MobaGenetics

Core DS Knowledge Model

Organization	ELIXIR Norway
Created by	Korbinian Bösl (korbinian.bosl@uib.no)
Based on	RI gap analysis, 0.0.1 (elixir.no:ri-elixir-norway:0.0.1)
Project Phase	Before Submitting the Proposal
Created at	11 Mar 2021

I. Administrative details

Report

Indications

Answered (current phase)	0 / 1
Answered	9 / 11

Metrics

No metrics for this chapter.

Questions

1 Contributors

Each person contributing to creating or executing the data management plan should be added as a contributor. A project probably should have a Contact Person, and a Data Curator.

🏷 Tags: *maDMP, Science Europe DMP*

Answers

1.b.1 Name

🏷 Tags: *maDMP, Science Europe DMP*

✔ *Gutorm Høgåsen*

1.b.2 E-mail address

🏷 Tags: *maDMP, Science Europe DMP*

✔ *GutormThomas.Hogasen@fhi.no*

1.b.3 ORCID Identifier

🏷 Tags: *maDMP, Science Europe DMP*

✘ **This question has not been answered yet!**

1.b.4 Affiliation

🏷 Tags: *Science Europe DMP*

✔ *NIPH - Norwegian Institute for Public Health*

Role

Roles in a project should be given as they are defined by [datacite](#).

You should specify at least one "Contact Person". If your project has a work package for data management, identify the leader of that work package as "Data Curator".

🏷️ Tags: *maDMP, Science Europe DMP*

✔️ *c. Data Curator*

---

2 **RI**

Add each of the project(s) that are you will be working on and for which the data and work are described in this DMP. Give each project a small identifying name for yourself.

🏷️ Tags: *maDMP, Science Europe DMP*

✖️ **This question has not been answered yet!**

---

3 **To execute the DMP, is additional specialist expertise required?**

🏷️ Tags: *Science Europe DMP*

✔️ *b. Yes, we will be training existing staff*

> 3.b.1 **What kind of training?**
>
> 🏷️ Tags: *Science Europe DMP*
>
> ✔️ *QC/publish/documentation*

---

4 **Do you require hardware or software in addition to what is currently available in the participating institutions?**

✔️ *b. Yes*

> 4.b.1 **What specific hard/software do you need, and why?**
>
> 🏷️ Tags: *Science Europe DMP*
>
> ✔️ *Storage, memory, cpu, possibly hpc*

II. Re-using data

Before you decide to embark on any new study, it is nowadays good practice to check all options to re-use existing available data, either collected or generated by yourself in an earlier project, or data from others (Barend Mons calls this "Other PEople's Data And Services" or OPEDAS). This can include reusable data that have been created for an earlier study, and also so-called "reference data" which is used by many projects.

It is not because we can generate massive amounts of data that we always need to do so. Creating data with public money is bringing with it the responsibility to treat those data well and (if potentially useful) make them available for re-use by others. And the circle is only complete if such data is actually re-used.

Report

Indications

Answered (current phase)	6 / 6
Answered	8 / 9

Metrics

Metric	Score
Reusability	1

Questions

1 **Describe the utility of data produced at the RI; to whom might it be useful?**

✔ *MobaGenetics is by design a general purpose genetic research database (https://www.fhi.no/en/studies/moba/for-forskere-artikler/genetic-data-from-the-norwegian-mother-and-child-cohort-study-mobagenetics/)*

2 **Is there pre-existing data?**

Are there any data sets available in the world that are relevant to your planned research?

🏷 Tags: *maDMP, Science Europe DMP*
📋 Data Stewardship for Open Science: *atq*

✔ *b. Yes*

2.b.1 **Will you be using any pre-existing data (including other people's data)?**

Will you be referring to any earlier measured data, reference data, or data that should be mined from existing literature? Your own data as well as data from others?

