Ethics and data protection

14 November 2018

Contents

I. Introduction

Data protection is both a central issue for research ethics in Europe and a fundamental human right. It is intimately linked to autonomy and human dignity, and the principle that everyone should be valued and respected. For this principle to guide the development of today's information society, data protection must be rigorously applied by the research community.

The right to data protection is enshrined in the EU Charter of Fundamental Rights and the Treaty on the Functioning of the European Union, which give effect to individuals' right to privacy by providing them with control over the way information about them is collected and used.[1]

In research settings, data protection imposes obligations on researchers to provide research subjects with detailed information about what will happen to the personal data that they collect. It also requires the organisations processing the data to ensure the data are properly protected, minimised, and destroyed when no longer needed.

Depending on the setting or information in question, the failure to protect personal data against loss or misuse can have devastating consequences for the data subjects. It may also have serious legal, reputational and financial consequences for the data controller and/or processor.[2] Many recent examples of unethical research practices have involved the unauthorised collection and/or (mis)use of personal data, resulting in enforcement action by regulators.

While individual EU-funded research projects processing personal data must comply with EU and national data protection laws, the objective of this guidance note is to ensure that, in addition to respecting legal obligations, all projects are guided by ethical considerations and the values and principles on which the EU is founded.

Particular attention should be paid to research involving special categories of data (formerly known as 'sensitive data'), profiling, automated decision-making, data-mining techniques, big-data analytics and artificial intelligence, as such processing operations may pose higher risks to the rights and freedoms of data subjects (see Table 1). The increasing impact of these and other new technologies on our everyday lives and activity is reflected in the letter and spirit of the EU's 2016 General Data Protection Regulation GDPR).

While the EU's ethics review process is primarily concerned with ethics issues, your project must demonstrate compliance with the GDPR. However, the fact that your research is legally permissible does not necessarily mean that it will be deemed *ethical*.

Crucially, if your research proposal involves the processing of any personal data, whatever method is used, you – and all of your partners, collaborators and service providers – must, if called upon, be able to demonstrate compliance with both legal and ethical requirements. Such requests could come from data subjects, funding agencies or data protection supervisory authorities.

When developing and implementing your proposal, it is your responsibility to identify the appropriate legal provisions and ensure compliance. All EU projects processing personal information about identifiable human research subjects are subject to the GDPR. The principle of accountability is central to the GDPR and requires data processors to establish and document data protection compliance processes. Comprehensively addressing data protection issues in your research proposal, which will become part of your contract if selected for funding, can make an important contribution to the accountability of the project.

---

1 Article 8, EU Charter of Fundamental Rights.

2 Regulators may impose fines of up to €20 million or 4 % of the global turnover of the entity (whichever is higher).

Note that in addition to the GDPR, national legislation or related EU measures could also apply to your research:

- if your proposal uses data processed or provided by authorities responsible for preventing, investigating, detecting or prosecuting criminal offences, Directive (EU) 2016/680 may also apply;
- if your project uses personal data generated or processed by electronic networks (e.g. data relating to 'cookies', internet usage or electronic network traffic), the EU's e Privacy Directive (currently under revision) may also apply;
- EU Member States have laid down their own rules on data processing, e.g. the processing of special categories of data (such as genetic, biometric and/or health data) may be subject to additional national legal requirements, such as prior notification of regulators or data protection authorities. It is your responsibility to ensure that your research complies with the data protection laws in all the Member States in which your research data are processed, as well as the GDPR.[3]

[3] See in particular Articles 9(4), 8 and 89(3) GDPR.

[Box 1] Key issues, concepts and definitions

'**Personal data'** are defined extremely broadly and include **'any information relating to an identified or identifiable natural person'**. An **'identifiable natural person'**, or **'data subject'**, is **'one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'** (Article 4(1) GDPR).

Personal data include data such as internet protocol (IP) addresses (unique identifiers that can be used to identify the owner of devices connected to the internet) and data from 'smart meters' monitoring energy usage by addresses linked to identifiable persons.

'Special categories of personal data' (formerly known as 'sensitive data') are subject to more stringent data-protection safeguards. They include ' **personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation'**(Article 9(1) GDPR).

If your project involves the processing of special categories of data, it is more likely to raise significant ethics issues. You must therefore justify the inclusion of this kind of data in your project.

The definition of **'data processing'** is very broad. It includes **'any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'** (Article 4(2) GDPR).

It is highly likely that if your project involves any data about identifiable persons, even if they are not directly participating in the research, you are processing personal data and must comply with EU and national law. Only data that have been fully and irreversibly anonymised are exempt from these requirements. Importantly, while **pseudonymisation** can provide individual data subjects with a degree of protection and anonymity, pseudonymised data still fall within the scope of personal data because it is possible to re-identify the data subject (see below).

Even if your project is using only **anonymised data**, the origin or acquisition of the data may still raise significant ethics issues.

The GDPR places obligations on both:

- the **'data controller'**, which **'alone, or jointly with others, determines the purposes and means of the processing of personal data'**; and
- the **'data processor'**, which **'processes personal data on behalf of the controller'**.

You must ensure that any partners, contractors or service providers that process research data at your request and on your behalf comply with the GDPR and the H2020 ethics standards. Where you share with consortium partners the responsibility for processing personal data collected in the course of your research project, your project may have **joint data controllers**. In this case, you and your partners must set out your respective responsibilities in an agreement available to data subjects and provide them with a single point of contact.

