terms/policy

save researchers hours of further work down the road and may even help them avoid legal responsibilities.

4 Choosing a legal basis for further processing

Once researchers/innovators become the controllers of the data gathered from social networks, they have to decide on the legal basis that will legitimate further processing of those data as soon as possible.

These are some criteria that should be kept in mind for this purpose:

- The necessity or usefulness of the use of the data obtained from the social networks for the achievement of the purpose or interest of the processing must be sufficiently justified in the lens of the legal basis selected.
- The data controller must carefully weigh up (1) the basis of entitlement used, against (2) the possible risks arising from the data processing.
- In addition, the controller should consider all adequate safeguards so as to ensure that the interests, rights and freedoms of the data subject are adequately preserved. This balancing must be particularly careful if the data subject's consent acts as the legal basis for processing.

The following tables provide brief overviews of the various alternative bases of legitimation under Article 6 and their relation to the processing of data from social media Consent is the most traditional legal basis for data processing in the context of social networks. However, where a controller seeks to process personal data for research purposes, public interest might be an excellent option. Unfortunately it requires that certain conditions apply. Legitimate interest, on the other hand, is an alternative suitable legal basis for processing in this context, but one cannot assume it will always be appropriate[5]. It is likely to be most appropriate where controllers use people's data in ways they would reasonably expect and which have the least possible relevant impact on data protection or privacy issues, or where there is a compelling justification for the processing.[6]

Possible Legal bases (Art. 6 GDPR)

Legal bases for processing	Use in the context of social networks
6.1.a –consent	Probably, the most popular legal basis for

[5] Ad ex., Public authorities can only rely on legitimate interests if they are processing for a legitimate reason other than performing their tasks as a public authority, so "public task" is a better legal basis in these situations (ICO: Legitimate interests, at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/).

[6] ICO: Legitimate interests, at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/

	data processing, although its widespread use is increasingly being questioned[7] (see following section)
6.1.e - processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller	It may be applicable, but the following cautions should be observed: - The public interest purpose must be clearly identified as well as the connection to the research, - Reasons must be given as to why the use of data from social media is necessary or highly desirable for the objectives pursued. -The basis for the processing has been laid down by Union law; or a Member State law to which the controller is subject.
6.1.f - processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child	It may be applicable, and indeed is the best alternative to consent as a basis for legitimacy. The following cautions should be observed: - the data controller must carry out and give reasons for an appropriate balancing of (1) the legitimate interest pursued and (2) the impact on the fundamental rights and freedoms of the data subject; this balancing must be carried out with particular care if data from minors are involved

Special consideration to re-use of data

End users of social networks are often unaware of the fact that their data are used for purposes other than those that they pursue when they provide those data. However, most social networks ensure that the data subjects provide consent to this further processing and their Developer Policies will surely cover this issue.

Researchers and developers willing to process the data obtained from social networks for research purposes might obtain a new consent from the data subjects. This, of course, is hard and not always necessary. They could rely on the original consent provided by the data subject to the social network. However, **the researchers/innovator should, however, ensure that the processing they are willing to perform is allowed by the consent originally provided by the data subject or find an alternative legal basis (by asking for a new consent or using legitimate interest or public interest as an alternative, for instance).** Consulting the terms of use of the

[7] See, on data processing for health purposes in the American privacy system, Charlotte A. Tschider, ´The consent myth: improving choice for patients of the future´ (2019) 96 Washington University Law Review 1506.

social network and the consent gathered originally is an excellent way to check if the secondary use of data could be considered compatible with the purposes for with data were originally collected.

If the research involves using data gathered from different social networks, researchers should focus on designing intra-provider and eventually inter-provider privacy risk evaluation mechanisms that take into account personal data revealed for all data processing activities for a concrete social network and for all OSNs that a data subject uses, respectively.

Last, but not least, since researchers will be processing data that have not been obtained from the data subject, they shall provide the data subject with the information requested by article 14 unless any of the circumstances quoted in its point 5 apply

5 Fairness and Transparency issues

Fairness is an essential principle in the GDPR. In the case of data gathered through the use of social networks, it is particularly important to avoid biases related to gender, race, age, sexual orientation, national origin, religion, health and disability, etc. This might be problematic since it is possible that some of the data gathered via social networks do not correspond to real users, or their sensitive data are not at all accurate. This might create hidden biases. In order to avoid biases, **critical assessment of the provenance of the data is required**. To this purpose, organisational measures should be implemented to guarantee the accuracy and reliability of the gathered data, while still ultimately deferring to the right of users to withhold private information (e.g. confirming whether or not a record is accurate). Furthermore, performing an audit devoted to detecting biases in raw data or in the inferred or derived datasets is required especially when controllers are using datasets produced via social networks.

