**Data protection & scientific research manual**

*Background information document[[1]](#footnote-1)*

# **1.** **A special GDPR regime for scientific research?**

The GDPR affords a **special regime** with a **degree of flexibility** for **public interest-driven scientific research projects** that operate within an ethical framework, and that aim to grow society’s collective knowledge and well-being. The special regime entails **consequences** for the GDPR **data processing principles** including the **lawfulness** requirements and **purpose limitation**, and for **some of the data subject access rights**. It permits some **derogations** from the **usual data controller obligations**. This includes the **presumption of compatibility** of data processing of data for **scientific research purposes** that was collected in commercial and other contexts, provided that appropriate safeguards are in place.

This flexibility is afforded on the assumption that research emerging within a framework of ethical oversight **serves**, as a principle, the **public** interest. The **accountability** **principle** remains key; it requires data controllers to assess and to manage responsibly the risks inherent to their research project.

Data protection **obligations** generally **scale up** according to the **risk** that the **processing activities** may pose to the **data subject**, but the GDPR recognises that **scientific** **research-related rights** and **interests** can **justify** some **derogation** from some of GDPR’s **general principles**. The **scope** of derogation is however **restricted,** because the **essence** of the right to **data protection** must be **preserved**.

In this background document, we unpack **who** in the context of **D3I projects** is responsible for **compliance** with the **GDPR**, and how the **derogations** work in **practice**. We elaborate **other important GDPR requirements** for researchers to consider in **separate documentation** about “**consent**” and “**Data Protection Impact Assessments**”.

# **2. An overview: GDPR requirements & scientific research in D3I**

## 2.1 Controllers and data subjects in D3I projects

The **GDPR** is largely about **assigning roles** and **responsibilities** across the stakeholders involved in data processing. In the context of **D3I**, the processing of **personal data** potentially entails **risks** for the **person to whom the data relates**. It is therefore important to know who is ***accountable*** and ***responsible*** for (parts of) for these risks, and how they can be addressed.

In the **D3I-project**, three key stakeholders are involved: (1) **companies** possessing and processing personal data (e.g., Youtube, Facebook); (2) **participants** whose personal data is made available to the researcher, the (3) **researcher** who carries out the investigations in data that was made available by the participants. There can be a (4) **fourth** stakeholder, which is a party that does something with the personal data **on behalf of** the researcher (e.g., store it, make it accessible) – this can be a **platform**.

First, we give some **key definitions** in the **GDPR**:

'***Data processing***’ in the context of the GDPR means **everything you do with** **personal data** or with **sets of personal data**, such as collection, recording, organisation, structuring, storage, adaptation, retrieval, use, disclosure by transmission, computation, structuring, storage, dissemination, etc. Note that GDPR’s understanding is **broader** compared to how it is **usually understood** by **computer scientists** and other **tech folk** (they often consider data processing only as performing computation or analytics over data).

Because ***personal data*** – **any information relating to an identified or identifiable natural person** – are processed in D3I-projects, the GDPR applies.

The GDPR terms the **persons whose personal data is processed** as ***data subjects***. In the context of D3I, the **participants** sharing their data are the **data subjects**.

The GDPR defines***data controllers*** as “the natural or legal **person** […] which, **alone** or **jointly** with others, determines the ***purposes*** and ***means*** of the processing of personal data [...]”. In the context of D3I, **researchers** (and ultimately the university they are working for) are considered data controllers, because they determine the purpose for which the personal data will be processed. **Companies** processing data of research participants are also regarded as data controllers.

Because the researcher and the companies determine for what purpose the personal data will be processed, the **GDPR puts** **responsibilities** on researchers and companies. **Researchers** must comply with GDPR’s data protection principles, with the rights of data subjects and with a range of **other** **obligations**. The same applies to companies processing participant data. Before elaborating the responsibilities, we first introduce another role that may be relevant in D3I-research contexts: that of the data processor.

***Data processors*** are stakeholders that process personal data ***on behalf of*,** or as tasked by, **data controllers**. Where a **platform** is part of D3I projects, it likely processes personal data on behalf of the **researcher**. As long as the platform does **not** **determine any purpose for the data processing itself**, it acts as a data processor. As soon as the platform determines its own purpose for the processing of participant data, it acts, for that part, as a data controller. **Processors** do have responsibilities too, but **most GDPR-obligations** regarding the data processing fall on the data **controller**.