II. Identifying and addressing ethics issues in your research proposal

All research proposals that involve the processing of personal data must provide information about the data protection provisions in their proposal. It is more likely that your project raises higher ethics risks if it involves:

- Processing of 'special categories' of personal data (formerly known as 'sensitive data');
- processing of personal data concerning children, vulnerable people or people who have not given their consent to participate in the research;
- complex processing operations and/or the processing of personal data on a large scale and/or systematic monitoring of a publicly accessible area on a large scale;
- data processing techniques that are invasive and deemed to pose a risk to the rights and freedoms of research participants, or techniques that are vulnerable to misuse; and
- collecting data outside the EU or transferring personal data collected in the EU to entities in non-EU countries.

[Table 1] Indicators of data processing operations that may entail higher ethics risks

Types of personal data	* racial or ethnic origin * political opinions, religious or philosophical beliefs * genetic, biometric or health data * sex life or sexual orientation * trade union membership
Data subjects	* children * vulnerable people * people who have not given their explicit consent to participate in the project
Scale or complexity of data processing	* large-scale processing of personal data * systematic monitoring of a publicly accessible area on a large scale * involvement of multiple datasets and/or service providers, or the combination and analysis of different datasets (i.e. big data)
Data-collection or processing techniques	* privacy-invasive methods or technologies (e.g. the covert observation, surveillance, tracking or deception of individuals) * using camera systems to monitor behaviour or record sensitive information * data mining (including data collected from social media networks), 'web crawling' or social network analysis * profiling individuals or groups (particularly behavioural or psychological profiling) * using artificial intelligence to analyse personal data * using automated decision-making that has a significant impact on the data subject(s)
Involvement of non-EU countries	* transfer of personal data to non-EU countries * collection of personal data outside the EU

Further indicators of the kinds of project and data-processing operation that may be considered higher-risk are provided throughout this note (see in particular the sections on data protection impact assessments and profiling, tracking, surveillance, automated decision-making and big data).

If your research entails higher-risk data processing, you must provide a detailed analysis of the ethics issues raised by your project methodology. This should comprise:

- an overview of all planned data collection and processing operations;
- identification and analysis of the ethics issues that these raise; and
- an explanation of how you will mitigate these issues in practice.

You must ensure that these issues are duly included and addressed in the research protocol that you submit to your research ethics committee. You may also be required to conduct a data protection impact assessment (DPIA) in line with Article 35 GDPR and supplementary guidance on DPIAs (see below).

If your institution has appointed a data protection officer (DPO), you should involve them in all stages of your project and seek their advice on data privacy issues. This will help in the implementation of your proposal and grant agreement (EU grants are subject to full compliance with data privacy rules).

Where complex, sensitive or large-scale data processing is envisaged or data are to be transferred outside the EU, you should consult the DPO on the compatibility of the data protection arrangements with the host institution's policies and applicable legislation.

You should include the opinion and/or advice of the DPO in your proposal. If your host institution does not have a DPO, it is recommended that you seek the advice of a suitably qualified expert.

III. Pseudonymisation and anonymisation

One of the best ways to mitigate the ethical concerns arising from the use of personal data is to anonymise them so that they no longer relate to identifiable persons. Data that no longer relate to identifiable persons, such as aggregate and statistical data, or data that have otherwise been rendered anonymous so that the data subject cannot be re-identified, are not personal data and are therefore outside the scope of data protection law.

However, even if you plan to use only anonymised datasets, your proposal may still raise significant ethics issues. These could relate to the origins of the data or the manner in which they were obtained. You must therefore specify the source of the datasets you intend to use in your proposal and address any ethics issues that arise. You must also consider the potential for misuse of the research methodology or findings, and the risk of harm to the group or community that the data concern.

Where it is necessary to retain a link between the research subjects and their personal data, you should, wherever possible, pseudonymise the data in order to protect the data subject's privacy and minimise the risk to their fundamental rights in the event of unauthorised access. Pseudonymisation and anonymisation are not the same thing and it is important that you are aware of the difference between them, as the GDPR requires you to use them wherever possible or feasible (Article 89 GDPR).

[Box 2] Pseudonymisation and anonymisation: understanding the difference

Pseudonymisation entails substituting personally identifiable information (such as an individual's name) with a unique identifier that is not connected to their real-world identity, using techniques such as coding or hashing. However, if it is possible to re-identify the individual data subjects by reversing the pseudonymisation process, data protection obligations still apply. They cease to apply only when the data are fully and irreversibly anonymised.

Anonymisation involves techniques that can be used to convert personal data into anonymised data. Anonymisation is increasingly challenging because of the potential for re-identification.

Re-identification is the process of turning pseudonymised or anonymised data back into personal data by means of data matching or similar techniques.

While anonymised data are no longer considered personal data, anonymisation processes are challenging, particularly where large datasets containing a wide range of personal data are concerned. This is because it is very difficult to create fully anonymous datasets that retain the granular information needed for research purposes.[4] As far as your research proposal is concerned, if there is a significant prospect of re-identification of persons whose data have been collected, the information should be treated as personal data. It is difficult to assess the risk of re-identification with absolute certainty and you should always err on the side of caution. A growing body of case studies and research publications in which individuals are identified from 'anonymous' datasets has demonstrated the fundamental constraints to anonymisation as a technique to protect the privacy of individuals.