🏷 Tags: *maDMP, Science Europe DMP*
📋 Data Stewardship for Open Science: *ezi*

✔ *b. Yes*

2.b.1.b.1 **What reference data will you use?**

Much of todays data is used in comparison with reference data. You may be comparing your own data with a "standard set" which is maintained as a collection by someone else. Or you could be determining differences to a stanndard (in bioinformatics, a genome is often compared with a reference genome to identify genomic variants). If you use reference data, there are several other issues that you should consider. What are the reference data sets that you will use?

🏷 Tags: *Science Europe DMP*
📋 Data Stewardship for Open Science: *quc*

Answers

No answer items

2.b.1.b.2 **Will you use non-reference data sets?**

🏷 Tags: *Science Europe DMP*

✔ *b. No*

2.b.1.b.3 **Will you couple existing (biobank) data sets?**

✔ *a. No*

2.b.2 **Do you need to harmonize different sources of existing data?**

If you are combining data from different sources, harmonization may be required. You may need to re-analyse some original data.

📋 Data Stewardship for Open Science: *wht*

✔ *b. Yes*

> **2.b.2.b.1** **Will you be making your harmonization results available to others?**
>
> By publishing either exactly what you did or (better) make sure that the harmonized data is available for reuse, you may save others the effort
>
> ✔ *b. Yes*

2.b.3 **Will you be using data that needs to be (re-)made computer readable first?**

Some old data may need to be recovered, e.g. from tables in scientific papers or may be punch cards.

▤ Data Stewardship for Open Science: *pth*

✔ *a. No*

III. Creating and collecting data

We will make sure that we know what data will be generated at the RI and when it will be generated. We also need to make sure that there will be adequate storage space to deal with it, and that all the responsibilities have been taken care of.

Report

Indications

Answered (current phase)	11 / 14
Answered	30 / 35

Metrics

Metric	Score
Accessibility	0
Interoperability	1
Good DMP Practice	1

Questions

1 **What data formats/types will you/your users be using?**

Have you identified types of data that you will use that are used by others too? Some types of data (for example "images" or "tables") are used by many different projects. For such data, often common standards exist (in our example "PNG" and "CSV") that help to make these data reusable. Are you using such common data formats?

You should make sure also to list the formats used in any data sets that you are re-using.

🏷 Tags: *Science Europe DMP*
▤ Data Stewardship for Open Science: *njy*

✖ **This question has not been answered yet!**

2 **Will you/your users be using new types of data?**

Sometimes the type of data you collect can not be stored in a commonly used data format. In such cases you may need to make your own, keeping interoperability as high as possible.

▤ Data Stewardship for Open Science: *ikk*

✖ **This question has not been answered yet!**

> **3** **How will you/your users be storing metadata?**

For the re-usability of your data by yourself or others at a later stage, a lot of information about the data, how it was collected and how it can be used should be stored with the data. Such data about the data is called metadata, and this set of questions are about this metadata.

SEEK is a webtool to store (meta)data and provenance. The public global instance FAIRDOMHub is free to users in Norway. SEEK can can be integrated with the data storage and analysis platform for users in Norway NeLS .

▤ Data Stewardship for Open Science: _rhm_

⧉ External Links: _SEEK_

✖ **This question has not been answered yet!**

> **4** **Please specify what data you will acquire using measurement equipment**

You can use any name for the data set, make sure that it identifies the data set to yourself.

🏷 Tags: _Science Europe DMP_

Answers

> **4.b.1** **Who will do the measurements? And where?**

🏷 Tags: _Science Europe DMP_

✔ _d. External party_

consider making them partner in the project

> **4.b.1.d.1** **Has formal ownership of the data been established?**

🏷 Tags: _Science Europe DMP_

✔ _c. We have made other arrangements_

> **4.b.1.d.1.c.1** **What other arrangements?**

🏷 Tags: _Science Europe DMP_

✔ _Project is obliged to return results to us - and we own the results_

> **4.b.1.d.2** **Has responsibility for long term safe keeping of the raw data been established? Who will deal with data publication?**

✔ _d. We have made other arrangements_

> **4.b.1.d.2.d.1** **What other arrangements?**

✔ _"Measuring" project uses data as soon as they want, we usually don't have embargo period before we can publish them to other research groups. (However, there can be long delays before publishing due to extensive/complicated QC)_

> **4.b.2** **Instruments used for data collection**

Specify what technical instruments you are using to collect the data.

