

D2.5: Guidelines on Open Access and Restricted Data (draft)

Author(s)	Helen Frew (LIBER), Rob Baxter (EPCC), Catherine Inglis (EPCC), Maria Francesca Iozzi (UiO), Simon Lambert (STFC), Heiko Tjalsma (DANS)
Status	Final
Version	v1.0
Date	30/08/2016

Abstract:

This report focuses on personal data as an important category of restricted data, and looks in particular at the legal requirements, both current and future, on EUDAT service providers storing and processing personal data. Guidelines and recommendations for EUDAT service providers are made on the basis of the newly-agreed EU General Data Protection Regulation (GDPR). A second version of this report will extend guidelines to cover broader categories of restricted data.

Document identifier: EUDAT2020-DEL-WP2-D2.5	
Deliverable lead	LIBER
Related work package	WP2
Author(s)	Helen Frew (LIBER), Rob Baxter (EPCC), Catherine Inglis (EPCC), Maria Francesca Iozzi (UiO), Simon Lambert (STFC), Heiko Tjalsma (DANS)
Contributor(s)	Melanie Imming (LIBER), Hege van Dijke (LIBER)
Due date	31/08/2016
Actual submission date	30/08/2016
Reviewed by	Pawel Kamocki, Damien Lecarpentier
Approved by	PMO
Dissemination level	PUBLIC
Website	www.eudat.eu
Call	H2020-EINFRA-2014-2
Project Number	654065
Start date of Project	01/03/2015
Duration	36 months
License	Creative Commons CC-BY 4.0
Keywords	Policy, personal data, GDPR

Copyright notice: This work is licensed under the Creative Commons CC-BY 4.0 licence. To view a copy of this licence, visit <https://creativecommons.org/licenses/by/4.0>. 

Disclaimer: The content of the document herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the document is believed to be accurate, the author(s) or any other participant in the EUDAT Consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

TABLE OF CONTENT

EXECUTIVE SUMMARY	5
1. INTRODUCTION – THE OPENNESS OF DATA	6
1.1. Constraints on openness.....	6
1.2. Why EUDAT must be prepared	6
1.3. Open data in European research projects	7
1.4. Scope and approach of this report.....	7
1.5. Structure of the report.....	8
2. ESTABLISHED EUDAT POLICIES FOR OPEN DATA	9
2.1. Summary of EUDAT open data policy	9
2.2. Openly discoverable.....	9
2.3. Openly accessible.....	9
2.4. Openly useable.....	10
3. MANAGEMENT OF PERSONAL DATA	11
3.1. Current data protection rules in the European Union	11
3.1.1. Article 29 Working Party clarifications	12
3.2. Data Protection Reform	13
4. THE APPLICATION OF DATA PROTECTION IN EUROPE: THREE CASE STUDIES.....	14
4.1. Data protection in Europe.....	14
4.2. Country case studies	14
4.3. Implications for EUDAT	15
5. FUTURE LEGISLATION ON PERSONAL DATA – THE GDPR	16
5.1. Data protection at EU level	16
5.2. Personal Data and Processing	16
5.2.1. Personal data	16
5.2.2. Special categories	16
5.2.3. Processing	17
5.3. Data controller and data processor	17
5.4. Applicable law	17
5.5. Principles and obligations	18
5.6. Legal grounds and consent	18
5.6.1. Consent	19
6. CLASSIFYING DATA IN THE EUDAT CDI AND RELATED SERVICES.....	20
6.1. Administrative data.....	20
6.2. Content data	22
6.3. A simple classification scheme.....	22
7. DATA CONTROLLERS AND DATA PROCESSORS IN EUDAT	24
8. RECOMMENDATIONS	26
8.1. General guidance for the EUDAT CDI and Collaboration	26
8.2. Considerations and issues for specific EUDAT Services	27
8.2.1. B2ACCESS.....	27
8.2.2. B2SHARE, B2DROP	27
8.2.3. B2FIND	28
8.2.4. B2SAFE, B2STAGE.....	28
8.2.5. B2HANDLE.....	28
8.2.6. Future Services	28

9. CONCLUSION AND PLANS FOR FURTHER WORK.....	29
10. REFERENCES.....	30
ANNEX A. COUNTRY CASE STUDIES (UK, NETHERLANDS, NORWAY).....	31
ANNEX B. TEMPLATE FOR PRIVACY POLICY AND DISCLAIMER	36
ANNEX C. TEMPLATE AGREEMENT FOR DATA CONTROLLER-DATA PROCESSOR	38
ANNEX D. GLOSSARY.....	43

LIST OF FIGURES

Figure 1. Mind map of the types of data stored within the EUDAT CDI, seen through the lens of the GDPR. The annotations OK, NO, PC, RC, PR and SR are described in the classification section below.	21
Figure 2. Simplified picture of the relationship between data subject, data controller and data processor.	24
Figure 3. Potential flows of personal data between EUDAT service providers (arrows indicate "from-to"). Orange denotes administrative data, as collected directly by EUDAT services; blue indicates content data, which could include personal data of either the illustrated user or another data subject. Gold boxes indicate EUDAT services (properly, service providers). The grey box illustrates a community data repository making use of EUDAT services. Service providers are labelled DP if they take the role of data processor, DC if data controller.	25

EXECUTIVE SUMMARY

This report focuses on personal data as a special, and important, category of restricted data, and looks in particular at the legal requirements, both current and future, on information service providers storing and processing personal data. Guidelines and recommendations for EUDAT service providers are made on the basis of the newly-agreed EU General Data Protection Regulation (GDPR).

A distinction is drawn between personal data in *content* that might be stored in EUDAT services and *administrative* personal data that EUDAT services might collect or manage as part of their operation. In the latter case, we summarise a number of key challenges under the GDPR that EUDAT service providers will need to be aware of: the principle of data minimisation in service design; new clarity on the rights of users over their personal data; explicit and active consent from users for the use of their personal data; protection of the privacy of children; legal separation of the roles of data controller and data processor.

For the former case, the complexity of subjects' personal data in uploaded content will require greater care in service design (impact assessments) and in deciding which research data repositories to "accept" into the EUDAT Collaborative Data Infrastructure (CDI). A risk-based classification scheme is introduced as a first step in supporting EUDAT service providers in their approaches to personal data of particularly sensitive or restricted nature. We note that certain kinds of data may be outside both the risk appetite of EUDAT and its stated mission to support open data.

The main impacts of the GDPR are analysed in the context of EUDAT's core CDI services; the largest impact falls on user-facing services, particularly B2ACCESS and B2SHARE. Consent for use of personal data such as names, online identities or email addresses must be explicitly collected and recorded, or their use must be minimised or eliminated. Service providers for B2ACCESS and B2SHARE will find themselves in the role of data controllers; their interactions with other EUDAT service providers must be codified in due course in a contract or other legal form.

The new regulations provide the opportunity to create a harmonised privacy policy and service disclaimer for EUDAT. Suggested wording for these is provided in annexes to the main report; we remark that, while based on existing policies, the authors do not have legal training and final versions must be subject to suitable legal review.

1. INTRODUCTION – THE OPENNESS OF DATA

In the context of EUDAT, it hardly needs reiterating that there is a powerful and universal trend for openness of data—including but not restricted to research data. The G8 Open Data Charter [1] declares that “open data are an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of our citizens and allow innovation and prosperity to flourish”, and sets out five principles that will be the foundation for access to, and the release and re-use of, data made available by G8 governments. Science and research is recognised as one of the areas of high-value data. The European Commission published in 2011 a Communication entitled “Open data: An engine for innovation, growth and transparent governance” [2] that singles out the acceleration of scientific progress as one of the reasons why open data is crucial for Europe.

Funding agencies increasingly require open access to research data in the investigations that they support. For example, Research Councils UK has a set of principles [3] starting with “Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner.”

Such statements and initiatives are not mere aspirations or impositions from on high. The Research Data Alliance has mobilised over 4,000 individuals in pursuit of its mission to build “the social and technical bridges that enable open sharing of data.” [4] EUDAT itself aims to support sharing and reuse of open data through its services, while recognising that not all data will be completely unrestricted.

1.1. Constraints on openness

There are some restrictions on the general openness of data that must be acknowledged. Many stakeholders have an interest in controlling or restricting access to some digital material, and not only for selfish reasons. The enforcement of restrictions might be underpinned by legislation, or by policies and practices of particular organisations. In any case the aim is to prevent unauthorised access to digital material that might cause harm of some kind, whether to national security, the lives of individuals, or commercial or scientific interests.

At the highest level, there are laws in place concerning disclosure of official secrets. In the United Kingdom, for example, the Official Secrets Acts 1911–1989 provide the main legal protection against espionage and the unauthorised disclosure of information. Their scope is information concerned with national security, defence, international relations, criminal activities and the like. The protection of personal data is also enshrined in legislation all over the world, respecting the rights of the individual on the collection and processing of data that is (or can be) linked to them.

Within the world of scientific data, all is not necessarily openly and instantly available. The Common Principles on Data Policy of Research Councils UK recognises that “there are legal, ethical and commercial constraints on release of research data” and acknowledges that “those who undertake Research Council funded work may be entitled to a limited period of privileged use of the data they have collected to enable them to publish the results of their research.” Even when there is no embargo period, user registration may be required simply in order to track who has accessed datasets. Legal constraints include the acquisition and processing of personal data, while ethical issues might arise where release of data might have unwanted consequences: for example, through revealing the location of archaeological sites or of rare animal or plant species.

In the commercial world, it goes without saying that there is much digital material that its owners wish to keep secret since its release would give competitors an advantage.

1.2. Why EUDAT must be prepared

Such considerations will certainly become relevant in EUDAT when some communities will wish to use EUDAT services to store and manage restricted data as well as open data. The EUDAT Collaborative Data Infrastructure (CDI) must be able to handle this situation, and the mechanism chosen is to produce

“consistent guidelines for restricted data access to be adopted in the EUDAT CDI”. In particular, the changing nature of data protection legislation in Europe sees a strengthening of emphasis on giving citizens control of their personal data, requiring explicit consent for their use and placing more formal restrictions on the data controllers and data processors who handle them.

In assessing the impact of data restrictions on EUDAT, and in particular EUDAT’s management of personal data, we draw a distinction between data that is submitted by external parties to the CDI, designated *content data*, and the data that EUDAT itself gathers and processes as part of its operations, *administrative data*, which might be personal in nature, for example about individual users of the CDI. EUDAT needs to be able to handle both of these, though in different ways. In handling administrative data like emails or login names EUDAT service providers will become “data controllers”; for storing content data they are clearly data processors. We address these distinctions in EUDAT’s case in Chapter 7.