If you intend to anonymise the data you collect for use in your research project, the timing of the anonymisation process is paramount. You are collecting 'anonymised' data only if the anonymisation happens at the point and time at which the data are collected from the research subject, so that no personal data are actually processed. If anonymisation takes place at a later stage, e.g. you intend to remove personally identifiable information during the transcription of audio recordings or at the point at which survey data are fed into a database, the raw data are still personal data and your proposal must include provisions for their protection up until the point at which they are deleted or rendered anonymous.

In some instances, your host institution, funding body or publisher may require you to keep the raw data for auditing, accountability or research integrity purposes. There may be other scenarios in which a host institution has a raw dataset which it makes available to its researchers and partners in anonymised form. In these instances, while the recipients of the anonymised data may – subject to the mitigation of the risk of re-identification – be exempt from data protection requirements, the host institution is still processing personal data and must therefore ensure appropriate protection for the raw (personal) data. This includes technical and organisational measures to protect the data and the means to identify the data subjects (e.g. the keys, codes or applications used to anonymise the data) against unauthorised access or use. If you are in any doubt as to the adequacy of the technique(s) that you intend to use, you should seek advice from your DPO or a suitably qualified expert. As noted below (see Box 5), for sensitive or complex processing scenarios involving pseudonymisation or anonymisation, it may even be necessary to conduct a DPIA in order to ensure an appropriate level of data protection and minimise risk to the data subjects' rights.

[4] See also *Opinion 05/2014 on anonymisation techniques*, Article 29 Working Party (adopted 10 April 2014).

IV. Data protection by design and default

To innovate ethically and responsibly, researchers and developers have long been encouraged to apply the concept of 'privacy by design', which provides a framework for focusing the design of systems, databases and processes on respect for data subjects' fundamental rights. A wider concept of 'data protection by design', now included in the GDPR, requires data controllers to implement appropriate technical and organisational measures to give effect to the GDPR's core data-protection principles (articles 5 and 25 GDPR). Data protection by design is one of the best ways to address the ethics concerns that arise from your research proposal at the design stage of your project.

In a research and development context, measures to achieve data protection by design could include:

- the pseudonymisation or anonymisation of personal data;
- data minimisation (see Box 3);
- applied cryptography (e.g. encryption and hashing);
- using data-protection focused service providers and storage platforms; and
- arrangements that enable data subjects to exercise their fundamental rights (e.g. as regards direct access to their personal data and consent to its use or transfer).

When considering whether and how to apply the principle of data protection by design, you should take into account:

- the nature, scope, context and purposes of processing;
- the severity of the risks to the data subjects' fundamental rights should you fail to protect their information; and
- the cost and availability of the technologies and applications you may need.

You must apply the principle of data protection by design where it could mitigate the ethics risks raised by the data processing in your research project, and explain in your research proposal how this will be achieved. This approach is underscored by the principle of data protection by default. **Wherever you have the possibility to enhance the level of data protection afforded to your research subjects, you should apply such measures by default rather than just considering them or making them available as an optional extra.**

Where your research involves complex, sensitive or large-scale data processing, your proposal should include a description of the measures you will take to apply the principles of data protection by design and default, and/or to enhance security so as to prevent unauthorised access to personal data or equipment.

[Box 3] Data minimisation

Data processing must be lawful, fair and transparent. **It should involve only data that are necessary and proportionate** to achieve the specific task or purpose for which they were collected (Article 5(1) GDPR).

You should therefore **collect only the data that you need to meet your research objectives**. Collecting personal data that you do not need for your research project may be deemed unethical and unlawful.

If you are in any doubt as to whether you actually need all of the data you intend to collect, you should **conduct a data minimisation review**. This should be designed and conducted by the research team to ensure that data are collected on a **'need to know' basis**, i.e. the data are required for a specific purpose that is relevant and limited to your project's objectives and methodology.

Data minimisation applies not only to the amount of personal data collected, but also to the extent to which they may be accessed, further processed and/or shared, the purposes for which they are used, and the period for which they are kept. You must minimise the processing as far as possible.

If you are unable fully to identify the purpose of the data processing at the time of data collection or you need to keep the data beyond the duration of your project, **you must explain and justify the data collection and retention arrangements**.

You must also explain how you will apply the principles of data minimisation and data protection by design in practice. In particular, you must ensure that:

- you pseudonymise or anonymise the data wherever possible (see Box 2);
- the data are securely stored; and
- where appropriate, policies and procedures are established to limit the use of the data and protect the fundamental rights of the data subjects.

V. Informed consent to data processing

Informed consent is the cornerstone of research ethics. It requires you to explain to research participants what your research is about, what their participation in your project will entail and any risks that may be involved. Only after you have conveyed this information to the participants – and they have fully understood it – can you seek and obtain their express permission to include them in your project (Articles 4(11) and 7 GDPR).[5]

In principle, living individuals should not be the subject of a research project without being informed, even in the relatively rare cases where research methods, conditions or objectives dictate that they are not made fully aware of the nature of the study until its completion. However, the advent of the internet and the widespread use of social media platforms and other ICTs have dramatically expanded opportunities for researching human behaviour without the express consent of the subjects. In turn, this has created a range of ethical dilemmas and challenges for the research community.