🏷 Tags: _Science Europe DMP_

Answers
No answer items

4.b.3 **Is the equipment completely standard and well described?**

If the technology is very much under development, you may want to come back later to understand exactly how the measurements have been made. Is the measurement equipment and protocol sufficiently standard that you will be able to explain how it is done or refer to a standard explanation?

🏷 Tags: *Science Europe DMP*

✔ *a. Very well described and known*

4.b.4 **Is special care needed to get the raw data ready for processing?**

Where does the data come from? And who will need it? Sometimes the raw data is measured somewhere else than where the primary processing is taking place. In such cases the ingestion or transport of the primary data may take special planning. You also need to make sure that data is secure and that data integrity is guaranteed.

✔ *b. Yes, lets explore this*

> **4.b.4.b.1** **Is the data format established?**
>
> Has the storage and transport format of the primary data been established between the people responsible for the measurement and the people responsible for the processing?
>
> ✔ *b. Yes*

> **4.b.4.b.2** **How will the raw data be transported?**
>
> ✔ *c. Via the network*
>
> > **4.b.4.b.2.c.1** **Is sufficient network capacity available?**
> >
> > Can the volume of data be accommodated by the standard network connection? Has a special network connection (e.g. light path) that is needed been reserved?
> >
> > ✔ *a. Yes, has been taken care of*

> **4.b.4.b.3** **Is data integrity guaranteed during this stage?**
>
> Do you have any means of identifying whether the raw data has been transferred error free and has not been tampered with?
>
> ✔ *a. No*

> **4.b.4.b.4** **Is data security guaranteed during this stage?**
>
> Are the raw data encrypted or otherwise protected from theft or leaks at either site or during transport? You could e.g. use a light path or a virtual private network if you transport the data over the net.
>
> ✔ *b. Yes*

4.b.5 **Will you be using quality processes?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

> **4.b.5.b.1** **Are you calibrating measurements?**
>
> 🏷 Tags: *Science Europe DMP*
>
> ✔ *b. Yes*

4.b.5.b.2 **Are you running repeat samples or are you repeating measurements?**

🏷 Tags: *Science Europe DMP*

✔ *a. No*

4.b.5.b.3 **Are you running standardized data capture or recording?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

4.b.5.b.4 **Are you doing Data Entry validation?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

4.b.5.b.5 **Are you using data peer review?**

🏷 Tags: *Science Europe DMP*

✔ *a. No*

4.b.5.b.6 **Are you using controlled vocabularies?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

4.b.5.b.7 **Are you using any other quality processes?**

🏷 Tags: *Science Europe DMP*

✔ *a. No*

---

5 **Do you have any non-equipment data capture?**

Does the data you collect contain non-equipment captured data such as questionnaires, case report forms, electronic patient records?

🏷 Tags: *Science Europe DMP*
📄 Data Stewardship for Open Science: *ybw*

✔ *b. Yes*

5.b.1 **Will you be collecting questionnaires?**

🏷 Tags: *Science Europe DMP*
🔗 External Links: *Nettskjema for collecting sensitive data*

✔ *b. Yes*

5.b.2 **Will you be collecting case report forms?**

🏷 Tags: *Science Europe DMP*
📄 Data Stewardship for Open Science: *hfg*

✔ *a. No*

5.b.3 **Will you be collecting data from electronic patient records?**

🏷 Tags: *Science Europe DMP*

✔ *a. No*

5.b.4 **Please specify how non-equipment data sets will be collected**

🏷 Tags: *Science Europe DMP*

✖ **This question has not been answered yet!**

6 **Is there a data integration tool that can handle and combine all the data types you are dealing with in your RI?**