1.3. Open data in European research projects

In order to encourage the acceptance of opening access to research data, the Horizon 2020 programme includes a pilot action, the Open Research Data Pilot [5]. Participating projects must develop a Data Management Plan (DMP) specifying which data will be openly accessible. Projects in certain areas, e.g. Future and Emerging Technologies, and Research infrastructures (including e-Infrastructures) are automatically part of the pilot; projects in other areas can also opt in. The Open Research Data Pilot aims to make the research data generated by these projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access.

The Open Research Data Pilot applies to two types of data:

- the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
- other data, including associated metadata, as specified and within the deadlines laid down in a data management plan.

As an exception, participants do not have to ensure open access to specific parts of their research data if the achievement of the action’s main objective would be jeopardized by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access. These reasons might be ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related.

In the area of policy, a useful framework contribution has come from the RECODE project (Policy Recommendations for Open Access to Research Data in Europe) [6] which produced a set of recommendations for facilitating open access to research data targeted at key stakeholders in the scholarly communication ecosystem. RECODE analysed a broad range of legal and ethical challenges: the former included intellectual property and data protection issues, while the latter focused on unintended secondary uses and misappropriation, dual use, violations of privacy and confidentiality, unequal distribution of research results, commercialization, restriction of scientific freedom.

RECODE recommends that policies should accommodate closed data when ethical, copyright, confidentiality, security and similar issues are demonstrably of key concern, and they should take into consideration different disciplinary practices. The policy should be explicit on which data should be open. Open access should be required for research data used to validate scientific claims in publications, while open access to other data produced in the project may be required to be open as well, including associated metadata. While open access to the research data itself may not always be possible, deposit in repositories/data centres with open metadata should be required.

1.4. Scope and approach of this report

This report focuses on personal data and the legislative environment that protects it. The reason for this initial focus is the ubiquity of personal data, even within the scientific research world in which EUDAT plays

a part, and the well-established legislation that protects it, which while it varies from country to country is often underpinned by the same principles and in the case of the European Union by its own directives and regulations—in particular the General Data Protection Regulation (GDPR) which will come into force in 2018.

The target audience for the guidelines in this report are the service providers within the EUDAT Collaborative Data Infrastructure. They will wish to be sure that their policies are well founded, defensible, and coherent with the policies of other service providers—particularly as EUDAT services may be distributed across multiple providers. The aim is to tease out some of the questions, the risks and impacts to make recommendations and pointers.

The policies under consideration here are high-level policies to be adopted by an organisation and that will have some influence on the organisation's behaviour. This is distinct from other more specialised uses of the word “policy” in EUDAT, such as the data management policies relating to B2SAFE.

Ultimately the aim is to produce a consistent and acceptable position on what restricted data may be stored and how; who may access it, when and how; and assurance that indeed it is secure.

This report will be followed by another (D2.8, “Guidelines on Open Access and Restricted Access Data (final)” due in September 2018. The second report will expand the scope to other types of open and restricted data, and introduce ethical issues (rather than purely legislative ones).

1.5. Structure of the report

In considering policies for open and restricted data, EUDAT is not starting from scratch: it already has established policies and guidelines, and **Chapter 2** summarises these. **Chapters 3, 4 and 5** review the legislative frameworks around personal data, considering the current EU directive and future General Data Protection Regulation, and examining several national cases to understand how they reflect the EU requirements and how they might differ. This leads into **Chapter 6** which looks at the nature of data currently and potentially stored within the CDI, and **Chapter 7** which highlights the roles of data controller and data processor and sketches the potential flows of personal data within the CDI. **Chapter 8** outlines a set of recommendations for EUDAT, both overall guidelines for the CDI and issues for the services within it, and **Chapter 9** concludes. **Annexes** provide a number of practical templates for different policies or agreements between parties that relate to handling personal data.

2. ESTABLISHED EUDAT POLICIES FOR OPEN DATA

2.1. Summary of EUDAT open data policy

Since its inception under the Framework 7 EUDAT project, the EUDAT Collaborative Data Infrastructure consortium has believed fundamentally in open access. By open access we mean the free availability of data on the public internet, permitting any user to reproduce and redistribute them for any purpose, and in particular for the purpose of non-commercial research, without financial, legal or technical barriers. The only allowable constraint on reproduction and redistribution should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

One of the prime motivations for the CDI is to create a single domain of registered, well-described, cross-disciplinary data, connecting collections and data centres across Europe and harmonising access to them – harmonising access not just in the technical sense but in the policy sense. In this, EUDAT subscribes to the ideas of intelligent openness as described in the 2012 Royal Society report “Science as an Open Enterprise” [7] and summarised as accessible, useable, assessable and intelligible. To this, we add the desirable property of discoverable. Consequently, all sites joining the CDI under the 2016 Collaboration Agreement are strongly encouraged to adopt open access policies towards their collections in return for the benefits of EUDAT replication and management services.

In the final version of the Sustainability Plan from the Framework 7 EUDAT project [8] a number of common policies for CDI sites were defined which have helped to steer the ongoing development of the common data infrastructure. In particular, EUDAT has adopted the following policies and principles that are directly relevant to its open data agenda.

2.2. Openly discoverable

All data objects deposited in the CDI will be assigned a unique, persistent identifier (a “CDI-assigned PID”) at a suitable level of granularity, and these PIDs will be communicated to the data depositor. EUDAT adopts globally unique Handles to identify digital objects within the CDI. The Handle System [9], the system behind DOI [10] and other well-known identification mechanisms, is administered by the Digital Object Naming Authority (DONA) and is used worldwide. EUDAT works with the European PID Consortium (EPIC) [11] to ensure all data objects registered in the CDI receive a unique, persistent Handle.

CDI sites will ensure that resolution of a CDI-assigned PID results in common, defined and stable behaviour. A CDI-assigned PID should be all a user of CDI services needs to retrieve the associated metadata record and (where authorised) data object from any CDI site.

CDI sites will ensure that data deposited in the CDI are documented with an agreed common metadata baseline to support discovery, citation and provenance. EUDAT strongly encourages adoption of the OpenAIRE application of the DataCite version 3.1 mandatory metadata schema [13]. CDI sites that store data will ensure that all metadata records are discoverable by (at least) the CDI metadata catalogue B2FIND. Metadata collected by EUDAT services are made available using the OAI-PMH publication standard as recommended by the OpenAIRE Guidelines for Data Archives [14].

2.3. Openly accessible

All data in the CDI should, in time, become full open access. Open access is the norm for CDI data. Nevertheless, where necessary or required, embargo periods for original producers are fully supported, on condition that such data become openly accessible when the embargo period expires.

CDI sites will ensure that metadata and (where authorised) data are accessible by users of CDI services over the Internet through common, defined and stable access methods.

CDI sites that store data will be agnostic to any particular data format or set of formats. Users of CDI services will be encouraged to deposit data in open (i.e. non-proprietary) formats appropriate to the content, but no format will be proscribed.

EUDAT's notion of open access follows closely the "Open Definition" [15].

2.4. Openly useable

CDI sites will encourage depositors of data in the CDI to licence their data for open access under the Creative Commons Version 4.0 Attribution licence scheme (CC BY 4.0) [16].

CDI sites that store data will ensure that all rights associated with data objects within the CDI are respected and that access to data objects not openly licensed is subject to appropriate authorisation checks. The reason for recommending adoption of the OpenAIRE Guidelines for metadata is because of their requirement for including a rights statement in the DataCite metadata record for each accessible data object [13].

3. MANAGEMENT OF PERSONAL DATA

This version of EUDAT's Guidelines on Open Access and Restricted Data focuses principally on personal data protection legislation, not least because of the recently adopted EU-wide General Data Protection Regulation (GDPR) [17] on personal data which will come into force within the next two years. It is important that EUDAT service providers are prepared to comply with the changes brought about by the GDPR. Our aim in this report is to prepare service providers in good time for the legislation in its current form, and review ways in which EUDAT already complies and the areas which will need to be updated.

3.1. Current data protection rules in the European Union

Until the GDPR comes into force on 25 May 2018, the main EU wide instrument in the field of data protection law continues to be the 1995 Data Protection Directive (95/46/EC). While there are other EU level instruments that cover personal data¹, we focus on the Data Protection Directive as it sets out the principles and conditions for data processing, as well as the rights of data subjects and the obligations of data controllers and data processors. These principles are significant for EUDAT because they both reflect the current state of affairs for data protection across the EU and, crucially, underpin the GDPR.

Directive 95/46/EC identifies and defines certain key terms:

'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

'processing of personal data' ('processing') shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;

'personal data filing system' ('filing system') shall mean any structured set of personal data which are accessible according to specific criteria, whether centralized, decentralized or dispersed on a functional or geographical basis;

'controller' shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law;

'processor' shall mean a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller;

'third party' shall mean any natural or legal person, public authority, agency or any other body other than the data subject, the controller, the processor and the persons who, under the direct authority of the controller or the processor, are authorized to process the data;

'recipient' shall mean a natural or legal person, public authority, agency or any other body to whom data are disclosed, whether a third party or not; however, authorities which may receive data in the framework of a particular inquiry shall not be regarded as recipients;

'the data subject's consent' shall mean any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed.

¹ For example: Directive 97/66/EC on "the processing of personal data and the protection of privacy in the telecommunications sector;" the ePrivacy Directive 2002/58/EC; and the Council Framework Decision 2008/977/JHA on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters.

It is important to note that EUDAT activities do fall under the processing definition. However, personal data processing is not a large part of the work carried out in our data centres. We therefore have policy in place to ensure compliance with the limited data processing of this type that we do carry out as well as to prepare us for future instances where such data might be held.

Within the framework of this Directive, it is for Member States to determine more precisely the conditions under which the processing of personal data is lawful along a set of principles covering:

Data Quality i.e. ensuring that personal data is processed fairly and lawfully; collected for specified, explicit and legitimate purposes; accurate and, where necessary, kept up to date and permitting identification of data subjects for no longer than is necessary.

Criteria for Making Data Processing Legitimate; i.e. so that personal data be processed only if the data subject has unambiguously given consent.

Special Categories of Processing; Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life. However there are some exceptions to this rule set out in the Directive.

Information to be Given to a Data Subject; i.e. relating to information about the controller, purpose of processing and recipients of the data.