Whenever you collect personal data directly from research participants, you must seek their informed consent by means of a procedure that meets the minimum standards of the GDPR. This requires consent to be given by a clear affirmative act establishing a freely given, specific, informed

[5] For research involving clinical trials, data processing should also comply with Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC (OJ L 158, 27.5.2014, p. 1).

and unambiguous indication of the subject's agreement to the processing of their personal data.[6] This may take the form of a written statement, which may be collected by electronic means, or an oral statement.

Wherever possible, this process should be integrated into a broader informed consent procedure that meets the standards set out in the Commission's *Guidance note on informed consent*. However, for projects involving particularly complex or sensitive data-processing operations, or intrusive methods such as behavioural profiling, audio/video recording or geo-location tracking, you should implement a specific informed consent process covering the data-processing component of your project.

You must keep records documenting the informed consent procedure, including the information sheets and consent forms provided to research participants, and the acquisition of their consent to data processing. These may be requested by data subjects, funding agencies or data protection supervisory authorities.

For consent to data processing to be 'informed', the data subject must be provided with detailed information about the envisaged data processing in an intelligible and easily accessible form, using clear and plain language. As a minimum, this should include:

- the identity of the data controller and, where applicable, the contact details of the DPO;
- the specific purpose(s) of the processing for which the personal data will be used;
- the subject's rights as guaranteed by the GDPR and the EU Charter of Fundamental Rights, in particular the right to withdraw consent or access their data, the procedures to follow should they wish to do so, and the right to lodge a complaint with a supervisory authority;
- information as to whether data will be shared with or transferred to third parties and for what purposes; and
- how long the data will be retained before they are destroyed.

The data subjects must also be made aware if data are to be used for any other purposes, shared with research partners or transferred to organisations outside the EU (see article 13 GDPR).

As with any research project involving human subjects, **if the data processing entails potential risks to the data subjects' rights and freedoms, they must be made aware of these risks during the informed consent procedure**.

The consent process(es) and the information you give to the data subjects should cover all the data-processing activities related to their participation in your research. From a research ethics perspective, and in accordance with the principles of fair and transparent data processing, if you intend to use or make their data available for future research projects, it is best practice to obtain their additional, explicit consent to the secondary use of the data.[7] If you do plan to use the data in

---

[6] See also Article 7 GDPR and *Guidelines on consent under Regulation 2016/679*, Article 29 Working Party (adopted 28 November 2017).

[7] e.g. 'A university research department conducts an experiment analysing changes of mood on 50 subjects. These are required to register in an electronic file their thoughts every hour, at a given time. The 50 persons gave their consent for this particular project, and this specific use of the data by the university. The research department soon discovers that electronically logging thoughts would be very useful for another project focused on mental health, under the coordination of another team. Even though the university, as controller, could have used the same data for the work of another team without further steps to ensure lawfulness of processing that data, given that the purposes are compatible, the university informed the subjects and asked for new consent, following its research ethics code and the principle of fair processing' (*Handbook on European data protection law: 2018 edition*, EU Fundamental

multiple projects or for purposes other than your research, you must give the data subjects the opportunity to opt out of the further processing operation(s).

If in the course of your research project you wish to make any significant changes to your methodology or processing arrangements that have a bearing on the data subjects' rights or the use of their data, you must make them aware of the intended changes, and seek and obtain their express consent; it is not enough to offer them the opportunity to opt out. This must be done *before* you make the changes.

If your project involves complex and large-scale data processing, if you plan to use the data in multiple projects or for multiple purposes, or if it is not possible fully to identify the purpose of the data processing at the time of data collection, it may be appropriate to use a consent management application. Various service providers now offer ethically robust, secure informed consent platforms that can help you to manage, document and evidence your consent processes.

VI. Collecting data on children

All research involving children and young people raises significant ethics issues, as they may be less aware of the risks and consequences of their participation. This is also true as regards the processing of their personal data.

If your research project involves collecting data from children, you must follow the *EC Guidance note on informed consent*, in particular the provisions on obtaining the consent of a parent/legal representative and, where appropriate, the assent of the child. As that guidance makes clear, it is imperative that any information you address to a child is in age-appropriate and plain language that they can easily understand. You must also apply the principle of protection by design to research data concerning children and minimise the collection and processing of their data as far as possible.

The GDPR establishes special safeguards for children in relation to 'information society services', a broad term covering all internet service providers, including social media platforms.[8] These include a requirement for *verified* parental consent in respect of information society services offered directly to children aged under 16. Individual Member States may provide for this threshold to be lowered to 13. If you are collecting data from children using ICTs (e.g. from social media platforms or apps), you must ensure that you observe the national and EU law safeguards and explain in your proposal how you will obtain and verify the parent/legal representative's consent.

VII. Use of previously collected data ('secondary use')

As noted above, some of the most high-profile breaches of ethics standards have concerned the use of data collected for one purpose and then used for other research or targeting processes, without the knowledge or consent of the data subject. If you are processing personal data in your research without the express consent of the data subjects, you must explain how you will obtain the data, justify their use in your project and ensure that the processing is fair to the data subject.