The main focus of **transparency** is to inform **data subjects** up-front of the existence of the processing and its main characteristics. In the case of using data from social networks, **it is necessary to point out that, in general, Article 14 of the GDPR will be applicable at some point. Thus,** data subjects should be made fully aware that their data is being shared with third parties This could be done in different ways. For instance, the CNIL advised that data controllers could either include all third-parties in an exhaustive privacy notice, but periodically updated, or insert a link in this notice and redirect individuals to the list with the third-parties and their own privacy policies. **I**mplementing the so-called Transparency Enhancing Tools (TETs)[8] might be an excellent option to guarantee that the Transparency principle rules, especially when massive or automated data processing is expected.

It is necessary, however, to mention that sometimes it might be extremely difficult for a controller who has gathered the data from a social network to inform data subjects about the processing. If this is the case, he/she might recall article 14.5 (b), Thus, in principle controllers could avoid providing information about the processing to the data subjects

[8] TETs can be subdivided into 'ex ante' and 'ex post'- TETs. Ex ante-TETs guide the user's decision making process before she makes her choice pertaining to disclosing any personal data to a data controller. Conversely, ex post-TETs visualise disclosed personal data in such a way as to make transparent the processes that have taken place since the user has disclosed her data (see P. Murmann; S. Fischer-Hübner, `Usable Transparency Enhancing Tools – A Literature Review´ (2017), working paper. At: http://www.diva-portal.org/smash/get/diva2:1119515/FULLTEXT02.pdf).

if this is rendered impossible, but only if they *take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available.*

Some essential tips related to fairness are:

- Perform internal/external audits aimed at detecting biases in the datasets built and/or the conclusions of the analysis
- Perform audits aimed at detecting biases in the datasets built and/or the conclusions of the analysis

Some essential tips related to transparency are:

- Provides the data subjecs with complete information about the processing and their rights
- Ensure that the information is provided concisely, transparently, intelligibly, and in an easily accessible way. It is clear and redacted in plain language.
- If providing the information is rendered impossible, take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available.
- If a DPIA has been conducted, ensure that it has been published openly and is publicly available
- Since the personal data were not provided by the data subject, provide the data subjects with all the information listed in Article 14.1 GDPR;
- Since the personal data is not provided by the data subject, provide them with the information:
  - within a reasonable period after obtaining the personal data, but at the latest within one month;
  - if the personal data are to be used for communication with the data subject, at the latest at the time of the first communication to that data subject;
  - if a disclosure to someone else is envisaged, at the latest when the personal data are first disclosed
  - Document all the information regarding these issues

6 Data governance: minimization, purpose limitation and storage limitation principles

6.1 Minimization principle

The minimization principle states that personal data shall be **adequate**, **relevant and limited to what is necessary in relation to the purposes** for which they are processed. When researchers/innovators gather data from social networks, they might end up processing far more personal and sensitive data than they really need for the specific purposes of the research. There are some ways through which such a scenario might be avoided. In principle, controllers **should promote the use of anonymised data.** If they do not need personal data, they could ask the social network to provide them with anonymized data. However, researchers/innovators should keep in mind that anonymization might be hard to reach. Thus, **controllers should not presume that their anonymization processes will serve well to preserve data subjects' privacy. Indeed, they should perform DPIAs and risk assessments to ensure such a belief.**

An alternative to anonymization as such is the use of **aggregated data**. **When the purpose of the processing can be achieved using aggregated data, this is recommendable.** Indeed, sometimes a specific research only needs aggregated data and has no need of the raw data collected in the social networks. Therefore, **controllers must delete raw data as soon as they have extracted the data required for their data processing**. As a principle, deletion should take place at the nearest point of data collection of raw data (e.g. on the same device after processing).