The Data Subject's Right of Access to Data; i.e. that data subject has certain rights to access data collected, to remove data or to be notified of third party disclosure.

Exemptions and Restrictions; i.e. in the interest of national or public security.

The Data Subject's Right to Object; i.e. to the processing of personal data in a way that might result in disclosure to third parties or for marketing purposes.

Confidentiality and Security of Processing; i.e. measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network

Further principles included relate to notification, judicial remedies, liability and sanctions, transfer of personal data to third countries and encouragement of codes of conduct.

3.1.1. Article 29 Working Party clarifications

The Article 29 Working Party of European information commissioners ("WP29") has provided, and continues to provide, valuable clarifications of some of the terms and definitions noted here. In particular, we would highlight the following additional sources of relevance not only to the 1995 Directive but also to the new 2016 GDPR:

- WP29's opinion 4/2007 on the concept of personal data:
 - http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf
- WP29's opinion 15/2011 on the definition of consent:
 - http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf
- WP29's opinion 03/2013 on purpose limitation:
 - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf
- WP29's opinion 06/2014 on the notion of legitimate interests... :
 - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf
- WP29's opinion 05/2014 on Anonymisation Techniques:
 - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

3.2. Data Protection Reform

One of the reasons for reforming data protection under the GDPR is because the current Directive 95/46/EC allows for Member State differences in the way that law is implemented, resulting in administrative costs and legal uncertainty. Using the examples of the UK, Norway and the Netherlands, this variation will be explored in Chapter 4, while in Chapter 5 we will look at the harmonized approach under the GDPR and the implications for EUDAT.

4. THE APPLICATION OF DATA PROTECTION IN EUROPE: THREE CASE STUDIES

EUDAT services currently span 14 European countries, most of which are within the EU but some of which are outside. The application of data protection rules is not currently uniform in these countries, though some common similarities have emerged which we highlight in this chapter. Case studies have been selected from countries both inside and outside the EU, with the UK now falling somewhere in between². We summarize how these countries in which EUDAT operates currently apply data protection legislation, showing that there is already a high level of data protection awareness within EUDAT and best practice that we can take forward in order to comply with the GDPR.

Further details on these case studies can be found in Annex A.

4.1. Data protection in Europe

Despite differences in national data protection laws in Europe, important common principles can be discerned. A fundamental point is the definition of personal data (in the EU Directive): these are data, in fact any information, “relating to an identified or identifiable natural person”. This covers everything that can tell something about you: physically, mentally, economically, culturally or socially.

As a basic level it is not permitted for anyone to process personal data without permission: the consent of the person involved. However, the laws generally stipulate circumstances or conditions which allow processing of personal data even without consent obtained: personal data may be processed for the purposes of the legitimate interests of the person involved or others as well out of compliance with a legal obligation. Processing is defined very broadly; it may include collecting, recording, (re-)organizing, storing as well as disseminating personal data.

For research data there is a so-called “purpose extension” in the 1995 Directive, broadly adopted in national laws, that can (and has) been used to cover the processing of personal data for “historical, statistical or scientific use”: if data were legitimately collected for purpose A then they can be processed – without any extra formalities like the data subjects’ consent – for a compatible purpose B, where historical, statistical and scientific purposes are regarded as compatible purposes. Regarding “sensitive personal data”, however, it seems that a “research exception” by purpose extension is only possible if the research serves “substantial public interest” (art. 8(4) of the Directive). A balance must always be struck between the protection of individual personal data and the public or scientific interest for these data.

4.2. Country case studies

The legislation concerning personal data has been looked into more closely for three countries: Norway, the United Kingdom and the Netherlands.

When summarising the differences as well as the similarities it can be seen very clearly that the essential division of personal data into a “normal” or regular category and a sensitive or special category is present in all the laws. Consequently, the processing of the latter category must satisfy more stringent conditions.

The handling of sensitive personal data is stricter, although this varies slightly from country to country. As a consequence, the processing of medical research data, patient data, can be considered as the strictest category of all, often, as in the Netherlands, regulated by separate laws. In particular, possible archiving of these data is subject to such stringent rules that this is often not possible at all.

Also the concepts of data controller and data processor come back, not surprisingly, in the laws of these three countries. An important point, also for EUDAT, is the handling of personal data by a data controller: this requires a written data processing contract or agreement between controller and processor. This agreement should contain what the data processor is allowed to do with the data. The data processor must provide sufficient security measures. In Norway a risk assessment of the data processing procedures also has

² Final editing of this report took place in July/August 2016, shortly after the UK voted in a public referendum to leave the EU.

to be carried out. In this risk assessment the probability and the consequences of security breaches have to be determined.

In the UK a risk assessment is also recommended best practice when personal data are stored “in the cloud”, as is a written contract, and special care should be paid to the country where the cloud provider could store the data. More generally speaking the transfer of personal data within EU member states is allowed as the EU member states are considered as having the same level of protection. This is not necessarily the case with countries outside the EU, but there is a list of approved third countries. In particular, with the collapse of the Safe Harbour agreement under the Schrems judgement [19] the position of the United States in this respect is at the moment uncertain. The replacement for Safe Harbour, Privacy Shield, has been agreed and implemented between the EU and the US, but the Article 29 Working Party remains concerned [20]. They have adopted a 12 month “wait and see” approach.

Another common element is the obligation to give notice to the national authority on handling personal data, the Data Protection Authority, or, as possible in the Netherlands, the local Data Protection Officers, for example of a university. In Norway a licence is mandatory for handling sensitive personal data.

4.3. Implications for EUDAT

Because EUDAT service providers operate in 14 countries, we have already developed processes to comply with the varying levels of data protection highlighted above. We will now look at the changes to data protection laws across the EU under the GDPR including how nationally developed EUDAT policy might be of benefit in preparing for EU wide legislative change.

5. FUTURE LEGISLATION ON PERSONAL DATA – THE GDPR

The European Union adopted the General Data Protection Regulation (GDPR) on 27 April 2016. It replaces Directive 95/46/EC and aims “to give citizens back control of their personal data and create a high, uniform level of data protection across the EU fit for the digital era.” It will apply from 25 May 2018 with Member States having to transpose it into their national law by 6 May 2018. Those dealing with personal data will have to work quickly to comply. This chapter will look at what the GDPR means for EUDAT users whose rights are being strengthened and clarified and, importantly, what our service providers need to be aware of to comply with these rules.

5.1. Data protection at EU level

Protection of personal data is recognised as a fundamental right under Article 8 of the Charter of Fundamental Rights of the European Union. The General Data Protection Regulation (2016/679/EC) (GDPR) fully harmonizes EU data protection law in order to give individuals greater control over their personal data including in the following ways:

- The right to be forgotten (Article 17);
- Better control over who holds one’s private data (Article 7);
- The right to switch one’s personal data to another service provider (Article 20);
- The right to be informed in clear and plain language (Articles 12, 13, 14);
- The right to know if your data has been hacked (Articles 33 and 34);
- Clear limits on the use of profiling (Article 21);
- Special protection for children (Article 8);
- Privacy as the norm.

5.2. Personal Data and Processing

Two key concepts trigger the data protection regime, namely personal data and processing.

5.2.1. Personal data

Article 4(1) of the Regulation (2016/679/EC) defines personal data as “any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person 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;”.

For EUDAT when processing data, it will be important to be aware of whether a person is identified or identifiable. There are some grey areas here for example an IP address might be a marker of identification to an internet service provider but not to the layman. Secondly, data that may not constitute personal data by itself, might in combination with other data enable identification. Other media such as photographs may also constitute personal data. Given the above it may be safer for EUDAT to assume that data processed in these contexts may constitute personal data and to follow steps to determine whether GDPR rules apply.

5.2.2. Special categories

Under the GDPR, a stricter regime applies to special categories of data, that constitute sensitive data; according to Article 9, these include personal data revealing:

- racial or ethnic origin;
- political opinions;
- religious or philosophical beliefs;
- trade-union membership, and;
- data concerning health or sex life.

EUDAT, as a general-purpose data infrastructure, cannot assume that it will not have to deal with data under these categories. Archive recordings of data subjects discussing unusual lifestyles, for instance, might form part of a social science archive of which EUDAT preserves a replica copy.

5.2.3. Processing

It is the processing of personal data that triggers the application of the GDPR. The definition of processing has been interpreted fairly broadly to mean “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”.

It is safe to regard all EUDAT services as data processing services; analysis of the use of personal data within EUDAT can proceed accordingly.

5.3. Data controller and data processor

The GDPR emphasises that when dealing with personal data, organisations must identify the data controller and data processor. This relates to the obligations and liabilities under the Regulation which are primarily aimed at the data controller with some responsibilities for the processor. The roles of controller and processor are more formal than under the 1995 Directive; under the GDPR their relationship must be covered by a contract or other legal instrument. (This follows the current practice in Norway, for example.)

According to Article 4(7), the controller is “the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law”.

Article 4(8) regards the data processor to be the “natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller”. As follows from the definition of controller (‘alone or jointly’), multiple parties might be considered to be the controller of the data, which are then regarded as “co-controllers”.

When a controller chooses a processor to process data on his behalf, this does not discharge the controller from obligations relating to the security of the data. Article 28(1) GDPR provides that the controller must “use only processors providing sufficient guarantees” relating to the technical and organisational security of the processing of the personal data and he must also ensure that these measures are complied with.

Chapter IV of the GDPR sets out the general obligations of the controller and processor and their implementation of the appropriate technical and organisational measures such as making records of processing activities, security against risk, data breach communication and the designation of a data protection officer. The Regulation also encourages organisations to draw up a code of conduct.

Chapter 7 of this report examines the implications of the controller–processor relationship for EUDAT in greater detail.

5.4. Applicable law

An additional point brought in by the Regulation is the territorial scope of the data protection. According to Article 3, the Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not. The protection also applies to the processing of personal data of data subjects who are in the Union, by a controller or processor not established in the Union under specified conditions.

For EUDAT this will apply to service providers in two countries with data centres outside of the EU: Norway and (in time) the UK. In both of these cases the service providers will need to operate under GDPR standards.

5.5. Principles and obligations

Any processing of personal data must comply with the six main principles provided for by Article 5(1) GDPR and shall be:

- a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');
- b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');
- c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');
- d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy');
- e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');
- f) processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality').