✔ *a. No*

6.a.1 **Can all data be brought into the same format?**

✔ *b. Yes*

7 **Will you be storing physical samples?**

📖 Data Stewardship for Open Science: *kuz*

✔ *b. Yes*

You might want to contact Biobank Norway for advice

8 **Will you need consent for any newly collected personal data?**

🏷 Tags: *maDMP, Science Europe DMP*
🔗 External Links: *NSD Information and consent, REC Informed consent*

✔ *e. Yes, we will collect consent for our use of the data, and will anonymize the data afterwards for reuse*

9 **How is the ownership of the collected data arranged?**

🏷 Tags: *Science Europe DMP*

✔ *c. All data will be owned by the institute*

IV. Data sensitivity

Ethical and legal issues

adapted from 2019 version of NSD DMP tool and Tryggve Checklist on ELSI issues and GDPR compliance

Report

Indications

Answered (current phase)	40 / 75
Answered	40 / 77

Metrics

Metric	Score
Good DMP Practice	1

Questions

1 **Will you collect or generate data about people?**

✔ *a. Yes*

1.a.1 **Will you collect and/or process personally identifiable data?**

What is personally identifiable data?

Personal data is any information that can be connected to a person e.g. name, address, phone number, e-mail address, IP-address, car registration number, images, fingerprints, iris patterns, head shape (for facial recognition) and birth number, or through a combination of background information. Information about behavioral patterns may also be considered as personal data.

Sensitive personal data is information relating to racial or ethnic origin, political, philosophical or religious beliefs, that a person has been suspected, charged or convicted of a crime, health, sex life, and union membership.

Read more about personal and sensitive data at the Data Inspectorate and at the NSD - Data Protection Services.

Sensitive data has to be stored and analysed using appropriate measures and infrastructure. (such as TSD) - You can apply for quotas through: contact@bioinfo.no

✔ *a. Yes*

1.a.1.a.1 **Please specify**

Answers

1.a.1.a.1.b.1 **Give a short description of the data and the contained personal information**

✔ *Genetic data is in itself a marker to identify a person. However, questionnaire data and other phenotypes are not*

1.a.1.a.1.b.2 **Have you been in contact with a data protection official or other inspectorates?**

If your are to archive personally identifiable data, you must document that you are allowed to archive the data in accordance with current regulations. If you do not have permission to store personally identifiable data, data must be anonymised so that anonymous version can be archived before the original data is deleted.

Consider using EGA or a local branch of it.

✔ *a. Yes*

1.a.1.a.1.b.2.a.1 **Please indicate which inspectorate your project has been in contact with.**

Answers

1.a.1.a.1.b.2.a.1.b.1 **NSD - Data Protection Services**

⧉ External Links: *NSD Data Protection Official for Research*

✔ *a. Yes*

1.a.1.a.1.b.2.a.1.b.1.a.1 **Case number**

✘ **This question has not been answered yet!**

1.a.1.a.1.b.2.a.1.b.2 **REC - Regional Committees for Medical and Health Research Ethics**

☑ External Links: *Regional Committees for Medical and Health Research Ethics*

✔ *a. Yes*

> 1.a.1.a.1.b.2.a.1.b.2.a.1   **Case number**
>
> ✖ **This question has not been answered yet!**

1.a.1.a.1.b.2.a.1.b.3   **The Norwegian Data Protection Authority**

☑ External Links: *Norwegian Data Protection Authority*

✔ *a. Yes*

> 1.a.1.a.1.b.2.a.1.b.3.a.1   **Case number**
>
> ✖ **This question has not been answered yet!**

1.a.1.a.1.b.2.a.1.b.4   **Other**

✖ **This question has not been answered yet!**

1.a.1.a.1.b.3   **Where are you going to process active personally identifiable research data?**

ELIXIR Norway provides sustainable free storage to the Norwegian life sciences community in collaboration with Sigma2 on TSD upon application.