6.2 Purpose limitation

Purpose limitation is a key concept when processing data obtained from social networks and most platforms include it in their Developer Policies. Researchers and innovators shall strictly follow such policies. On the other hand, it is often true that data subjects are not truly aware of the permissions they provide the social networks for the processing. This is a particularly important reason why controllers using those data should not process the data for purposes that could be considered as incompatible with the initial consent.

6.3 Storage limitation

The principle of storage limitation obliges data controllers not to store personal data for 'longer than is necessary for the purposes for which the personal data are processed' and to introduce pseudonymization and anonymization measures that reduce/eliminate the identifiability of data subjects as soon as possible for such purposes. The problem here is that usually social networks might use the stored data for different purposes

In order to avoid unlawful storage, a necessity test must be carried out by each and every stakeholder in the provision of a specific service in the social network, as the purposes of their respective processing can in fact be different. For instance, personal data communicated by a user when he subscribes to a specific service in the social network should be deleted as soon as the user puts an end to his subscription. Similarly, information deleted from his/her account by the user should not be retained. When a user does not use the social network for a defined period of time, the user profile should be set as inactive.

Some essential tips regarding these principles are:

- Assess what data is necessary and proportionate, and anonymised or pseudonymised any other data.
- Document anonymisation and pseudonymisation methods
- Ensure that if special categories of data are used, a necessity analysis has been carried out and document it
- Use the data only for the purposes you collected them, unless a legal basis allows their lawful processing.
- Do not store personal data for 'longer than is necessary for the purposes for which the personal data are processed and made data subjects aware of the lifespan.
- Check and documented the utility of the stored data for the intended purpose of the research.

- Store data in a way that hinders personal data processing as much as possible and document the reasons that made you select such policy.
- Document all the information regarding these issues.

7 Accountability and oversight[9]

Accountability consists of two requirements for controllers:

- **Compliance** with the principles of the GDPR;
- **Demonstration of compliance.**

Compliance is achieved by implementing ***technical and organizational measures*** that are adequate compared to the risks to the rights and freedoms of data subjects, correspond to the state of the art of technology, and are cost-effective. Every description of the principles has provided examples of such technical and organizational measures. For a systematic application of these measures, controllers can create ***data protection policies***. ***Approved codes of conduct***, where available, are similar but are pre-approved and usually address an entire sector. Compliance is not a state that is reached once, but **a continuous process** that spans the whole life cycle of a processing activity.

Demonstration of compliance is predominantly achieved by **documentation**. Documentation should be continuous like the process of compliance. Every implemented measure, including data-protection-relevant considerations and decisions, should be documented. The GDPR requires two formal documents as part of demonstrating compliance towards supervisory authorities: the **register of processing** and, where the risks are likely to be high, a **data protection impact assessment.** Certification can support the demonstration of compliance.

7.1 Data Protection Officer

In most cases, ICT research based on data from social networks involves operations that, by virtue of their nature, their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale. Therefore, the appointment of a DPO is compulsory according to the conditions settled by Article 37(1) apply. Even if this is not the case, **it is always recommendable to proceed to do so, at least in terms of transparency.**

7.2 Data protection Impact Assessment

Performing a DPIA is often compulsory in the case of social networks since it involves a systematic monitoring of a publicly accessible area on a large scale (Article 35(3) of the GDPR). In order to see if a DPIA is necessary:

[9] This part of these Short Guidelines was originally written by Bud Brugger, see: https://guidelines.panelfit.eu/the-gdpr/main-principles/accountability/

- Determine the jurisdictions where data-processing activities will take place.
- Check if those jurisdictions have enacted lists indicating the processing that requires a mandatory DPIA and checked if the intended data processing is covered by those provisions.
- If you are unsure of the necessity of carrying out a DPIA, you must consult with the DPO or, in lieu of, the legal department of the controller.
- If necessary, file a prior consultation with the appropriate supervisory authority.

There is no standard way to perform a DPIA. However, Article 35.7 GDPR calls for specific elements that shall always be present. These are:

- a systematic description of the envisage processing operations;
- the purposes of the processing operations;
- an assessment of the necessity of the processing operations in relation to the purposes;
- an assessment of the proportionality of the processing operations in relation to the purposes;
- an assessment of the risks to the rights and freedoms of data subjects;
- the technical and organizational measures envisaged to address the risks.