In addition, Articles 13 and 14 provide that the data controller must inform the data subject of, inter alia, his identity, the purposes of the processing of data, the recipients and categories of personal data, and the existence of the data subject's right of access and right to rectify his data.

For EUDAT it will be important to show that the service providers know of the principles and obligations mentioned above and that there are provisions to inform data subjects where Articles 13 and 14 apply. The classification scheme sketched in Chapter 6 is designed to assist service providers in this.

5.6. Legal grounds and consent

Any processing of personal data requires a legal ground. Article 6 provides a limitative list of six grounds that legitimise the processing of personal data:

- a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;
- b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
- c) processing is necessary for compliance with a legal obligation to which the controller is subject;
- d) processing is necessary in order to protect the vital interests of the data subject or of another natural person;
- 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;
- 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.

The key considerations for EUDAT here hinge on consent.

5.6.1. Consent

Article 4 (11) of the GDPR defines ‘consent’ of the data subject as meaning 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 relating to him or her. Conditions for consent are explained in Article 7.

An important consideration for EUDAT is that explicit consent for use of personal data must now be sought, *and must be recorded* so that compliance can be demonstrated at a later date.

6. CLASSIFYING DATA IN THE EUDAT CDI AND RELATED SERVICES

In assessing the impact of the GDPR on EUDAT data services and data stored across multiple sites, we begin by classifying data in the CDI to be either *content* – data uploaded by EUDAT users into a service for storage or other processing – or *administrative* – data collected (perhaps automatically) by a service as part of its normal operation. From here, we use a mind mapping technique to create a tree of possible personal and non-personal data of which the CDI service operators will need to be aware (Figure 1).

6.1. Administrative data

Administrative data are those which are collected directly from users of EUDAT services, including PIDs, email addresses, accounting data, login names, IP addresses, file checksums and other file attributes, and so on. They are under the direct responsibility of – and the direct control of – EUDAT service providers and are, in many ways, the easier type to deal with.

Our primary concern here is with personally identifiable administrative data from EUDAT service users. We assume that an EUDAT service user is not deceased for these purposes (a reasonable assumption). Whether these data are automatically collected (e.g. by logging IP addresses from service requests) or user-supplied (e.g. a login name or email address), key considerations come down to three points.

Legal basis – whatever personal data EUDAT services are collecting, the service provider must be able to demonstrate a legal basis for doing so, and this must be reflected in a clear, public *privacy policy* for each service and each service provider: “we collect the following pieces of personal data for the following reasons”. Certainly none of the personal data that EUDAT service providers might require should fall into the *special categories of personal data* (so-called “sensitive data”) – race, religion, sexuality etc.

Consent – the key to compliance with the GDPR is seeking consent from users for the specific use of their personal data. This can be achieved by obtaining a user’s consent to the privacy policy defining the legal basis for data collection. Another key point is that consent cannot be assumed from use; consent must be explicit (e.g. by a tick-box or active click), and must be recorded in order to demonstrate compliance with the GDPR.

Age – children, meaning a person up to an age of between 13 and 16, depending on jurisdiction, cannot give consent under the GDPR. If a child creates a login account on the B2ACCESS service, the EUDAT service provider must take “reasonable steps” to ensure that they (the service provider) have consent from that child’s *legal guardian* for them to do so lawfully. This raises a significant issue for EUDAT services (see later).

All administrative data collection is also subject to the blanket *personal rights* of a user: the right of access (to any personal data stored); the right to be forgotten (to have personal data deleted); the right to move elsewhere; the right to be informed of hacking.

The challenges of complying with some aspects of this framework will colour the way EUDAT services evolve in the future. This is, of course, in line with the GDPR’s *principle of minimisation*: data services should be designed to collect or process the minimum required set of personal data to deliver their service. Chapter 8 considers some of these possible impacts on EUDAT services.



6.2. Content data

Where content data has no data subject and no personal data records as part of it, or if the data subject in question is deceased, then storage or other processing is not an issue under the GDPR.

Personal content data face similar challenges to administrative data but add a few more of their own. To begin with can we divide content data into two classes: those with a data subject, and those without. Our definition of content data here includes user-supplied metadata of any kind; metadata are data too.

Thus, previously impersonal data (measurements of rainfall, for example) may be rendered personal by the addition of relevant metadata (measurements of rainfall made on 17 June 2016 by Dr John Smith, submitted for archival on 28 July 2016 by Dr Jane Doe). The contributor may add their own name and contact details when they create or upload a content file (e.g. as DataCite *creatorName* and *nameIdentifier*). These personal data will then appear in the metadata associated with the initial impersonal data, and consent for the lawful storage of personal content metadata like these must be obtained in the same way as for administrative metadata noted above.

Content data with a data subject can raise additional issues on top of the three principal issues – legal basis, consent and age – discussed in the previous section. These are: special categories of personal data; and processing for research or statistical purposes. They are related.

Special categories – special categories of personal data cannot easily be processed by general-purpose data services like those in the EUDAT CDI without important specialisation. Where a community data provider (e.g. in the social or medical sciences) has data of this nature and wishes to use the CDI for data storage and management, a specific solution will need to be designed between them and a designated service provider, following the principle of minimisation. Even where service providers have obtained specific consent for the storage or processing of data of this nature, strong arguments can be made against storing or processing them on any Internet-connected system such as the CDI; the impact of leakage or unauthorised access is extremely high.

Research use – the principle of minimisation also applies to the storage of personal data for historical, statistical or research purposes (we use the term “research” to cover these specific adjectives from the GDPR). Storing data for a particular data subject (e.g. a voice recording) for research purposes covers the fact that specific consent for all possible future reuse scenarios for those data cannot be obtained at the time the data are recorded; it does provide a get-out clause against minimisation principles. This means that EUDAT may need to design specific services to store personal data of this nature (e.g. end-to-end encryption with a user’s public key, requiring no shared secret between user and service provider).

It is difficult to offer more specific guidance on these topics here. For use of special categories of personal data in research, the GDPR is non-specific and refers instead to national laws and community guidelines. There is, of course, as yet no case law to which to refer. Also, confidentialisation techniques for different kinds of data quickly become technologically specific³: de-identifying medical image data stored in DICOM formats, for instance, is both a topic of active research and an expert area for sophisticated specialist companies. There is no easy answer, but EUDAT’s open culture of service building with community drive lends itself well to future co-design of suitable services.

6.3. A simple classification scheme

Using this analysis we can define a simple classification scheme for types of data, personal and otherwise, which we can use as a tool in assessing the risks associated with processing certain data within the CDI. We bear in mind here the principal feature of the CDI is as an Internet-connected distributed platform for research data with a mandate for open access.

³ We use the term “confidentialised data” after the 2016 ANDS report *Publishing and sharing sensitive data* [21]: “when data has been modified to remove or reduce the risk that people or subjects of the data can be identified”.

“Risks” mean risks to EUDAT service providers of finding themselves in breach of the GDPR. Further, in the case of consented data we assume that “consent for processing” does *not* include consent for making personal data openly accessible. If a user has, in fact, consented to their data being made openly accessible then encryption, for example, is probably unnecessary, and risks for the management of such data can safely be reassessed as *low*. We draw our levels of risk (no or low-risk, risk, and high risk) from the commentary and recitals of the GDPR itself⁴.

Note that this classification only covers personal data as defined under the GDPR; it does not yet assess data which may be restricted or sensitive for other reasons, nor does it address copyright or other intellectual property rights.

Label	Nature	Examples	Notes
OK	Impersonal data; data relating to deceased persons; unrestricted in GDPR terms.	Actual or simulated measurements of natural phenomena; physical file or object metadata.	Minimal risk.
PC	(Personal; Consented). Personal data with data subject’s explicit consent for processing.	Names, addresses, logins, email addresses, other contact details with processing consent.	Low risk. Care needs to be taken to secure such data in normal operations.
SC	(Special; Consented). Special categories of personal data with data subject’s explicit consent for processing.	Race, religion, trade union activities, health data.	Risk. Must be processed subject to high security: e.g., encryption at rest and in flight; restricted access. May require new services or service customisation.
PR	(Personal; Research). Personal data without consent, collected for research purposes.	Names, addresses, voice recordings, facial images.	Risk. Must be processed subject to minimisation: e.g., confidentialisation or de-identification; encryption at rest and in flight; restricted access. May require new services or service customisation.
SR	(Special; Research). Special categories of personal data without consent, collected for research purpose.	Race, religion, trade union activities, health data, criminal records.	High risk. Must be processed subject to minimisation: e.g., confidentialisation or de-identification; encryption at rest and in flight; restricted access. Will require new services or service customisation.
NO	Unacceptable data.	Special personal data of children; personal data without consent where research use cannot be applied.	Unacceptable risk.

⁴ A good summary of the history of and approach to risk in the GDPR can be found at <https://iapp.org/news/a/risk-and-high-risk-walking-the-gdpr-tightrope/>

7. DATA CONTROLLERS AND DATA PROCESSORS IN EUDAT

In assessing the impact of the GDPR on EUDAT CDI, the roles of the EUDAT data service providers with regard to the data handled must be clearly identified. While the ownership of the data is unequivocally assigned to the data subject, discerning the role of a service provider in term of Data Controller or Data Processor is not always simple, but it is crucial in order to adopt the correct measures to comply with the GDPR.