☑ External Links: *TSD, HUNTCloud, SAFE, EUTRO*

✔ *a. TSD*

1.a.1.a.1.b.4   **Do you have a consent form the research subjects already?**

The content of informed consents needed to be valid under different laws (e.g. GDPR, the Health Research Act or other ethical legislation) might differ.

☑ External Links: *GDPR consent, The Health Research Act - Participant consent*

✔ *a. Yes*

> 1.a.1.a.1.b.4.a.1   **What are the limitations of use defined in the informed consent, if any?**
>
> Data archival and transfers have to be explicitly included in the consent. e.g. only for research on certain types of diseases, sharing only within certain geographical boundaries etc. Alternatively, state if there are no limitations.
>
> ✖ **This question has not been answered yet!**

> 1.a.1.a.1.b.4.a.2   **What is the stated intended research purpose?**
>
> ☑ External Links: *Art. 5 GDPR - Principles relating to processing of personal data*
>
> ✖ **This question has not been answered yet!**

> 1.a.1.a.1.b.4.a.3   **Is the intended research purpose within the scope of the limitations of use that is defined in the ethics approval(s) and/or the informed consent(s)?**
>
> ✔ *a. Yes*

`1.a.1.a.1.b.5` **Who will be the data controllers of the personal data processed in the dataset?**

⬀ External Links: *See also*

Answers

`1.a.1.a.1.b.5.b.1` **Data controller**

GDPR Article 4 (7): "'controller' means 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; […]." For cross-border collaborative projects the controllers of different datasets should be identified. Also, if joint controllership is considered, make sure that all parties understand their obligations, and it is probably good to define the terms for this in an agreement between the parties.

⬀ External Links: *Art. 4 - GDPR Definitions*

✘ **This question has not been answered yet!**

`1.a.1.a.1.b.6` **What is the legal basis for processing the personal data?**

GDPR - Article 6 (1) lists under what conditions the processing is considered lawful. Of these, *Consent* or *Public interest* are relevant when it comes to research. You should determine what legal basis (or bases) you have for processing the personal data in your project. Traditionally, consent has been the basis for processing personal data for research, but under the GDPR there cannot be an imbalance between the processor and the data subject for it to be considered to be freely given. In some countries the use of consent as the legal basis for processing by universities for research purposes is therefore not recommended. In those cases, public interest should probably be your legal basis. Note that if your legal basis for processing is consent, a number of requirements exists for the consent to be considered valid under the GDPR. Consents given before the GDPR might not live up to this. Also note that even if public interest is the legal basis, other laws and research ethics standards might still require you to have consent from the subjects for performing the research. Please consult with the Data Protection Officer of your organisation on which legal basis to apply to your data.

⬀ External Links: *Art. 6 GDPR - Lawfulness of processing*

✔ *b. Consent*

`1.a.1.a.1.b.6.b.1` **Are consents in compliance with the GDPR?**

✔ *a. Yes*

`1.a.1.a.1.b.7` **What are the exemptions for the prohibition for processing of special categories of data (such as health and genetic data)?**

Processing of certain categories of personal data is not allowed unless there are exemptions in law to allow this. Among these categories ("sensitive data") are "[…] data revealing racial or ethnic origin, […] genetic data, […] data concerning health'". Most types of personal data collected in biomedical research will fall under these categories. Article 9 (2) lists a number of exemptions that apply, of which consent and scientific research are most likely to be relevant for research. Please consult with your Data Protection Officer of your organisation.

⬀ External Links: *Art. 9 GDPR - Processing of special categories of personal data*

✔ *a. Scientific research*

`1.a.1.a.1.b.8` **Have data processing agreements been established between the data controller(s) and any data processors?**

A data processing agreement has to contain the obligatory clauses specified in Art 28.3 of the GDPR. The agreement should also regulate the use of any sub-processors.

⬀ External Links: *Art. 28 - GDPR Processor*

✔ *a. Yes*

`1.a.1.a.1.b.8.a.1` **List Processors**

✖ **This question has not been answered yet!**

`1.a.1.a.1.b.9` **Have Data Protection Impact Assessments (DPIA) been performed for the personal data?**

List DPIAs done and for which parts of the data. *Note:* All Nordic Data Protection Authorities have identified that most types of research projects on health or genetic data require a DPIA.