7.3 Design your Privacy Policy and prepare the documentation of processing

The Privacy Policy is the public document that explains how a research project processes personal data and how it applies data protection principles, according to articles 12-14 of the GDPR. All data subjects must have access to this Privacy Policy. It should be documented. A non-official, but recommendable template can be found here: https://gdpr.eu/wp-content/uploads/2019/01/Our-Company-Privacy-Policy.pdf

Controllers must always keep in mind that, in the case of data gathered from social networks, they might end up mixing different datasets or create inferred or derived data. The traceability of the processing, the information about possible re-use of data, and the use of data pertaining to different datasets in either the same or different stages of the life cycle must be ensured by the records. Whoever processes personal data (including both controllers and processors) needs to document their activities primarily for the use of qualified/relevant Supervisory Authorities. This must be done through records that are maintained centrally by the organization across all its processing activities, and additional documentation that pertains to an individual data processing activity.

The main decisions made by the data controller "have to be documented in order to comply with the requirement of data protection by design" (of Article 25 of the GDPR). Indeed, an organization who is processing personal data (including both, controllers and processors) needs to documents its activities primarily for consumption by the competent Data Protection Supervisory Authorities (DPA). This includes the ***records of processing*** that is maintained centrally by the organization across all its processing activities and **additional documentation** that pertains to an individual data processing activity.

Records of processing can be kept in written or electronic form[1]. So expect to either fill in an organization-specific form or enter your information into some (data protection) management system.

To provide an initial idea, the minimal content of the records of processing for controllers includes the following items[2]:

- The **name** and **contact** details **of the controller**, the controller's **representative** and the **data protection officer**;
- the **purposes** of the processing;
- a description of the **categories of data subjects** and of the **categories of personal data**;
- the **categories of recipients** to whom the personal data have been or will be disclosed;
- where applicable, **transfers of personal data to a third country** together with the documentation of suitable safeguards;
- where possible, the envisaged **time limits for erasure of the different categories of data**
- where possible, a general description of the **technical and organizational *security* measures**

Keep in mind that:

- Your organization may use a different set of items since on one hand, it already is in possession of some of this information (such as the first bullet), and on the other hand, it may require additional information (such as the contact of the person responsible for the single processing activity at hand).
- It is possible that the legally required record keeping is combined with the management needs of the organization, such as an internal inventory of computing and computing resources.
- Your organization may also use multiple systems, e.g. depending on whether it is acting as a controller or as a processor; or distinguishing between permanent data processing activities (such as communication systems and accounting) and temporary ones (such as those linked to temporary projects or assignments). The creation and maintenance of records across multiple systems is not prohibited under the GDPR.

Some essential tips are:

- Data protection (like security) is a process, not a final state. Continuously document that process rather than only the final characteristics of the processing activity.
- When applying data protection by design[4], the processing activity can be seen as the results of a series of many considerations and decisions. It is these considerations and decisions that should be documented.
- Deciding on a structure and format to systematically collect this information at the point of time when you conceive your processing activity.
- Where the documentation itself contains personal information (see below), make sure to protect is sufficiently and limit its further use to the purpose of demonstrating compliance with the GDPR.

The first stages of the project development are the perfect moment to set up a systematic way of collecting the necessary documentation, since it will be the time when the organization conceives and plans the processing activity[10].

[10] Article 25(1) of the GDPR calls this "the time of the determination of the means for processing".

Last but not least, controllers must keep in mind that ethics committees will probably play a key role in personal data processing. However, this might change considerably between sectors and countries. Controllers shall ask their DPO about this topic.

Finally, controllers shall not forget that there might be ethical implications beyond legal compliance. Consultation with an expert in the ethics of social networks is always recommended.

8 Data subjects rights

Since we have already analysed the right to information and the right not to be Subject to Automated Decision-Making has been extensively addressed in the "Human Agency" section of this document, we will focus now on the remaining rights. Some general tips regarding data subject rights include:

- Introduce the necessary procedures to ensure that the data subject rights are adequately satisfied, no matter if they are the end-users or third parties.
- Introduce the necessary procedures to ensure that the data subject rights are satisfied in time (maximum one month after request, extendable by two additional months with regard to the complexity of the task and the number of requests). If you need this additional time, inform the data subject about.
- Introduce efficient tools to ensure that data subjects are able to exercise their rights in a practical manner, for instance by introducing data interoperability standards.
- Provide the data subjects with remote access to their personal data. Particularly, controllers which provide online services based on personal data have provided an online tool for this purpose.
- Ensure that data subjects have easy access to the procedures to exercise their data rights and the contact details of the DPO or person responsible to handle data requests
- Document all the information regarding these issues.