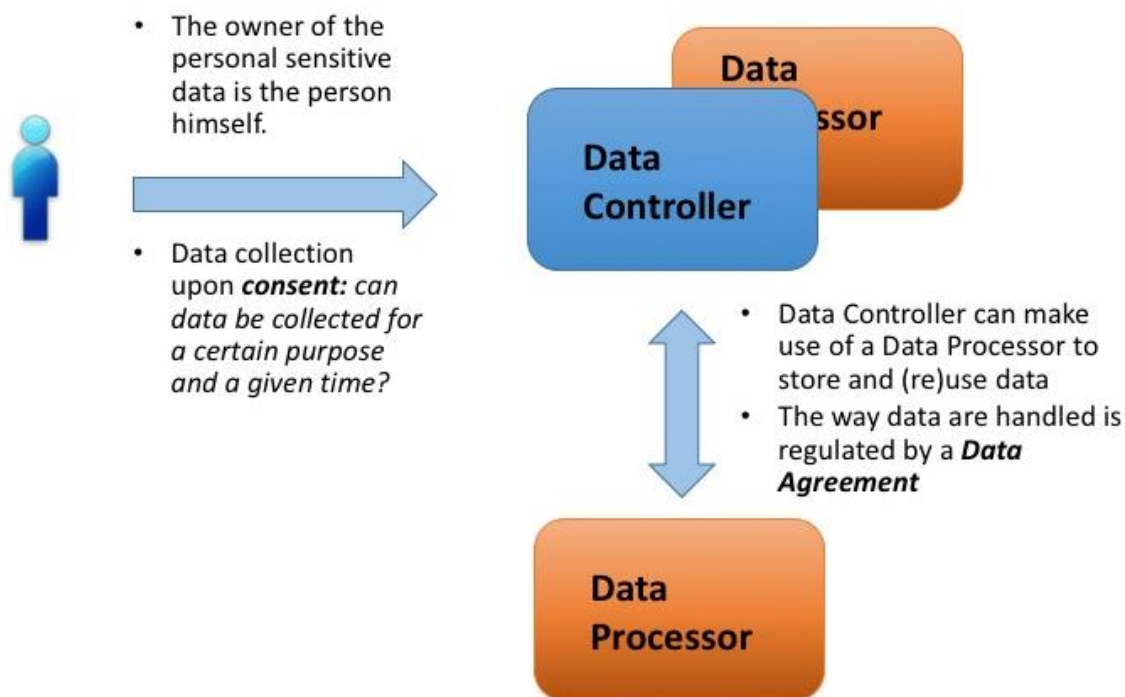


Figure 2. Simplified picture of the relationship between data subject, data controller and data processor.

In the simplified scenario shown in Figure 2, the data controller is the person, organisation, authority or agency who determines the purposes for which, and the manner in which, any personal data are processed. Prior to collecting the data, the data subject must give his or her consent to the collection of a given set of data, for a given purpose and a given time. The data controller might process the data or engage other service to process the data on its behalf. The data processor is the person, organisation, authority or agency who processes the data on behalf of the data controller. The data agreement signed by both parties (data controller and data processor) ensures that the data are processed according to certain standard and states the relative roles and responsibilities of the two parties. An example of a data agreement is given in Annex C. Often the data controller and data processor are the same institution, and therefore a data agreement is not needed. Furthermore a service provider that acts as data controller for a given set of data might be the data processor for other data.

The main significance of the GDPR for EUDAT at an organisational level is the requirement to formalise in law agreements between data controllers and data processors. Under the 1995 Directive, controllers and processors had to have a data agreement; under the GDPR this agreement must now be in the form of a contract or equivalent legal undertaking.

As noted in the general case above, service providers in the EUDAT CDI will take one, or both, of these roles. Figure 3 indicates potential flows of personal data (both administrative and content-specific) between principal EUDAT service providers, and the likely roles this will require from each.

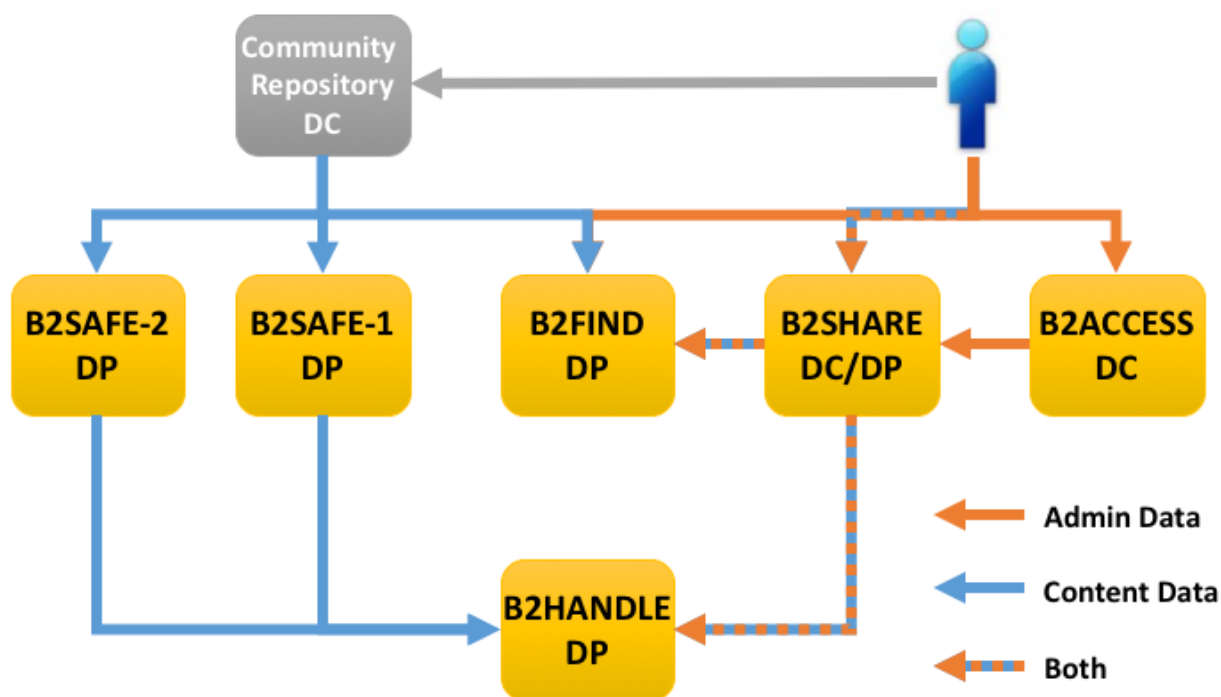


Figure 3. Potential flows of personal data between EUDAT service providers (arrows indicate "from-to"). Orange denotes administrative data, as collected directly by EUDAT services; blue indicates content data, which could include personal data of either the illustrated user or another data subject. Gold boxes indicate EUDAT services (properly, service providers). The grey box illustrates a community data repository making use of EUDAT services. Service providers are labelled DP if they take the role of data processor, DC if data controller.

In general, it is safe to assume that non-user-facing services like B2SAFE and B2HANDLE will operate solely as data processors. User-facing services that require user authentication – B2ACCESS and B2SHARE – will almost certainly operate in the role of data controller through their handling of user identity data. B2FIND, while user-facing, does not require authentication and can be used anonymously, although administrative data such as IP addresses can, in principal, be collected for statistical purposes. However, since B2FIND does not transmit non-confidentialised data onwards for further processing, we do not categorise it as a data controller.

The requirement for formal contracts between these service providers needs to be included in the future legal framework of the EUDAT CDI.

8. RECOMMENDATIONS

The EUDAT CDI, built as it is from existing data repositories, already complies with national data protection legislation. In order to comply with the GDPR some areas will require renewed attention. These relate to **consent**, **age**, responsibility (**data processor/controller relationship**) and **content audit**. We summarise the general findings below, followed by an orthogonal view on a service-by-service basis.

8.1. General guidance for the EUDAT CDI and Collaboration

Overall, to ensure compliance with existing laws and with the future requirements of the GDPR, EUDAT service providers will have to make informed decisions on the following:

For protection of EUDAT users and consent:

1. EUDAT service providers *must* ensure that service users are made aware of their rights in a clear understandable format. This can be done through a standard *Privacy Policy* (cf. Annex B).
2. EUDAT services *must* have a method to gather and record consent from users – “freely given, specific, informed and unambiguous” – over the use of EUDAT services, and in particular acceptance of the Privacy Policy – the equivalent of “I accept cookies” and/or “I have read and agree to the terms and conditions”.
3. EUDAT service providers *must* ensure that there are procedures on how to handle requests about personal data. An appropriate contact should be noted in the Privacy Policy, and clear lines of action should be created.
4. EUDAT service providers *must* ensure privacy policy notices are reviewed regularly.
5. EUDAT service providers *must* put in place procedures to detect, report and investigate personal data breaches. This could be added to the responsibilities of the EUDAT security officer, or form part of the new role of data protection officer.

On processing and identifying processors and controllers:

6. EUDAT service providers *must* be aware of the change in EU legislation and their responsibilities as data processors, data controllers, or both. This report serves as a foundation document here.
7. EUDAT service providers in the role of data controllers *must* arrange contracts or other legal forms with any and all data processors to whom they transmit personal data.
8. EUDAT service providers *must* ensure that there are measures in place to uphold the principles and obligations of the GDPR in Article 5, vis: ‘lawfulness, fairness and transparency’; ‘purpose limitation’; ‘data minimisation’; ‘accuracy’; ‘storage limitation’; and ‘integrity and confidentiality’.
9. EUDAT service designers *must* be able to demonstrate that there is *data protection by design*, ideally through impact assessments.
10. EUDAT management *must* ensure that these GDPR rules apply to EUDAT data centres outside EU Member States.

On data held (both administrative and content data):

11. EUDAT service providers *must* ensure that there is a record and understanding of personal data held, and that this is in line with the stated Privacy Policy.
12. EUDAT service providers *must* ensure that there is a mechanism in place to identify personal data including the special categories of personal data.
13. EUDAT management *should* instigate a documented review of the various types of data processing EUDAT carries out and the legal basis for carrying it out.

Protecting children:

14. EUDAT service providers *must* ensure that systems are in place to verify individuals' ages and to gather parental or guardian consent for the data processing activity if required. A paragraph in the *Disclaimer* section of the Privacy Policy is necessary but possibly insufficient.

Other:

15. EUDAT *should* appoint a Data Protection Officer to monitor compliance with these policies.

8.2. Considerations and issues for specific EUDAT Services

The GDPR will impact generally the way EUDAT CDI services are designed, implemented and operated. In this section we summarise the probable impacts on each of the B2- services specifically, and offer some initial guidelines on future service design. Note that these are initial assessments and guidelines, not points of law; the actual impact of a number of these issues may not be known until the GDPR has been tested in the courts – hopefully by entities outwith the EUDAT CDI Partnership!

8.2.1. B2ACCESS

Consent for user identity management: B2ACCESS, the common authentication service, is likely to see the largest impact, given that it manages personal data as a matter of design. Consent must be obtained from each and every user of B2ACCESS, and must be recorded. It must be made clear what personal data are collected and for what purpose; this must be explained in a clear *Privacy Policy*.

Age: As noted, children are unable to consent to use of their personal data. While it's unlikely that a child would be interested in creating either a federated or indeed simple identity on B2ACCESS, nevertheless some form of age verification for the service may be inescapable. How this could or should be done is unclear. As a guard, a disclaimer could be included in the Privacy Policy stating that EUDAT services are not intended for use by children.