## 2.2 Data processing principles

The GDPR nails down **six data processing principles** with which researchers must comply:

1. lawfulness, fairness and transparency;
2. purpose limitation;
3. data minimisation;
4. data accuracy;
5. storage limitation; and
6. integrity and confidentiality, and the **accountability** **principle.**

We briefly mention what each principle means. More generally, a ‘principle’ means not a hard rule that exactly prescribes what you must do to be compliant. A GDPR-principle requires the researcher to consider and respond to the topic at hand (e.g. lawfulness), and to design their data processing in such a way that the design reflects that technical, organisational and legal measures were taken to best safeguard the rights and interests of the participants.

* **Lawfulness** means that personal data must be processed in a lawful way. In the context of D3I, this means that participant data will generally be processed on the basis of (the lawful ground of) **consent**.
* **Fairness** means that the researcher has considered how the **processing can affect** the **participants**. The researcher can **justify** **any adverse impact.** The **researcher** only handles **participant** data in ways they would **reasonably expect**, or he or she can explain why any **unexpected processing** is justified. The researcher does not deceive or mislead participants when their personal data are collected.
* **Transparency** means that the researcher is open and honest, and that he or she complies with the participant’s right to be informed about the data processing.
* **Purpose limitation** means that researchers must **from the start** be clear about what their **purposes** for **processing participant** **data** are. **Researchers** must record their purposes as part of their **privacy policy** and/or their **consent form**.
* **Data minimisation** means that the researcher collectsthe minimum amount of personal data that is needed to deliver a research result: it means that researcher should **not collect more**, but also that the researcher should **not collect less** data than they need to perform their research.
* **Data accuracy** means that the researcher should **take reasonable steps** to ensure the personal data he or she holds is **not incorrect**. If a researcher discovers that personal data is incorrect, he or she must take reasonable steps to **correct** or erase it as soon as possible.
* **Data storage** means that a researcher should **not keep personal data for longer than they need** it. Researchers need to think about a- and be able to justify – how long they keep personal data. This largely depends on the purposes for the processing of the data.
* **Integrity** and **confidentiality** mean that the researcher must ensure that appropriate security measures are in place to protect the personal data he or she is processing.
* **Accountability** requires researchers (and where that applies, the platform) to have **measures** and **records** in place to **reduce risks** for the **participant’s** **personal data,** and that the researcher can **demonstrate** this.

Compliance with these principles is important, as compliance is an important condition for **participant** **trust** in the **researcher’s processing** of **participant data**, and, not less important, for **societal trust** in **scientific research** with participants engaging in that research **more generally**.

In the context of **scientific research**, the principles of ***purpose limitation*** and ***storage limitation*** might apply **a bit less strictly – but note that they still apply!** We next explain why and how these derogations may apply in **D3I data processing contexts**.

### 2.2.1 Derogation 1: purpose limitation

As a rule, under the principle of purpose limitation, **personal data** must always be **collected** for **specified**, **explicit** and **legitimate purposes**. **Further processing** of the **same data** is **not permitted** for purposes that are **incompatible** **with the original purpose** for processing. But whenever **personal data** are **further used** for **compatible purposes**, ‘no legal basis separate from that which allows the collection of the personal data is required’. In other words, if personal data are **further** used for **compatible** **purposes**, **such use** will **not** need a **new** determination for a **legal basis**. This is **premised** on ‘the **reasonable expectations** of **data subjects** based on their **relationship** with the **controller** as to [the data’s] further use’.[[2]](#footnote-2)

The **GDPR** also establishes **criteria** for **determining** the **compatibility** of ***further*** **use** of personal data. Where further (or ‘secondary’) use is considered, **controllers** must consider ‘***any link*** between the **purposes** for which the **personal data** have been **collected** and the **purposes** of the **intended further processing**’ or the ***context*** in which **personal data** were **collected’**. If the controller **shares** or **further processes** the data for purposes ***incompatible*** with the original purpose, the **controller** must, as a rule, ***again obtain consent*** for the **new** purpose.

**Yet**, further processing for ***scientific research*** purposes or ***statistical purposes*** is usually **not** considered **incompatible** with the initial purposes.[[3]](#footnote-3) This presumption, however, **cannot stand by itself**; **whether** it **applies** depends on the **presence** of **effective** and **appropriate technical** and **organisational safeguards**, such as **pseudonymisation**, and **access limitations**.[[4]](#footnote-4) The presumption is moreover **not a general authorisation** to further process data in all cases for scientific purposes; **each case** must be **considered** on its **own merits** and **circumstances**.