Where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data, a so called Data Protection Impact Assessment (DPIA) - Article 35. To clarify when this is necessary, the Data Protection Authorities (DPAs) in Denmark, Finland, Norway and Sweden have issued guidance of when an impact assessment is required. Large-scale processing of sensitive data such as genetic or other health related data is listed by all DPAs as requiring DPIAs. The French DPA has made a PIA tool (endorsed by several other DPAs) available that can help in performing these impact assessments. Please also consult your Data Protection Officer of your organisation.

⬀ External Links: *Art. 35 GDPR - Data protection impact assessment, Datatilsynet: When must an impact assessment be carried out?, The open source PIA software helps to carry out data protection impact assesment*

✔ *b. Needs to be investigated*

`1.a.1.a.1.b.10` **What technical and procedural safeguards have been established for processing the data?**

To ensure that the personal data that you process in the project is protected at an appropriate level, you should apply technical and procedural safeguards to ensure that the rights of the data subjects are not violated. Examples of such measures include, but are not limited to, pseudonymisation end encryption of data, the use of computing and storage environments with heightened security, and clear and documented procedures for project members to follow.

✔ *pseudonymisation. highly secure IT-platforms (TSD/HUnt Cloud)*

`1.a.1.a.1.b.11` **What happens with the dataset after project completion?**

The GDPR states that the processing (including storing) of personal data should stop when the intended purpose of the processing is done. There are, however, exemptions to this e.g. when the processing is done for research purposes. Also, from a research ethics point of view, research data should be kept to make it possible for others to validate published research findings and reuse data for new discoveries. This is also governed by what the data subjects have been informed about regarding how you will treat the data after project completion. The recommendation is to deposit the sensitive data in the appropriate controlled access repositories if such are available, but this requires that the data subjects are informed and have agreed to this. Other considerations

✔ *c. The dataset will be archived in a controlled access repository for re-use.*

Note. Subjects should be informed about this, and that this should be stated in the informed consents and/or ethical approvals.

`1.a.1.a.1.b.11.c.1` **Which archive?**

⬀ External Links: *European Genome-phenome Archive (EGA), Norwegian Federated EGA node*

✔ *c. Other*

`1.a.1.a.1.b.11.c.1.c.1` **Where?**

✔ *MobaGenetics - wherever it will be hosted. Currently on TSD/Hunt Cloud*

`1.a.1.a.1.b.11.c.2` **In what form will the data sets be stored?**

✔ *b. Pseudonymised*

1.a.1.a.1.b.12 **Are there other relevant national legislation considerations that has to be taken into account?**

⬈ External Links: *Lov om helseregistre og behandling av helseopplysninger (helseregisterloven) , Lov om medisinsk og helsefaglig forskning (helseforskningsloven), Forskrift om organisering av medisinsk og helsefaglig forskning, Merknader til forskrifter til Helseforskningsloven, Veileder til Helseforskningsloven, Forskrift om befolkningsbaserte helseundersøkelser, Lov om helsepersonell mv (helsepersonelloven), Lov om pasient- og brukerrettigheter (pasient- og brukerrettighetsloven), Lov om legemidler mv (legemiddelloven), Forskrift om klinisk utprøving av legemidler til mennesker, Norm for helsedata, Lov om humanmedisinsk bruk av bioteknologi mm (bioteknologiloven), Lov om arkiv [arkivlova], Lov om organisering av forskningsetisk arbeid (forskningsetikkloven)*

✔ *a. Yes*

> **1.a.1.a.1.b.12.a.1** **Which?**
>
> ✔ *Lov om helseregistre og behandling av helseopplysninger (helseregisterloven) + more (I don't know all the legalese)*

1.a.1.a.1.b.13 **Are there other Terms & conditions for data access (in particular if presenting obstacles for cross-border processing of health data)?**