Consideration of identity federation: B2ACCESS can combine user identities from multiple sources. This almost certainly brings it under the scrutiny of the minimisation principle: a data breach of such multiple identities is more serious than that of a single identity, and so thought must be given to how user identities are recorded in the service databases. A security assessment should be carried out to help understand the risks.

Data controller – data processor: B2ACCESS both federates user identities from external providers and allows users directly to create an “EUDAT” identity. This almost certainly places the B2ACCESS service provider in the role of both data processor to the identity providers' data controllers, and data controller to other EUDAT sites making use of the service for single sign-on. These controller-processor relationships will need to be codified in contracts or other legal instruments.

8.2.2. B2SHARE, B2DROP

Consent: B2SHARE and B2DROP offer user-facing web-sites requiring authentication, either via B2ACCESS or directly. In either case, consent for use of personal data (logins) must be obtained and recorded. A clear *Privacy Policy* must be in place for each site (ideally a common policy with B2ACCESS).

Age: As noted in B2ACCESS, age verification for consent is a potential issue. Given that B2SHARE is designed for “long-tail” science and “citizen science”, and one potential group of citizen scientists is classes of schoolchildren, the likelihood of this occurring is perhaps higher than we might expect. Devolving all authentication issues to B2ACCESS is one way to collect the problems – and solutions – in one place.

Content (PC, PR, SC, SR): Users can upload anything into B2SHARE; some of this could be personal data, potentially even special categories of personal data. There is no policing of content, nor is there likely to be in the future. B2SHARE should publish a clear *Disclaimer* to this effect as part of the Privacy Policy (cf. Annex B).

8.2.3. B2FIND

Content (PC, PR, SC, SR): As with B2SHARE, the B2FIND service provider has little or no control over personal data that may find their way into the service. Consequently, a clear *Disclaimer*, possibly common with other services like B2SHARE, needs to be in place on the site. The current service provider, based in Germany, provides a disclaimer in German, noting that the ruling law is that of the Bundesrepublik Deutschland; a non-binding English translation should be provided alongside.

8.2.4. B2SAFE, B2STAGE

Content (PC, PR, SC, SR): As a multi-site service, the operation of B2SAFE between, say, a community site and one or more data centres is or will be covered by service level agreements (SLAs). Where data to be replicated or otherwise processed might be personal data, these agreements will need to reflect appropriate minimisation principles, probably on a case-by-case basis. Any data transfer mechanisms should be secure (ssh; sftp; https); data may need to be stored encrypted. There is almost certainly no one-size-fits-all solution here.

Data controller – data processor: If a community site has collections which include personal data (e.g. interviews with data subjects) then they automatically take the role of data controller for those data. An EUDAT CDI site receiving those data by onward transmission through B2SAFE, for instance, automatically takes the role of data processor; the B2SAFE agreement between controller and processor (between community site and data centre) must now be codified in a contract or other legal instrument; a sub-legal agreement is no longer sufficient.

8.2.5. B2HANDLE

Data controller - data processor: B2HANDLE is currently used behind the scenes by a number of other services for PID creation and has no user-facing interface. However, *if* the B2HANDLE PID schema records any personal data (e.g. *creatorName* from the DataCite schema) *then* this would potentially place the B2HANDLE service provider in the role of data processor to a PID-requesting data controller, requiring a contractual agreement between the two parties.

Consent: Again, *if* the B2HANDLE PID schema records any personal data, explicit consent must be obtained from the relevant data creator. The request for consent must make clear that this personal data will propagate into the global Handle System.

8.2.6. Future Services

The minimisation principle must be applied to all future service design, not only as necessary to comply with the law but also as a defensive mechanism. Consider, for example, the propagation of (*creatorName*, *nameIdentifier*) into 1,000 or more Handles and the impact of a subsequent invocation by that creator of their right to be forgotten.

9. CONCLUSION AND PLANS FOR FURTHER WORK

EUDAT deals with open data, some of which is covered by data protection regulation. EUDAT has already developed national solutions under existing legislation but some updates will be required to comply with the EU wide GDPR as it comes into force in 2018. EUDAT will be drawing on established best practice as a source of recommendation for our service providers. Key to our approach is ensuring that there is the awareness of our obligations following the principles of the GDPR, knowing where responsibility lies and upholding user rights including that there is explicit consent given for the processing of personal data. This deliverable has therefore identified the areas in which EUDAT already complies with the GDPR and set out further recommendations for our service providers. It will undergo expert legal review before further sections are to be added covering ethical issues in D2.8, “Guidelines on Open Access and Restricted Access Data (final)” due in September 2018.

10. REFERENCES

- [1] G8 *Open Data Charter*, 2013. <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>
- [2] European Commission, *Open data: An engine for innovation, growth and transparent governance*, 2011, COM/2011/0882 final. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52011DC0882>
- [3] RCUK *Common Principles on Data Policy*, 2011 (revised 2015). <http://www.rcuk.ac.uk/research/datapolicy/>
- [4] Research Data Alliance website, <http://www.rd-alliance.org>
- [5] OpenAIRE website on the EU Open Data Pilot. <https://www.openaire.eu/opendatapilot>
- [6] RECODE project website. <http://recodeproject.eu/>
- [7] The Royal Society, *Science as an Open Enterprise*, The Royal Society Science Policy Centre report 02/12, ISBN: 978-0-85403-962-3.
- [8] R. Baxter et al, *EUDAT Sustainability Plan (final)*, EUDAT-DEL-WP2-D2.1.3 v1.0, May 2015.
- [9] Handle System website. <http://www.handle.net/>
- [10] Digital Object Identifiers website. <http://www.doi.org/>
- [11] European PID Consortium website. <http://www.pidconsortium.eu/>
- [12] The DataCite Consortium, *DataCite Metadata Schema for the Publication and Citation of Research Data*, version 3.1, October 2014, doi:10.5438/0010
- [13] OpenAIRE, *Guidelines for DataCite*, 2015. https://guidelines.openaire.eu/en/latest/data/use_of_datacite.html
- [14] OpenAIRE, *Guidelines for use of OAI-PMH*, 2015. https://guidelines.openaire.eu/en/latest/data/use_of_oai_pmh.html
- [15] *Open Definition*, version 2.1. <http://opendefinition.org/od/2.1/en/>
- [16] Creative Commons website. <http://creativecommons.org/>
- [17] EU, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*, 2016. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>
- [18] EU, *Directive 95/46/EC, Protection of personal data*, <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV:l14012>
- [19] ECJ, *The Court of Justice declares that the Commission's US Safe Harbour Decision is invalid*, October 2015, <http://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150117en.pdf>
- [20] Article 29 Working Party, *Statement on the decision of the European Commission on the EU-U.S. Privacy Shield*, 2016. http://ec.europa.eu/justice/data-protection/article-29/press-material/press-release/art29_press_material/2016/20160726_wp29_wp_statement_eu_us_privacy_shield_en.pdf
- [21] ANDS, *Publishing and sharing sensitive data*, 2016. http://www.ands.org.au/__data/assets/pdf_file/0010/489187/Sensitive-Data-Guide-2016.pdf

ANNEX A. COUNTRY CASE STUDIES (UK, NETHERLANDS, NORWAY)

A.1. The United Kingdom – the Data Protection Act 1998

The Data Protection Act 1998 is founded on eight principles, of which the first is that personal data shall be processed “fairly and lawfully”. From the principles there follow various restrictions on what may be done with the data and how it must be handled, as well as the rights of the data subject (the individual who is the subject of personal data).

The Act defines *personal data* as follows:

“Personal data means data which relate to a living individual who can be identified – (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.”

There is a further category of *sensitive personal data*:

“Sensitive personal data means personal data consisting of information as to (a) the racial or ethnic origin of the data subject, (b) his political opinions, (c) his religious beliefs or other beliefs of a similar nature, (d) whether he is a member of a trade union, (e) his physical or mental health or condition, (f) his sexual life, (g) the commission or alleged commission by him of any offence, or (h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.”

The processing of sensitive personal data must satisfy more stringent conditions.

Anonymised or aggregated data is not regulated by the Act, providing the anonymisation or aggregation has not been done in a reversible way.

Processing of data means “obtaining, recording or holding the information or data” or carrying out operations on it.

A further key distinction is between *data controller* and *data processor*.

“Data controller means ... a person who (either alone or jointly or in common with other persons) determines the purposes for which and the manner in which any personal data are, or are to be, processed.”

“Data processor, in relation to personal data, means any person (other than an employee of the data controller) who processes the data on behalf of the data controller.”

Note that in this context a “person” may be (and usually will be) an organisation. It is the data controllers who are responsible for ensuring that their processing complies with the Act. Data processors are not directly subject to the Act except as regards data they hold for their own administrative purposes.

Rights of the data subject include the right of access to a copy of the information held on them, as well as rights to object to certain uses of the data, to correct inaccuracies, etc.

Requirements on handling personal data include adequate security and proper disposal. The Act requires that appropriate technical and organisational measures are taken against unauthorised or unlawful processing of personal data and against accidental loss, destruction or damage.

Where a data processor is engaged by the data controller, the data processor must provide sufficient guarantees about its security measures, reasonable steps must be taken by the data controller to check that those security measures are being put into practice, and there must be a written contract setting out what the data processor is allowed to do with the personal data.

The Act says that:

“Personal data shall not be transferred to a country or territory outside the EEA unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.”

The Information Commissioner’s Office has published *Guidance on the use of cloud computing*⁵, noting that “By processing data in the cloud an organisation may encounter risks to data protection that they were previously unaware of.” An organisation choosing a cloud service for data storage will continue to be the data controller. The advice includes assessing the risks, selecting which data to move to the cloud, monitoring performance, establishing a written contract, protecting data (possibly with encryption), and carefully controlling access—and of course being careful about where (in which countries) the cloud provider will store the data.

A.2. Norway – the national legal framework on restricted data

According to the Norwegian Personal Data Act (PDA)⁶ a data controller (an institution) is obligated to assess several issues before starting the data processing:

- i. Provide an exhaustive list of the personal data that are going to be processed (personal data: any information and assessments that may be linked to a natural person) e.g. name, phone number, email address, user name, IP address and so on. (Remember any information that can be linked to a specific person.)

Also, one can only process data that "are adequate, relevant and not excessive in relation to the purpose of the processing," ref. section 11 d) of the personal data act. This means a research project can never process more data than necessary to obtain the purpose of the project.

Note! Data processing is, according to the PDA, any use of personal data such as collection, recording, alignment, storage and disclosure or a combination of such uses.