### 2.2.2 Derogation 2: storage limitation

Personal data should, as a rule, be ‘**kept** in a **form** which **permits identification** of **data subjects** for ‘**no longer** than is **necessary’**. However, the GDPR permits ‘storage for **longer periods**’ if the ***sole* purpose** is **scientific research** (or archiving in the public interest, historical research or statistical purposes). The derogation does however **not apply to infinite** storage terms of data in a scientific context. The **intention** **behind the limitation of this derogation** is to **dissuade** ***unlimited*** **storage** – even in the special regime for scientific research. The legislator also aimed to **prevent** **scientific** **research** as a ***pretext*** for **longer storage** for **other, private, purposes**.

## 2.3 Consent

The GDPR defines **consent** as “**any freely given, specific, informed** and **unambiguous** **indication** of the **data subject’s** wishes by which he or she, by a **statement** or by a **clear affirmative action**, signifies **agreement** to the **processing** of **personal data** **concerning him** or **her** […]”.[[5]](#footnote-5)

Some **researchers** have sought to rely on ‘**broad consent formulations**’ to the use of personal data of participants. The **GDPR** indeed **acknowledges** that - in the context of scientific research – it might **not always be possible** to **fully identify** the specific ***purpose*** of personal data processing for scientific research purposes ***at the time of data capture***.[[6]](#footnote-6)

Therefore, **data subjects** should, if the purposes cannot be made very specific at the time where consent is obtained, be allowed to **consent** to the **use** of their personal data in **specified areas** of **scientific research.** But that data must still be **kept** and **stored** in line with **recognised ethical standards** for **scientific research**. **Data subjects** should also have the **opportunity** to give their **consent** only to **certain areas** of **research**, or to **parts of research** projects to the extent **allowed** by the **intended purpose**.[[7]](#footnote-7)

Processing ***special categories of data***[[8]](#footnote-8) is prohibited, unless the participant has given his or her ***explicit* consent**.[[9]](#footnote-9) Explicit consent, **described** as ‘an **express statement of consent’** which can be **demonstrated** in the event of doubt, is required in situations where there may be particular risk to the rights of the data subject.

As **consent** is the **lawful ground** for **processing** in **D3I-context**, the **data subject** must be able to ***withdraw*** their ***consent*** at ***any time***. There is ***no exception*** to this requirement for **scientific research**. The **withdrawal** of consent **does not affect** the **lawfulness** of the **processing** based on consent ***before* the withdrawal**. In other words, if consent is **withdrawn**, the controller is **required** to **stop** the **processing** **actions** from the moment of consent withdrawal, but they **do not have to delete the personal data** that was processed **up until consent was withdrawn**.[[10]](#footnote-10)

If the controller wishes to process that data after consent was withdrawn, the controller can only do so if the data concerned can be ***anonymised*.**

## 2.4 Risk reduction & DPIAs

**Data protection impact assessments** (***DPIAs***) ***aim*** to provide data **controllers** with tools to ***manage data******protection risks***. While DPIAs initially emerged as voluntary measure, under the GDPR they are now a **mandatory** requirement – if certain conditions are met. They are particularly required in the case of a **systematic** and **extensive** **evaluation** of **personal aspects** **relating** to **natural persons**, where that evaluation is based on **automated processing**, including profiling, and on which **decisions** are based that **produce legal effects concerning the natural person**, or that **similarly, significantly, affect the natural person**.

Research carried out in the context of the **D3I project** will likely fall **within the scope of obligatory DPIA**s, as the personal data captured and used (young people, detailed insights of personal lives, with likely special category data potentially revealing a person’s ethnicity, religion or beliefs or sexual orientation involved) is generally of a **highly sensitive nature** (with GDPR's **special category** data involved).

A DPIA contains a systematic and subsequent description of

1. the **purposes** and the **means** through which the **purposes** as defined by the researcher should be realised, including the envisaged processing operations (Art. 35(7)(a) GDPR);
2. an **assessment** of the **risks to the rights and freedoms** of the participants (Art. 35(7)(c) GDPR);
3. an assessment of the **necessity** and **proportionality** of the **processing operations** in relation to the **purposes** (Art. 35(7)(b) GDPR);
4. and the **measures** envisaged to **address the risks**, including safeguards, security measures and mechanisms to ensure the protection of personal data, and to **demonstrate compliance** with the regulation, **taking into account the rights and legitimate interests of data subjects** and **other persons** concerned (Art. 35(7)(d) GDPR).

The DPIA-application will be further elaborated in the **DPIA manual**.