E.g. register data access policies (requirement of PI in the same country, moving data to other secure services)

✔ *a. Yes*

> **1.a.1.a.1.b.13.a.1** **Which?**
>
> ✘ **This question has not been answered yet!**

1.a.1.a.1.c.1 **Give a short description of the data and the contained personal information**

✔ *Genotypes*

1.a.1.a.1.c.2 **Have you been in contact with a data protection official or other inspectorates?**

If your are to archive personally identifiable data, you must document that you are allowed to archive the data in accordance with current regulations. If you do not have permission to store personally identifiable data, data must be anonymised so that anonymous version can be archived before the original data is deleted.

Consider using EGA or a local branch of it.

✔ *a. Yes*

> **1.a.1.a.1.c.2.a.1** **Please indicate which inspectorate your project has been in contact with.**
>
> **Answers**
> No answer items

1.a.1.a.1.c.3 **Where are you going to process active personally identifiable research data?**

ELIXIR Norway provides sustainable free storage to the Norwegian life sciences community in collaboration with Sigma2 on TSD upon application.

⬈ External Links: *TSD, HUNTCloud, SAFE, EUTRO*

✗ This question has not been answered yet!

1.a.1.a.1.c.4   **Do you have a consent form the research subjects already?**

The content of informed consents needed to be valid under different laws (e.g. GDPR, the Health Research Act or other ethical legislation) might differ.

⬈ External Links: *GDPR consent, The Health Research Act - Participant consent*

✔ *a. Yes*

1.a.1.a.1.c.4.a.1   **What are the limitations of use defined in the informed consent, if any?**

Data archival and transfers have to be explicitly included in the consent. e.g. only for research on certain types of diseases, sharing only within certain geographical boundaries etc. Alternatively, state if there are no limitations.

✗ This question has not been answered yet!

1.a.1.a.1.c.4.a.2   **What is the stated intended research purpose?**

⬈ External Links: *Art. 5 GDPR - Principles relating to processing of personal data*

✗ This question has not been answered yet!

1.a.1.a.1.c.4.a.3   **Is the intended research purpose within the scope of the limitations of use that is defined in the ethics approval(s) and/or the informed consent(s)?**

✗ This question has not been answered yet!

1.a.1.a.1.c.5   **Who will be the data controllers of the personal data processed in the dataset?**

⬈ External Links: *See also*

✗ This question has not been answered yet!

1.a.1.a.1.c.6   **What is the legal basis for processing the personal data?**

GDPR - Article 6 (1) lists under what conditions the processing is considered lawful. Of these, *Consent* or *Public interest* are relevant when it comes to research. You should determine what legal basis (or bases) you have for processing the personal data in your project. Traditionally, consent has been the basis for processing personal data for research, but under the GDPR there cannot be an imbalance between the processor and the data subject for it to be considered to be freely given. In some countries the use of consent as the legal basis for processing by universities for research purposes is therefore not recommended. In those cases, public interest should probably be your legal basis. Note that if your legal basis for processing is consent, a number of requirements exists for the consent to be considered valid under the GDPR. Consents given before the GDPR might not live up to this. Also note that even if public interest is the legal basis, other laws and research ethics standards might still require you to have consent from the subjects for performing the research. Please consult with the Data Protection Officer of your organisation on which legal basis to apply to your data.

⬈ External Links: *Art. 6 GDPR - Lawfulness of processing*

✔ *b. Consent*

1.a.1.a.1.c.6.b.1   **Are consents in compliance with the GDPR?**

✔ *a. Yes*

1.a.1.a.1.c.7   **What are the exemptions for the prohibition for processing of special categories of data (such as health and genetic data)?**

Processing of certain categories of personal data is not allowed unless there are exemptions in law to allow this. Among these categories ("sensitive data") are "[...] data revealing racial or ethnic origin, [...] genetic data, [...] data concerning health'". Most types of personal data collected in biomedical research will fall under these categories. Article 9 (2) lists a number of exemptions that apply, of which consent and scientific research are most likely to be relevant for research. Please consult with your Data Protection Officer of your organisation.