- ii. Document the purpose of the processing of the data. Data collected for one purpose may never be for purposes that are incompatible with the original purpose of the collection. (PDA Section 11 b).
- iii. Conditions for the processing of personal data. Personal data may only be processed if the data subject has consented thereto, or there is statutory authority for such processing. (PDA Section 8)

When processing sensitive data (see below) one has to fulfil the conditions stated in the PDA section 9 in addition to the condition in section 8.

Sensitive personal data: PDA section 2 8): any information relating to

- a) racial or ethnic origin, or political opinions, philosophical or religious beliefs,
- b) the fact that a person has been suspected of, charged with, indicted for or convicted of a criminal act,
- c) health,
- d) sex life,
- e) trade-union membership.

If the Data controller is using a data processor one needs to enter a data processor agreement. Also, one has to do a risk assessment of the data processing process. The data controller is responsible for making this happen. No processor may process personal data in any way other than that which is agreed in writing with the controller. Nor may the data be turned over to another person for storage or manipulation without such agreement.

⁵ <https://ico.org.uk/media/about-the-ico/documents/1042330/cloud-computing-guidance-for-organisations.pdf>

⁶ [https://datatilsynet.no/English/Regulations/Personal-Data-Act-/](https://datatilsynet.no/English/Regulations/Personal-Data-Act/)

iv. Risk assessment

The data controller shall carry out a risk assessment of the planned processing of personal data (the whole process) in order to determine the probability and consequences of breaches of security. A new risk assessment shall be carried out in the event of changes of significance for information security. (if it is a research project – it is the use of personal data in the specific project: how data is collected, stored, and so on).

[*Remark:* The University of Oslo (UiO) carries out risk assessments of the process used to process data by looking at the data processed (the list of personal information processed) - the more sensitive the data is the higher the requirements are to the information security provided. E.g. for health data UiO require the use of the specialised TSD service to store and process data. UiO do not have a blanket ban on the use of cloud services. Cloud services are assessed in the same way as other services: how it is built, what kind of security measures have to be taken, where data are stored, what kind of control UiO will have over the processing of the data and so on. This then informs the decision about what kind of data can be processed in the cloud service in question.]

v. Information to the data subject

When personal data is collected from the data subject himself, the controller shall on his own initiative inform the data subject about the project, why and what the data is collected for and so on. PDA section 19.

vi. Obligation to give notification and to obtain a license

When starting to process data, the data controller must consider if it is necessary, PDA section 31, to notify the Data Inspectorate or apply for a license to use the data from the Data Inspectorate, PDA section 33. The latter is required when processing sensitive personal data.

vii. Prohibition against storing unnecessary personal data

The controller shall not store personal data longer than necessary to carry out the purpose of the processing. PDA section 28

If the personal data shall not thereafter be stored in pursuance of the Norwegian Archives Act or other legislation, it has to be erased.

A.3. The Netherlands – national framework on restricted data

Acts on Privacy Issues

1. **Personal Data Protection Act** (Wet Bescherming Persoonsgegevens WBP)

Based on a European Directive. General law regulating the protection of personal data. Contains exceptions to the strict rules for preserving/using personal data for research, statistical or historical purposes. Will be replaced in the near future by the GDPR – General Data Protection Regulation, an EU law, uniform for all EU member-states. Finally formally approved on April 14th 2016; implementation however will take another two years from now on as complimentary national legislation is needed. The latter is particularly relevant for research.

2. **Obligation to Report Data Leakage Act** (Wet Meldplicht Datalekken)

Addition to the Personal Data Protection Act on making public personal data. Contains sanctions in the form of penalties.

3. **Medical Treatment Agreement Act** (Wet Geneeskundige Behandelings Overeenkomst WGBO)

Regulates the use and preservation of personal (patient) data in and for medical research.

4. **Scientific Medical Research with People Act** (Wet Medisch-Wetenschappelijk Onderzoek met Mensen WMO)

Regulates scientific research in the medical field, in particularly how to handle personal data. Makes ethical reviews compulsory for all medical research projects.

Codes Of Conduct

1. **Code of Practice for the use of personal data in scientific and scholarly research** (Gedragcode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek VSNU)

This code is based on the Personal Data Protection Act and prescribes how to handle personal data in the practical research context. Its validity has formally expired; legal status unclear.

2. **Code of Conduct for Medical Research** (Gedragcode Gezondheidsonderzoek)

This code is based on the Medical Treatment Agreement Act and regulates in more detail how researchers should handle medical personal data. Its validity has formally expired; legal status unclear.

There is no national standard set of data sensitivity categories in the Netherlands. Not even in the research world. The National Data Protection Authority of the Netherlands has created a set of five categories (ranging from zero risk to high risk) years ago, but does not endorse this set formally anymore. It is however certainly used in the field.

Using Personal Data for Research in the Netherlands

The Personal Data Protection Act (Wet Bescherming Persoonsgegevens WBP) makes an important distinction between regular and special personal data.

Regular personal data

Special personal data are clearly defined in the law, this is to a lesser degree the case with the regular personal data. Anyway it seems rather safe to assume that the name, title, gender, date of birth, address, telephone number, e-mail address and car registration number are considered to be regular personal data. Regular personal data may, in principle only be processed and/or made available (published) if this is in line with the aim for which they have been gathered. There are several exceptions to this. Public figures like politicians are less protected by the law in their public life. For research there is a general exception for using or preserving personal data for other purposes than the original aim they have been collected for. That is to say; they can be used for research, statistical or historical purposes.

Special personal data

Special personal data are data on someone's health (medical records), religion, political conviction, race, sexual orientation and personal identification number. The law is far more strict here. Processing and preserving these data is prohibited. There is again the exception for research, statistical or historical purposes, but this is applied much stricter. This is now only allowed when several conditions are met: the research should be in the general interest, processing the data is necessary for carrying the research, asking for consent of the people involved would be a disproportionate effort and no disproportionate damage to the person's privacy may be done.

Personal data from medical research

For personal data used in medical research, patient data, there are separate laws. On the whole these are again far stricter than for the "non-medical" personal data. Generally speaking, patient data may only be processed if the patient has given explicit "informed consent" and if a medical-ethical commission has given the green light for the whole research project. Even then preserving these personal data is mostly only allowed for a limited period. If medical research data are fully anonymised, it is possible to preserve them and use them for further research. If that is not case, they can only, under certain conditions, be kept for a

maximum of five years, in practice. And it is problematical whether they can be used in research projects other than the original one.

ANNEX B. TEMPLATE FOR PRIVACY POLICY AND DISCLAIMER

The following text can be used by service providers as a basis for a common privacy policy and disclaimer. It is derived from existing EUDAT statements (notably B2FIND and B2SHARE) and has been strengthened where necessary to align with the new requirements of the GDPR.

Privacy Policy

The service ... (hereafter “the Service”) is operated by ... (hereafter “we”, “us”, “the Service Provider”). To make use of the Service you must consent to processing by Us of your personal data as described below.

The careful and lawful handling of your data is important to us. This privacy policy explains what personal data the Service collects, for what purposes and how they are processed.

Please read the privacy policy thoroughly before you provide us with any personal data, or use the Service in any other way. If you do not consent to these terms, you may not use the Service. The following data privacy provisions for the Service are the current version.

Information We Collect

The Service creates technical information about service usage and status. This information (in the form of server log files) is automatically and manually monitored and processed. The Service Provider may collect information about the Service by analysing types and kinds of data stored in the Service and how these data are accessed.

The Service Provider may also combine information with other EUDAT service providers. Information achieved by collecting and analysing technical information and service usage information will only be used to check that the relevant service Terms of Use are being followed and for service development purposes.

When you use the Service the following data may be collected automatically to create anonymous statistics; for the purpose of monitoring data protection, data security; and to ensure the proper operation of our data processing systems:

- the IP address of your computer;
- the date and time of visit;
- the operating system and browser on your computer;
- the amount of data transmitted;
- the internet address of the website from which you have accessed this site.

When you use the Service the following data may also be collected if you choose to supply it in content you upload to the Service:

- contact information, such as your name, email address;
- your username and password;
- other personal information in content you provide to the Service;
- institutional or organisational affiliations.

Children's Privacy

Our Service is not aimed at children under the age of [13..16] and we do not knowingly collect personal information from children under the age of [13..16] through the Service. If we become aware that we have

inadvertently received personal information through the Service from a child under the age of [13..16], we will delete the information from our records.

Other Processing and Use of Personal Data

Your information will not be disclosed to third parties and will be used exclusively for the implementation of the Service and the processing of your request.

Your Rights

You have the right to withdraw your consent to the collection, processing and use of personal data at any time with effect for the future.

You have the right to receive free information about the personal data stored about you. On request we will inform you in writing in accordance with applicable law whether and what personal data is stored by us.

You have the right to update or correct inaccuracies in the personal data we hold about you.

You have the right to delete your personal data from the Service (your “right to be forgotten”).

You have the right to know if your personal data have been hacked or compromised.

To exercise these rights, please send requests to: [.....]

Changes to the Contents of this Privacy Policy

The Service Provider reserves the right to change the content of this Privacy Policy from time to time in accordance with legal data protection regulations. Changes to this Privacy Policy will become effective when those changes are posted to the Service.

Disclaimer

Service Website

The Service Provider endeavours to provide accurate and up-to-date information on the Service website. However, errors cannot be ruled out, and the Service Provider accepts no responsibility for the correctness or completeness of the information provided.

The Service Provider reserves the right to change the website in part or in whole without prior notice.

The Service website includes links to external sites. The Service Provider accepts no liability or responsibility for the accuracy, completeness or legality of the content of any linked external websites.

Third-Party Content

The Service Provider is not under any obligation to monitor third-party content for completeness, accuracy, or compliance with binding rules. This holds for all data accessible through the Service, regardless of whether they are stored on servers owned by the Service Provider or others. Accordingly, the Service Provider accepts no liability or responsibility for the completeness, accuracy or legality of third-party content.

Unauthorised Access by Third Parties

Despite all best efforts, no method of transmission over the Internet and no method of electronic storage can be guaranteed to be absolutely secure. The Service Provider accepts no liability for unauthorised access by third parties or for any possible transmission of computer viruses, Trojan horses or other malicious programs. The user is responsible for making arrangements to protect their own computer from such malicious programs, in particular by installing antivirus software and using the latest antivirus definitions.

If the Service Provider learns of a security breach, affected users will be notified immediately by email and via the Service website, so that they can take appropriate protective steps.