## 2.5 Data subject rights

One of the **essential motivations** behind the **GDPR** is to **(re-)empower data subjects,** and with this, participants donate their data for scientific research purposes in D3I contexts. They aim to strengthen participant’s rights pertaining to their data by giving them some control over the processing of their data. The GDPR addresses these rights, specifying in detail their respective modalities, conditions and exceptions:

*Table 1: overview of GDPR’s data access rights*

|  |  |
| --- | --- |
| **Article** | **Data right** |
| 12 | Transparent information, communication and modalities for the exercise of the rights of the data subject |
| 13 | Information to be provided where personal data are captured from the data subject |
| 14 | Information to be provided where personal data were not captured from the data subject |
| 15 | Right of access by the data subject |
| 16 | Right to rectification |
| 17 | Right to have one’s personal data erased |
| 18 | Right to restriction of data processing |
| 19 | Notification obligation concerning rectification or erasure of personal data or restriction of processing |
| 20 | Right to have one’s data ported to another data controller |
| 21 | Right to object against data processing |
| 22 | Right not to undergo automated individual decision-making, including profiling |

### 2.5.1 Core rights in D3I context

The data rights particularly relevant in the context of D3I are highlighted in Table 1.

* Article 12 – Transparent information, communication and modalities for data rights exercises

The data controller – i.e., the researcher, must take appropriate measures to **inform** the **participant** (see next, under Article 13) and to **communicate** that information in a **transparent** and **easily accessible form**, using **clear** and **plain language**. The researcher also **facilitates** the **exercises** of **data rights** (articles 15 – 22, see below).

* Article 13 – right to be informed

Individuals have the **right to be informed** about the **collection** and **use of their personal data**. They must be provided with information, including research purposes for processing their personal data and retention terms. This information should be made available to individuals **at the time a researcher collects** their personal data. The information you provide to people must be **concise**, **transparent**, **intelligible**, **easily accessible**, and it must use **clear** and **plain language**.

Researchers in D3I will be provided with a **template** for a **privacy policy** that aims to **inform potential research** **participants** about what may be done with their data: the **researcher** will **share** this privacy policy in order to inform the individual about the topics mentioned in the previous paragraph. We (the legal support to the D3i project) will **regularly review**, and where necessary, **update** the **privacy policy template**. Researchers also have to bring **any new uses** of an individual’s personal data to their attention, before they can start the participant’s data processing.

Getting the right to be informed correctly can help researchers to **comply** with other aspects of the GDPR and **build trust** with people. Getting it **wrong** can leave you open to **fines** and lead to **reputational damage.**

* Article 15 – data access

The right of access to one’s personal data gives individuals the **right to obtain** a **copy** of their **personal data**, as well as other supplementary information. It helps them to **understand** how and why a data controller – for instance, a company offering a search engine – is **using** their data, and check **whether** the **data** is processed **lawfully**.

* Article 16 – rectification

The GDPR includes a right for individuals **to have inaccurate personal data rectified**, or **completed** if it is incomplete. This right is closely related to the controller’s obligation under the accuracy principle mentioned in §2.2.

* Article 22 – automated decision-making and profiling

The GDPR applies **automated individual decision-making** (making a decision solely by automated means without any human involvement) and on **profiling** (automated processing of personal data to evaluate certain things about an individual). The GDPR has additional rules to protect individuals whenever a controller carries out solely automated decision-making that has **legal** or **similarly significant effects** on **individuals**. This type of decision-making can only be used by researchers in a D3I- context where the decision is based on the individual’s **explicit consent**.

We however expect that **research** in the context of **D3I** will **not** have **any legal** or **similar**, **significant effect** to the **participants** involved in the **research**. When in doubt, the researcher should consult the Data Protection Officer (DPO) of their faculty/university.

### 2.5.2 The key role of data rights exercises in D3!

Data rights sit at the core of the D3I-initiative. **Research participants** exercise their **right of access** to their data towards **social media providers** and other **organisations** holding their **personal data**. These organisations, being data controllers, **must comply** with such **data access** **requests** whenever they receive such a request, i.e., they have to respond to such requests within one month. Many organisations have some form of procedure in place that helps the organisation to identify and transmit the response to the data subject, i.e., the participant partaking in D3i-research. The **transmits** the responses to the **researcher**.

### 2.5.3 Exceptions for scientific research to data rights

**As a rule**, GDPR’s data rights also apply to scientific researchers. Whenever a researcher receives a data access request, they must respond without delay and **within one month of receipt** of the request. The response might be extended by a **further two months** if the request is **complex** or if the researcher receives **a number of requests** from the **individual**.