If the collection or use of data raises specific ethics issues (e.g. as regards consent and transparency, privacy and the rights and expectations of the data subjects), you must provide a detailed overview

---

Rights Agency, European Court of Human Rights, Council of Europe and European Data Protection Supervisor (2018); http://fra.europa.eu/en/publication/2018/handbook-european-data-protection-law).

[8] See also Directive (EU) 2015/1535 of the European Parliament and of the Council of 9 September 2015 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services (OJ L 241, 17.9.2015, p. 1).

of the planned data collection and processing operations and **explain how the ethics concerns will be mitigated**.

If you are using data that are publicly available, you must provide details of the source(s) and confirm that the data are openly and publicly accessible and may be used for research purposes. You must also do this where the data you intend to use have been manifestly made public by the data subject (see Box 4).

[Box 4] Using 'open source' data

The fact that some data are publicly available does not mean that there are no limits to their use.

On the contrary, **if you take 'open source' personal data about identifiable persons and create new records or files/profiles, you are processing personal data about them** and must have a lawful/legitimate basis for doing so.

You must ensure that the data processing is fair to the data subject and that their fundamental rights are respected.

If your research project uses **data from social media networks** and you do not intend to seek the data subjects' explicit consent to the use of their data, you must assess whether those persons actually intended to make their information public (e.g. in the light of the privacy settings or limited audience to which the data were made available).

It is not enough that the data be accessible; they must have been made public to the extent that the data subjects do not have any **reasonable expectation of privacy**. **You must also ensure that your intended use of the data complies with any terms and conditions published by the data controller**.

If you are in any doubt as to what you can and cannot do with this kind of data, you should seek advice from your DPO or a suitably qualified expert and include their opinion in your proposal.

If you intend to use personal data that were collected from a previous research project, you must provide details regarding the initial data collection, methodology and informed consent procedure. You must also confirm that you have permission from the owner/manager of the dataset(s) to use the data in your project.

Where the planned use of data is predicated on the 'legitimate interests' of the data controller, the nature and purpose of the dataset must be set out in detail, together with the safeguards (e.g. anonymisation or pseudonymisation techniques) that warrant its use in your project.[9]

If your intended data processing is based on national legislation or international regulations authorising your research, or a demonstrable overriding public interest (e.g. public health, social protection) allows you to use a particular dataset, your proposal must make reference to the relevant Member State or Union law or policy.

In principle, if you are using personal data provided to you by a third party and the data subjects have not expressly consented to its use in research projects, you must, in accordance with the GDPR, inform them that you have acquired the data and what you will be using them for (art.14 GDPR). You must also provide them with the same basic information about the data processing and their rights

[9] According to the GDPR, '[t]he legitimate interests of a controller, including those of a controller to which the personal data may be disclosed, or of a third party, may provide a legal basis for processing, provided that the interests or the fundamental rights and freedoms of the data subject are not overriding, taking into consideration the reasonable expectations of data subjects based on their relationship with the controller'. See also recital 47 and Article 89 GDPR.

as data subjects that you are obliged to provide to people you are collecting data from directly (see section V). These requirements do not apply only where it is not possible or would involve a disproportionate effort to contact the data subjects. However, in such cases you must implement appropriate safeguards, including technical and organisational measures to ensure respect for the principle of data minimisation (see Box 3) and protect the subjects' fundamental rights. Crucially, the GDPR requires that pseudonymisation or anonymisation techniques (see above) be implemented wherever viable (article 89 GDPR).

VIII. Data protection impact assessments

The risk-based approach to data processing upon which the GDPR is predicated can help researchers with complex, sensitive or large-scale data processing requirements to identify and address the ethics issues that arise from their methods and objectives.

The DPIA is a process designed to assess the data-protection impacts of a project, policy, programme, product or service and, in consultation with relevant stakeholders, to ensure that remedial actions are taken as necessary to correct, avoid or minimise the potential negative impacts on the data subjects.

Under the GDPR, a DPIA is mandatory for processing operations that are likely to 'result in a high risk to the rights and freedoms of natural persons'(art.35). These include in particular:

- a 'systematic and extensive' analysis of personal data in the context of automated processing, including profiling, where this has a significant effect on the data subject;
- large-scale processing of 'special categories' of personal data, or of personal data relating to criminal convictions and offences; or
- a systematic monitoring of a publicly accessible area on a large scale.

The EU's Article 29 Working Party (WP29) has produced a longer list of scenarios in which it is likely to be necessary to conduct a DPIA (see Box 5). The European Data Protection Board and national data protection supervisory authorities are expected to clarify further those processing operations for which DPIAs are mandatory. It is your responsibility to check whether you are required to conduct a DPIA in accordance with EU or Member State rules.

If your research objectives and methods require you to conduct a DPIA in accordance with the GDPR, then provision for this must be made in your proposal. This includes details of how, when and by whom the assessment will be conducted.

Crucially, if the DPIA indicates that the envisaged processing would result in a high risk to people's rights and freedoms in the absence of measures taken by the controller to mitigate the risk, you must seek the advice of your data protection supervisory authority as to whether the envisaged processing is permissible(art.36 GDPR).This may in turn have a significant bearing on the viability of your research proposal and must therefore be addressed in your risk assessment.