8.1 Right of access

This right is particularly important in the case of data gathered from social networks since data subjects are usually unaware of the existence of such data. Furthermore, inferred data might be created by the controller and these data might be of particular interest for the data subject.

Some essential tips are: Make sure that

- You are aware that you need to inform individuals of their right to access, in addition to including it in your privacy notice.
- Ensure that you have provided data subjects with clear information on how to exercise their access rights
- Make sure that you are able to recognize a subject access request and they understand when the right of access applies.

- Understand that the right of access is to be applied at each stage of the life cycle of the AI solution, if it uses personal data.
- Have a policy for how to record requests you receive verbally.
- Understand when you can refuse a request and be aware of the information they need to provide to individuals when doing so.
- Understand the nature of the supplementary information you need to provide in response to a subject access request.
- Have processes in place to ensure that you respond to a subject access request without undue delay and within one month of receipt.
- Be aware of the circumstances in which you can extend the time limit to respond to a request.
- Consider that there is a particular emphasis on using clear and plain language, especially if you are disclosing information to a minor. Consider who should be the subject of the information (the child? A representative?)
- Understand what you need to consider if a request includes information about others.
- Understand how to apply the right to access in training stages.

Right to rectification

Data subjects can provide false or inaccurate information due to a lack of understanding of the implications that it might have. Controllers are obliged to communicate the rectified data to each recipient to whom the personal data has been disclosed, unless this proves impossible or involves disproportionate effort. Controllers cannot argue that managing large datasets is too complex to ensure rectification in order to avoid this requirement. Some essential tips are[11]: make sure that

- You are aware that you need to inform individuals of their right to rectification, in addition to including it in your privacy notice.
- You know how to recognize a request for rectification and understand when this right applies
- If you receive a rectification request from a legal entity, please indicate that the request was not lodged by an individual;
- If the data subjects have not identified themselves in an adequate manner, please ask for further information to confirm identity
- You have a policy for how to record requests you receive (including verbally).
- You understand when you can refuse a request, and you are aware of the information you need to provide to individuals when asked to do so.

[11] These tips have been created on the basis of ICO (no date) Right to rectification. Information Commissioner's Office, Wilmslow. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-rectification/ (accessed 28 May 2020).

- Do you need a proof of inaccuracy or additional information to rectify the data? If yes, please ask for further information to the data subject. Remember not to place an unreasonable burden of proof on the data subject
- You are prepared to address the right of rectification of data subjects' data, especially those generated by the inferences and profiles made by the AI solution.
- You have processes in place to ensure that they respond to a request for rectification without undue delay and within one month of receipt.
- You are aware of the circumstances when they can extend the time limit to respond to a request.
- You have appropriate systems to rectify or complete information, or provide a supplementary statement.
- You have procedures in place to inform recipients if you rectify any data you have shared with them, unless this proves impossible or involves disproportionate effort.

8.2 Right to erasure

Data subjects have a right to ask controllers for the deletion of their personal. However, the use of cloud computing, the existence of diverse servers and repositories, the possibility that the data are processed by different processors and controllers, makes it hard to ensure that all backup copies and the personal data –and not only their encryption keys- are deleted. To avoid such results, controllers should monitor procedures carefully.

Finally, controllers shall keep in mind that this right does not cover processing necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller' or when it will 'adversely affect the rights and freedoms of others'. If deleting some data might cause severe damage to the rights and freedoms of others, erasure should not be allowed. Needless to say, this involves the need to balance the different interests involved. Some essential tips are[12]

- You are aware that you need to inform individuals of their right to erasure, in addition to including it in your privacy notice.
- You know how to recognize a request for erasure and they understand when the right applies (see article 17.1 GDPR).
- You are aware that if the request satisfy one of the exemptions provided by Article 17.3 GDPR you can inform and explain to the data subject that the request shall be denied
- You have a policy for how to record requests that you receive (even verbally).
- You understand when you can refuse a request and are aware of the information you need to provide to individuals when doing so.