The GDPR however outlines **specific conditions** under which **EU or Member State law** **may derogate** from the data subject’s **right of access** (Article 15), **right to rectification** (Article 16), **right to restriction** (Article 18) and **right to object** (Article 21). In the context of **research conducted in D3I**, only the right of **access** and the right to **rectification** are relevant. Depending on what the **national legislator determined**, these derogations apply **also** to **scientific research** and thus to research carried out in the context of **D3I**.

In the **Netherlands**, the legislator created a **possibility** for researchers to **derogate** from the data rights; however, it **left** to the **universities** how these derogations can be **practically** implemented.[[11]](#footnote-11) The **University of Amsterdam** leaves this **choice** to the **individual researcher**, but it **encourages researchers** to **comply** with the **requirements** (and to **not use** the **derogation**).

It is important to realise that the **rights of access** and **rectification** are set out in the **EU-Charter of fundamental rights** and that they are therefore generally considered **essential elements** of the right to the **protection of personal data**. In other words, they have the **status of fundamental rights**, which generally require **very weighty reasons to derogate from**.

The **right of access** is of particular **importance** as it **enables the data subjects** to **exercise all other data rights** provided for the GDPR. Therefore, **any derogation** from these essential data subject rights must be subject to a **high level of scrutiny**.

## Annex 1 – text of Article 89 GDPR (derogations for scientific research)

Article 89 of the GDPR provides for flexibility in the obligations on controllers and an emphasis on safeguards and accountability:

Article 89 GDPR

1. **Processing** for archiving purposes in the public interest, **scientific** or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that **technical and organisational** measures are in place in particular in order to ensure respect for the principle of **data minimisation**. Those measures may include **pseudonymisation** provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by **further processing** which does not permit or no **longer permits the identification** of data subjects, those purposes shall be fulfilled in that manner.

2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for **derogations** from the rights referred to in **Articles 15, 16, 18** and **21** subject to the **conditions** and safeguards referred to in **paragraph 1 of this Article** in so far as such rights are likely to render **impossible** or **seriously** **impair the achievement of the specific purposes**, and such derogations are **necessary** for the **fulfilment** of those purposes.

3. Where personal data are processed for **archiving purposes** **in the public interest**, Union or Member State **law** may provide for **derogations** from the rights referred to in **Articles 15, 16, 18, 19, 20** and **21** subject to the **conditions** and safeguards referred to in **paragraph 1 of this Article** in so far as such rights are likely to render **impossible** or seriously **impair the achievement** of the specific **purposes**, and such derogations are **necessary** for the **fulfilment** of those purposes.

4. Where processing referred to in paragraphs 2 and 3 serves **at the same time another purpose**, the **derogations** shall apply **only** to processing for the **purposes** referred to in **those paragraphs**.

Annex 2 – Bibliography

**Bibliography** on the derogation for scientific research of Article 89 GDPR:

Dove, Edward S. ‘The EU General Data Protection Regulation: Implications for International Scientific Research in the Digital Era’. *Journal of Law, Medicine & Ethics* 46, no. 4 (1 December 2018): 1013–30.

Ducato, Rossana. ‘Data Protection, Scientific Research, and the Role of Information’. *Computer Law & Security Review* 37 (1 July 2020): 105412.

European Data Protection Supervisor. *A Preliminary Opinion on Data Protection and Scientific Research. Brussels.* EDPS. 6 January 2020.

Vayena, Effy. *How the General Data Protection Regulation Changes the Rules for Scientific Research. Pdf Brussels*, Study PE 634.447. European Parliamentary Research Service. July 2019.

1. This background document is based on the European Data Protection Supervisor’s Preliminary Opinion on data protection and scientific research (6 January 2020). [↑](#footnote-ref-1)
2. Recital 50 GDPR. [↑](#footnote-ref-2)
3. Art. 89 (1) GDPR. [↑](#footnote-ref-3)
4. Art. 89 (1) GDPR. [↑](#footnote-ref-4)
5. Art. 4 (11) GDPR. [↑](#footnote-ref-5)
6. Recital 33 GDPR. [↑](#footnote-ref-6)
7. Recital 33 GDPR. [↑](#footnote-ref-7)
8. Art. 9(1) GDPR: "Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of generic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or concerning a natural's person's sex life or sexual orientation shall be prohibited". [↑](#footnote-ref-8)
9. See footnote 8. [↑](#footnote-ref-9)
10. Art. 7 (3) GDPR: "The data subject shall have the right to withdraw his or her consent at any time. The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. Prior to giving consent, the data subject shall be informed thereof. It shall be as easy to withdraw consent as to give consent.” [↑](#footnote-ref-10)
11. Art. 44 UAVG. [↑](#footnote-ref-11)