If you are unsure as to whether you are required to conduct a DPIA, you should seek advice from your DPO or a suitably qualified expert, and include their opinion in your proposal. Even if you are not required to conduct a DPIA in accordance with the GDPR, it is good practice to conduct such an assessment in order to ascertain and minimise risk wherever the envisaged data processing is complex, large-scale or sensitive.

Regardless of whether a DPIA is required or conducted, if the data processing that you envisage raises significant ethics concerns, you must provide a thorough evaluation of those risks in your proposal. As a minimum, this should include the risk of unethical conduct or harm to the wellbeing

or interests of research participants at both individual (e.g. research participants, their associates or other third parties) and group level (e.g. the potential for adverse impacts on the community that the data concern).

When assessing the ethical issues arising from your research, you must consider the risk of discrimination, stigmatisation, data breaches (i.e. exposing the identity or sensitive data of individuals or damaging their reputation through a breach of confidentiality), threats to the safety or security of participants and the potential for misuse of the research methodology or findings.

[Box 5] Scenarios in which you should conduct a data protection impact assessment

WP29 considers that **processing operations raising multiple data-protection concerns are more likely to present a high risk to the rights and freedoms of data subjects, and therefore require a DPIA**, regardless of the measures, which the controller intends to adopt. The Article 29 guidance note[10] gives the following examples:

Examples of processing	Possible Relevant criteria	DPIA likely to be required?
A hospital processing its patients' genetic and health data (hospital information system).	- Sensitive data or data of a highly personal nature - Data concerning vulnerable data subjects - Data processed on a large-scale	Yes
The use of a camera system to monitor driving behaviour on highways. The controller envisages to use an intelligent video analysis system to single out cars and automatically recognise license plates.	- Systematic monitoring. - Innovative use or applying technological or organisational solutions	Yes
A company systematically monitoring its employees' activities, including the monitoring of the employees' work station, internet activity, *etc*.	- Systematic monitoring - Data concerning vulnerable data subjects	Yes
The gathering of public social media data for generating profiles.	- Evaluation or scoring - Data processed on a large scale - Matching or combining of datasets - Sensitive data or data of a highly personal nature	Yes
An institution creating a national level credit rating or fraud database.	- Evaluation or scoring - Automated decision making with legal or similar significant effect - Prevents data subject from exercising a right or using a service or a contract - Sensitive data or data of a highly personal nature	Yes
Storage for archiving purpose of pseudonymised personal sensitive data concerning vulnerable data subjects of research projects or clinical trials	- Sensitive data - Data concerning vulnerable data subjects. - Prevents data subjects from exercising a right or using a service or a contract	Yes

[10] *Guidelines on data protection impact assessment (DPIA) and determining whether processing is 'likely to result in a high risk' for the purposes of Regulation 2016/679*, Article 29 Working Party (last revised and adopted 4 October 2017).

IX. Profiling, tracking, surveillance, automated decision-making and big data

The widespread use and vast research and development potential of information and communication technologies has created a new range of ethical challenges. These include potentially adverse or unforeseen consequences for individual data subjects, specific communities and society at large. These may relate to the implications of combining and analysing different datasets, the potential for misuse of applications, or the risk of institutionalised discrimination.

If your research project involves these techniques, you must provide a detailed analysis of the ethics issues raised by your methodology. This should comprise:

- an overview of all planned data collection and processing operations;
- identification and analysis of the ethics issues that these raise; and
- an explanation of how these issues will be addressed to mitigate them in practice.

If human participants are involved in your research, you must ensure that robust informed consent procedures are in place. **Your research involves human participants if you recruit them directly, or if your research activities consist of actively involving, influencing, manipulating or directing people in any way.**

If your project involves the large-scale processing of personal data using techniques such as data-mining, 'web crawling' or social network analysis, you should address both the ethics implications of the research methods and the GDPR compatibility of the data processing.

If your project involves the automated processing or profiling of personal data (see Box 6), your proposal should address the ethical implications of the objectives, methods and expected outcomes. You should also consider the legal, social and ethical impacts of any big-data analysis,[11] in particular its potential impact on people's right to equal treatment and non-discrimination.[12]

If your project involves developing or using technology that may be used for the surveillance or tracking of individuals, it may fall within the scope of the EC dual-use regulation (428/2009) , or be vulnerable to misuse. In such cases, you must consult the EC Guidance note-Research involving dual use items and/or EC Guidance note- Potential misuse of research.

If your project entails the intensive monitoring or tracking of research participants, for example with regard to their movements, behaviour, activities or emotions, your proposal must explain what measures will be taken to protect both their personal data and fundamental rights.

If the goal of the project is to develop surveillance technologies or techniques for law enforcement purposes, your proposal should explain why the surveillance can be deemed necessary and proportionate in a democratic society, in accordance with EU values, principles and laws.

As noted above, this kind of research may require a DPIA in accordance with the GDPR or supplementary guidance issued by supervisory authorities. If your planned research activities entail multiple or particularly complex ethics issues that cannot be resolved at the proposal stage, or subsequently through a DPIA, your proposal should make provision for a broader ethical impact assessment, which should in turn be subject to review by your research ethics committee or other appropriate body.

---

11 See also *Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data (Big Data Guidelines)*, Council of Europe (January 2017).