[12] ICO (no date) Right to erasure. Information Commissioner's Office, Wilmslow. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure/ (accessed 28 May 2020).

- You have processes in place to ensure that you respond to a request for erasure without undue delay and within one month of receipt.
- You are aware of the circumstances under which you can extend the time limit to respond to a request.
- You understand that there is a particular emphasis on the right to erasure if the request relates to data collected from minors
- You have procedures to inform recipients if they erase any data you shared with them, unless this proves impossible or involves disproportionate effort
- You have appropriate methods to erase information in robust, accountable and permanent way, which prevents you and any other party from (re-) accessing and (re-)processing the data;

8.3 Right to restrict the processing

Since a controller other than the social network who originally gathered the data is involved in data processing, it might be good to keep in mind that this right shall be exercised through any of the actors involved, who should inform the rest about the requirement and proceed accordingly. In this context, it can be very useful to develop data sharing agreements that help to clarify the responsibilities attributed to each of these roles in the performance of the specific data processing activities to be carried out, if the developer policies do not clarify this issue. Some essential tips are:

- If you receive a request to restrict data the processing from a legal entity, please indicate that the request was not lodged by an individual;
- If the individuals have not identified themselves, please ask for further information to confirm identity
- If the request does not fall within one of the scenarios laid down in Article 18.1 GDPR, please inform the data subject that the request shall be denied
- If the request cannot be fulfilled within one month, please inform why and how long will it take to process the request
- Remember that the restriction does not encompass the data storage;
- When restriction is pending, personal data can still be processed under the circumstances laid down in Article 18.2 GDPR;
- Communicate the restriction of the processing to each recipient to whom the personal data has been disclosed in compliance with Article 19 GDPR, unless this proves impossible or involves disproportionate effort.

8.4 Right to object

Data subjects must have a possibility to revoke any prior consent given to a specific data processing and to object to the processing of data relating to them. The exercise of such right must be possible without any technical or organisational constraints and the tools provided to register this withdrawal should be accessible, visible and efficient. Thus, researchers/innovators should make this option available for data subjects as soon as they start processing the data gathered from social networks. Some essential tips are:

- Have clear information in their privacy notice about individuals' right to object, which is presented separately from other information on their rights.
- Understand when you need to inform individuals of their right to object, in addition to including it in their privacy notice.
- Ensure that you know how to recognize an objection and they understand when the right applies.
- Be aware of the fact that if the request falls within one of the exceptions laid down in Article 21.2-6 GDPR, you shall inform the data subject that the request shall be denied.
- Build a policy for how to record objections you receive (even verbally).
- Understand when they can refuse an objection and are aware of the information they need to provide to individuals when doing so.
- Make sure that you have processes in place to ensure that you respond to an objection without undue delay and within one month of receipt.
- Be aware of the circumstances when you can extend the time limit to respond to an objection.
- Be able to check the data subject's particular situation aim at balancing its rights with the legitimate ones of others in processing their data.

8.5 Right to data portability

According to the GDPR, data subjects have a right to portability. However, it only applies to data 'concerning' the data subject and data they 'provided to' the controller. As a consequence, both anonymized and inferred or derived data are not included in the right to portability, since anonymized data do not concern the data subject, and inferred or derived data have not been provided by the data subject. **Some fundamental tips are:**

- Be aware that you need to inform individuals of their right to portability, in addition to including it in your privacy notice.
- Take into account the requirement for data portability from the earliest stages of conception and design of the AI processing. Otherwise, things will get seriously complicated if a data subject ask for this right.
- Make sure that you are able to recognize a request for data portability and understand when the right applies.
- Be aware of the circumstances that allow you refuse a request and be aware of the information you need to provide to individuals if you proceed with such refusal.
- If the portability request is made by several data subjects, make sure that all of them agree on the request
- If the information intertwines with the one from other individuals, please carry out a balancing test.
- Transmit data in structured, commonly used and machine-readable formats;

- The controllers inform users in advance when it is not technically possible to exercise the right of portability by means of a protocol.
- Transmit data in a secure way.
- Implement adequate processes to ensure that you respond to a request for data portability without undue delay and within one month of receipt. If it is going to take longer, inform the data subject about the delay and the time it will take to process the request.
- Be aware of the circumstances under which you can extend the time limit to respond to a request.