ANNEX C. TEMPLATE AGREEMENT FOR DATA CONTROLLER-DATA PROCESSOR

The following text can be used as a template for a legal agreement between the Data Controller and the Data Processor, where the latter processes personal data on behalf of the former. The draft is an adaptation of what it is presently in use in Norway. The legal formalisation of the agreement between controller and processor now present in the GDPR stems, at least prima facie, from the Norwegian model, hence our suggestion of this as a starting point. In the agreement the articles/act that the partners need to comply with have to be specified and normally consist of the national Personal Data Act and/or national Personal Health Data Act, but in the near future also the GDPR articles/acts might be invoked.

AGREEMENT ON THE PROCESSING OF PERSONAL DATA

Storage and Processing of research data in the *<Insert the name of the Data Processor>*

The text in blue italics must be removed and replaced with relevant text, in some cases by selecting one of several alternatives.

1. Parties to the Agreement

1.1 Parties

The Agreement is entered into between the party responsible for the data processing: *<Insert the name of the project/institution>* (Org. no.) (hereafter referred to as the Data Controller) and the party that processes the data: *<Insert the name of the service/Institution>* (Org. no.) (hereafter referred to as the Data Processor).

1.2 Contact persons

Contact person for the Data Controller: *<name, contact information, role>*,.....

Contact person for the Data Processor: *<name, contact information, role>*,.....

2. Purpose of the Agreement

The Data Processor offers storage services for researchers who conduct research on person-sensitive data, including health data.

The purpose of the Agreement is to regulate rights and duties pursuant to:

- *<list of the articles/act/regulations of the GDPR regarding the personal data and personal health data>*

The Agreement regulates the Data Processor's processing and securing of personal data and health data that have been made available by the Data Controller. It must be clearly stated whether the Data Processor is permitted to surrender data to other parties for storage, processing or other use.

The purpose of the processing shall not be changed by either of the parties without a new agreement being signed.

3. The parties' area of responsibility pursuant to *<list of the articles /act/ regulations of the GDPR regarding the personal data and personal health data>*

The Data Controller is to be deemed the unit responsible for the data processing pursuant to *<list of the articles of the GDPR to which the agreement aims at complying...>*

The Data Controller is responsible for ensuring the fulfilment of the requirements laid down in the *<list of the articles of the GDPR to which the agreement aims at complying...>*, including those relating to security. This entails the Data Controller being charged with ensuring that the requirements relating to the storage and use of health data and sensitive personal data are complied with by the Data Processor.

The Data Processor can only process health data and personal data that have been made available by the Data Controller in accordance with this Agreement. Any other use of health data and personal data shall be agreed with the Data Controller in advance and in writing.

The Data Processor shall ensure that health data and personal data made available by the Data Controller are kept separate from its own and others' data and services.

4. Description of the purpose of the use of the Data Processor

The Data Processor can only process personal data in accordance with the purposes that have been specified by the Data Controller and pursuant to the terms stated in this Agreement.

<This point MUST be filled in, and it must be stated clearly and precisely what the data are to be used for. Any link with other data sets must be approved by the Data Controller. An exception from this is when the links are made anonymous.>

State what the data is to be used for:>

5. Specification of the data that is to be processed

<Must be filled in, and must indicate the type of data that is to be processed, and whether these data are directly identifiable or have been made unidentifiable (i.e. whether the data appear as anonymous, but where it is actually possible to go back and find out who the data/information concerns).>

If the Data Controller finds it necessary to change the data that are to be processed, or to add a new type of data to those that are to be processed, he is under the obligation to make a new security assessment. If a material change is involved, the change cannot take place without a new data processor agreement being signed.

6. Requirements regarding data security

Pursuant to the provisions stated in the *<list of the articles of the GDPR to which the agreement aims at complying...>*, both parties shall at all times meet the requirements regarding data security and internal control, as well as those relating to access control.

The Data Processor shall ensure that all processing of health data and personal data encompassed by this Agreement is carried out in accordance with the acceptable level of risk defined by the Data Controller. As part of this the Data Processor shall submit risk assessments of its own security.

With regard to security, the Data Processor is required to have defined its objectives, strategy, organization and responsibility in accordance with the *<list of the articles of the GDPR to which the agreement aims at complying...>*, and is required to ensure that these are followed up by the necessary internal control system.

Any breach of security or any suspected breach of security shall immediately be reported to the Data Controller.

The Data Processor shall have clear procedures for logging errors and nonconformities in systems that are used to handle health data and personal data and that are included in this Agreement. If such errors or nonconformities are detected, the Data Processor shall notify the Data Controller of this as soon as possible and at the latest within 24 hours (48 hours if the incident arises at the weekend or on a public holiday). In such an event the Data Processor shall immediately take steps to minimize possible damage to the Data Controller.

The Data Controller can at any time demand documentation from the Data Processor as reassurance that the Data Processor is complying with all relevant requirements concerning data security stated in the *<list of the articles of the GDPR to which the agreement aims at complying...>*. The Data Controller can request access to the Data Processor's reports etc. on periodic audits of its procedures and routines.

The Data Processor shall be able to demonstrate good routines concerning data security, including in particular technical security, access control and physical security.

The Data Controller is responsible for adequate security at the units that are used for remote access to the Data Processor. With regard to updating and virus control this will in many cases mean that the units must be in the Data Controller's operating regime or that of parties closely related to the Data Controller.

7. The Data Controller's right to access, inspection and testing

The Data Controller shall have the right to access the solution and to verify how it is secured. In this context 'access' means documentation, interviews, meetings and any other forms of verification that may be appropriate. The Data Processor accepts that access can be exercised by the Data Controller or by the third party the Data Controller may select to carry this out as long as the access extends only to the area designated to the processing of the Data Controller's data. The right to access applies to all technical, organizational and administrative aspects that are relevant for security in the services that are delivered to the Data Controller.

The Data Processor is obliged at four weeks' notice to surrender security documentation relevant for the Data Controller, or otherwise to ensure access to such documentation.

If the Data Controller makes use of the right to access, and nonconformities are detected in the security of the Data Processor's systems, the Data Processor shall remedy the nonconformity as quickly as possible. The Data Processor shall give a written description of the remedial measures and the plan for implementing them.

8. Confidentiality obligation

The parties shall observe professional secrecy on all confidential information, people's personal circumstances, security and business matters, and information that may cause harm to one of the parties or that may be utilized by a third party.

The confidentiality obligation applies to the parties' employees and to others who act on behalf of the parties in connection with the implementation of the contract. All employees must have signed a non-disclosure declaration.

The parties are under the obligation to take the necessary precautions to ensure that others do not come into possession of material or information in conflict with this clause. Employees and others who resign from the service of one of the data processors shall be subject to confidentiality on the matters mentioned above also after their resignation.

This provision also applies after the termination of the Agreement.

9. Entry into force, duration and termination

9.1 Entry into force and duration

The Agreement comes into force when it has been signed by both parties.

<alt. 1>

The Agreement applies as long as the Data Processor processes personal data on behalf of the Data Controller in accordance with the purpose stated in this Agreement.

<alt.2>

The Agreement comes into force on and lasts until The Agreement can be terminated with months' written notice.

9.2 Termination

Unless otherwise agreed with the Data Controller, on termination of this Agreement the Data Processor undertakes to return all health data and personal data that have been received on behalf of the Data Controller and that are included in this Agreement.

The Data Processor shall delete all documents, data, hard disks, CDs and other storage media that contain information that is included in the Agreement. The deletion shall be carried out in a way that prevents the data being retrieved. This also applies to any back-up copies.

10. Breach of contract

Breach of contract occurs if one of the parties does not fulfil its duties according to this Agreement and when this is not due to circumstances for which the other party bears responsibility or risk. If one of the parties wishes to invoke breach of contract, the other party must be notified of this in writing without undue delay.

In the event of breach of contract, the injured party may withhold payment in return, although the amount withheld shall not be clearly higher than what seems necessary to remedy the effects of the breach, and only until the matter has been brought into accordance with the Agreement.

Should a material breach occur, the other party may – after having given written notice and a reasonable deadline for remedying the matter – terminate all or parts of the Agreement with immediate effect, and may demand compensation for any loss this has caused.

11. Transfer of rights and obligations

The Data Controller may, completely or partially, transfer its rights and obligations pursuant to this Agreement to another body, which is then entitled to equivalent terms and conditions. The Data Processor can require any additional expenses incurred by the transfer to be covered.

The Data Processor can transfer its rights and obligations pursuant to the Agreement with the written consent of the Data Controller. Such consent cannot be refused without reasonable grounds. The right to remuneration according to the Agreement can be freely transferred, but the transfer does not exempt the Data Processor from its obligations and responsibilities.

12. Governing law

The parties' rights and duties pursuant to this Agreement are determined in their entirety by the law of *<usually the country of the Data Controller>*.

13. Signing

This Agreement has been signed in 2 – two – copies, each party retaining 1 – one – copy.

<Place>, on

<Data Controller>

(signature)

.....

Name:

(in block capitals)

Position:.....

<Place>, on

<Data Processor>

(signature)

.....

Name:

(in block capitals)

Position:

ANNEX D. GLOSSARY

Term	Explanation
AAA	Authentication, Authorisation and Accounting
AAI	Authentication and Authorization Infrastructure
ANDS	Australian National Data Service
APARSEN	Alliance for Permanent Access to the Records of Science in Europe Network
CC	Creative Commons (an IPR licensing scheme)
CDI	Collaborative Data Infrastructure
CERT	Computer Emergency Response Team
DataCite	An organisation which standardises and assigns PIDs for data (cf. DOI)
DCC	Digital Curation Centre
DMP	Data Management Plan (or Planning)
DOI	Digital Object Identifier (a de facto standard PID for data)
DSA	Data Seal of Approval (a repository certification scheme)
EC	European Commission
IPR	Intellectual Property Right
ISO	International Organization for Standardization
ITSM	IT Service Management
NDS	National Data Service (USA)
OAI-PMH	Open Archives Initiative Process for Metadata Harvesting (a metadata exchange protocol)
OAIS	Open Archival Information System
OpenAIRE	Open Access Infrastructure for Research in Europe
OSCT	Operational Security Coordination Team
PID	Persistent Identifier
QA	Quality Assurance
QC	Quality Control
QoS	Quality of Service
RDA	Research Data Alliance
RI	Research Infrastructure
RP	Resource Provider
SLA	Service Level Agreement
SP	Service Provider
ToU	Terms of Use