12 Article 21, EU Charter of Fundamental Rights.

[Box 6] Automated processing and profiling

The GDPR includes specific safeguards related to the automated processing or 'profiling' of personal data that has, or is likely to have, a significant legal or material impact on the data subject (article 22 GDPR). Profiling and its impact are expressly linked to evaluation and scoring: **the more intrusive the profiling and the greater the potential effect of the result, the more likely it is to raise significant ethical and fundamental rights issues**. The GDPR safeguards are designed to enable the data subject to:

- understand that they are subject to profiling;
- know the logic behind the processing and any envisioned consequences of such processing;
- object to or opt out of the processing; and
- contest or seek human intervention in respect of the automated decision reached.

As research or development projects (rather than real-world applications), your activities may not have a significant legal or material impact on the data subject. However, in accordance with the principles of data protection by design and responsible research and innovation, **you must consider the automated processing and profiling safeguards required by the GDPR at the project development stage**. If you intend or expect your methodology to be used more widely (e.g. in a product, application or research context), you must develop the requisite safeguards.

Profiling safeguards include the use of proven and reliable mathematical and statistical methods, methods to ensure that data are as accurate as possible, and the development of models and techniques to minimise the risk of error or discriminatory effect. **Transparency and accountability to research participants are particularly important** and the GDPR also requires you to **provide data subjects with information on the automated processing, profiling and assessment, and recourse for those affected to obtain an explanation and challenge the decision reached**.

X. Data security

Whenever and however you collect personal data, you have both ethical and legal obligations to ensure that participants' information is properly protected. This is fundamental to safeguarding their rights and freedoms, and minimising the ethics risks related to the data processing.

The GDPR requires all data controllers and processors to implement appropriate technical and organisational measures to ensure a level of data security that is commensurate to the risks faced by the data subjects in the event of unauthorised access to, or disclosure, accidental deletion or destruction of, their data (art.32 GDPR).

Your proposal should provide details of the technical and organisational measures that will be implemented to protect the personal data processed in the course of your research, e.g. with reference to your host institution's and research partners' data protection and information security policies. Such measures may include the pseudonymisation and encryption of personal data, and policies and procedures to ensure the confidentiality, integrity, availability and resilience of processing systems.

Where higher-risk processing is envisaged (e.g. involving special categories or large-scale data), you should explain clearly how you will ensure an enhanced level of data security. In these scenarios, it is important that you choose appropriate research methods and data-processing tools (see Box 7).

This is vital where your research involves research subjects who are vulnerable or may be rendered vulnerable because of their participation in your research project. This may be the case, e.g. if you are collecting data on sensitive political issues or communicating with people in countries with repressive governments. Almost all communication is vulnerable to surveillance and interception,

but some channels are more susceptible than others. Wherever you believe there is a heightened risk to researchers and research participants, you should ensure that your communications are secure from unauthorised access.

[Box 7] Data security: 10 do's and don'ts

Do

- ✓ use GDPR-compliant tools to collect, process and store research subjects' personal data;
- ✓ take communications security seriously, and devise and implement dedicated protocols for your project as necessary;
- ✓ check the terms and conditions of all of the service providers you use (software, applications, storage, etc.) to process personal data within your project, in order to identify and mitigate risks to the data subjects;
- ✓ encrypt your research data and/or the devices on which they are stored, and ensure that keys/passwords are appropriately protected; and
- ✓ consult your DPO or a suitably qualified expert for advice on how to achieve a level of data security that is commensurate to the risks to your data subjects.

Don't

- ✗ collect data on a personal device such as a smartphone without ensuring that they are properly protected (e.g. consider the implications of automatic back-ups to the cloud, and the device's security features);
- ✗ use free services that may use your participants' data for their own purposes in lieu of payment, or collect data or communicate with research participants via social media platforms without first assessing the data protection implications;
- ✗ use unencrypted email, SMS or insecure 'voice over IP' platforms to communicate with vulnerable participants or those who may be subject to state surveillance;
- ✗ expose personal data to unauthorised access or use when accessing them remotely (e.g. by using insecure wifi connections) or travelling to countries where your devices may be inspected or seized; and
- ✗ assume that your research partners, collaborators or service providers have appropriate information security and data protection policies without checking that this is the case.

XI. Transfer of personal data to non-EU countries

Sending participants' personal data to partners, collaborators or service providers outside the EU raises ethical and legal issues that can be difficult to address in practice. Researchers based outside the EU may be subject to different ethical rules and their treatment of the data may fall short of EU standards.

Few non-EU countries have received an 'adequacy determination' from the European Commission indicating that they have a data protection framework offering a level of protection equivalent to that provided under EU law.[13] This means that your research subjects' data may not be adequately protected or may even be used in ways that undermine their fundamental rights. The EU requires that its ethics standards apply to all of the research it funds, regardless of the country in which it takes place. The transfer of personal data from non-EU countries is subject to strict data protection requirements under Chapter V GDPR.

[13] The list of countries covered by a Commission adequacy determination is available at: https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en

You **do not actually have to 'send' the data to a non-EU country for these provisions to apply**; if one of your partners or service providers is located outside the EU and is able to access the personal data you have collected, this amounts to a 'data transfer' in the context of the GDPR. You must give details of all envisaged data transfers to non-EU countries in your proposal. You must also ensure that the recipients of the data ensure the same level of data protection as is required under EU law.

For data transfers to non-EU countries to be lawful they must be predicated on one of the following grounds:

- the explicit consent of the data subject (which requires them to be informed in advance of any such transfers);
- an 'adequacy determination' by the European Commission in respect of the country in question;
- a data-transfer agreement containing EC standard contractual clauses giving effect to EU data protection law; or
- binding corporate rules covering both sender and recipient and approved by a national supervisory authority.

These requirements apply to all personal data transfers, regardless of the sensitivity of the data.

From a research ethics perspective, the transfer of research participants' data to non-EU countries should in principle always be based on their informed consent, which must be sought and obtained in accordance with the guidance set out above.

If your research proposal envisages the transfer of participants' data to non-EU countries without their express consent of the data subjects, then your proposal must clarify the legal basis for any such transfer. In such cases, you should seek the advice of your host institution's DPO as to the legality of the data transfer and include their opinion in your proposal. If your host institution does not have a DPO, you should seek the advice of a suitably qualified expert.

XII. Collection of personal data outside the European Union

Collecting personal data from research subjects in non-EU countries raises similar ethical issues, but these may be amplified by the need to ensure that the participants are:

- wholly comfortable with being part of a research project conducted by researchers from outside their own country;
- aware of what will happen to their data; and
- not subject to any undue pressure to participate.

As noted above, the EU's ethics requirements apply to all EU-funded research, irrespective of where it takes place. Similarly, the GDPR applies to all data-processing operations conducted by data controllers based in the EU, irrespective of where the processing takes place. This means that, even if you are collecting personal data outside the EU, you must still ensure and be able to demonstrate compliance with EU law.

You also have to comply with the laws of the country in which you are conducting your research, including any national data-protection laws. For example, you may be under an obligation to notify or seek permission for your research from national authorities or data protection regulators . Further authorisations may be required to transfer personal data outside the country in which the research takes place. 'Data sovereignty' provisions may even prohibit the transfer of certain kinds of information, such as health or patient data, out of the country.

It is your responsibility to determine what legal obligations apply to any research you conduct outside the EU and to take whatever action is necessary to comply with them. You must also be able to demonstrate compliance upon request. Again, if you are unsure as to how to handle issues related to international data transfers, you should seek the advice of your host institution's DPO, or a suitably qualified expert, and include their opinion in your proposal.

[Box 8] Checklist: international data transfers

Transferring personal data out of the EU

- ✓ ensure that any international data transfers fulfil at least one of the relevant conditions in Chapter V GDPR;
- ✓ check that any third-party services you intend to use (e.g. survey tools, data analytics, cloud storage, etc.) are incorporated in an EU Member State or legally represented in the EU in accordance with the GDPR;
- ✓ adopt legally binding and enforceable agreements with partners or service providers prior to data transfers;
- ✓ prohibit the onward transfer of personal data by members of your consortium and any other recipients outside the framework of such agreements; and
- ✓ implement appropriate organisational and technical measures to ensure that personal data are transferred securely.

Collecting personal data in non-EU countries

- ✓ ensure that processing, notification, consent and accountability provisions meet GDPR standards;
- ✓ identify any further data protection requirements in applicable laws in the country in which data are to be collected and explain in your proposal how you will comply with them;
- ✓ if applicable, ensure that research participants understand and consent to the export of the personal data they provide to an EU Member State or an non-EU country;
- ✓ use pseudonymisation or anonymisation techniques to minimise the risk to data subjects;
- ✓ implement appropriate organisational and technical measures to ensure that personal data are transferred securely.

XIII. Deletion and archiving of data

You may keep the personal data you collect only as long as it necessary for the purposes for which they were collected, or in accordance with the established auditing, archiving or retention provisions for your project. These must be explained to your research participants in accordance with informed consent procedures.

Recent high-profile cases involving the misuse of personal data have stemmed from data controllers' failure to delete personal data and ensure that third parties to whom the data were provided had done the same in accordance with the agreed terms of their use.

As soon as your research data are no longer needed, or subject to an established retention period, you must securely delete the data in their entirety and make sure that they cannot be recovered. Data retained for auditing processes should be stored securely and further processed for those purposes only.

If research data are held in the cloud or by a third-party service provider, you should ensure that it has securely deleted the data together with any back-ups. If data have been shared with partners or transferred to third parties in the course of your project, you should ensure that they have deleted the data, unless they have a legitimate basis for retaining them.

XIV. Data protection officers and other sources of help

If your institution has appointed a DPO, it is recommended that you seek their advice as to your data protection obligations and how to meet them. You must ensure that the DPO's contact details are made available to all the data subjects involved in your research.

If your project raises complex data protection issues due to the sensitivity of the data, or the scale or nature of the processing involved, you should consider appointing a data protection specialist/adviser to your project or research ethics board. If your host institution does not have a DPO, you should seek the advice of a suitably qualified expert in the preparation of your proposal and/or appoint such an expert to your project if necessary.

If you need help and advice addressing the broader ethics issues raised by the data processing in your project, you should contact the relevant institutional bodies or services (e.g. research office or research ethics committee) in your university or institution, relevant national bodies, members of your consortium, or colleagues in your personal network who may have relevant expertise and experience.

If you are uncertain about any aspects of ethics in your research, you should consider appointing an ethics advisor or engaging an ethics mentor to provide advice, oversee the ethical concerns in your research and ensure that it is fully ethically compliant.