⤢ External Links: *Art. 9 GDPR - Processing of special categories of personal data*

✔ *a. Scientific research*

[1.a.1.a.1.c.8] **Have data processing agreements been established between the data controller(s) and any data processors?**

A data processing agreement has to contain the obligatory clauses specified in Art 28.3 of the GDPR. The agreement should also regulate the use of any sub-processors.

⤢ External Links: *Art. 28 - GDPR Processor*

✔ *a. Yes*

> [1.a.1.a.1.c.8.a.1] **List Processors**
>
> ✘ **This question has not been answered yet!**

[1.a.1.a.1.c.9] **Have Data Protection Impact Assessments (DPIA) been performed for the personal data?**

List DPIAs done and for which parts of the data. *Note:* All Nordic Data Protection Authorities have identified that most types of research projects on health or genetic data require a DPIA.

Where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data, a so called Data Protection Impact Assessment (DPIA) - Article 35. To clarify when this is necessary, the Data Protection Authorities (DPAs) in Denmark, Finland, Norway and Sweden have issued guidance of when an impact assessment is required. Large-scale processing of sensitive data such as genetic or other health related data is listed by all DPAs as requiring DPIAs. The French DPA has made a PIA tool (endorsed by several other DPAs) available that can help in performing these impact assessments. Please also consult your Data Protection Officer of your organisation.

⤢ External Links: *Art. 35 GDPR - Data protection impact assessment, Datatilsynet: When must an impact assessment be carried out?, The open source PIA software helps to carry out data protection impact assesment*

✘ **This question has not been answered yet!**

[1.a.1.a.1.c.10] **What technical and procedural safeguards have been established for processing the data?**

To ensure that the personal data that you process in the project is protected at an appropriate level, you should apply technical and procedural safeguards to ensure that the rights of the data subjects are not violated. Examples of such measures include, but are not limited to, pseudonymisation end encryption of data, the use of computing and storage environments with heightened security, and clear and documented procedures for project members to follow.

✔ *TSD/Hunt Cloud*

[1.a.1.a.1.c.11] **What happens with the dataset after project completion?**

The GDPR states that the processing (including storing) of personal data should stop when the intended purpose of the processing is done. There are, however, exemptions to this e.g. when the processing is done for research purposes. Also, from a research ethics point of view, research data should be kept to make it possible for others to validate published research findings and reuse data for new discoveries. This is also governed by what the data subjects have been informed about regarding how you will treat

the data after project completion. The recommendation is to deposit the sensitive data in the appropriate controlled access repositories if such are available, but this requires that the data subjects are informed and have agreed to this. Other considerations

✔ *c. The dataset will be archived in a controlled access repository for re-use.*

Note. Subjects should be informed about this, and that this should be stated in the informed consents and/or ethical approvals.

1.a.1.a.1.c.11.c.1 **Which archive?**

⬀ External Links: *European Genome-phenome Archive (EGA),* *Norwegian Federated EGA node*

✘ **This question has not been answered yet!**

1.a.1.a.1.c.11.c.2 **In what form will the data sets be stored?**

✔ *b. Pseudonymised*

1.a.1.a.1.c.12 **Are there other relevant national legislation considerations that has to be taken into account?**

⬀ External Links: *Lov om helseregistre og behandling av helseopplysninger (helseregisterloven), Lov om medisinsk og helsefaglig forskning (helseforskningsloven), Forskrift om organisering av medisinsk og helsefaglig forskning, Merknader til forskrifter til Helseforskningsloven, Veileder til Helseforskningsloven, Forskrift om befolkningsbaserte helseundersøkelser, Lov om helsepersonell mv (helsepersonelloven), Lov om pasient- og brukerrettigheter (pasient- og brukerrettighetsloven), Lov om legemidler mv (legemiddelloven), Forskrift om klinisk utprøving av legemidler til mennesker, Norm for helsedata, Lov om humanmedisinsk bruk av bioteknologi mm (bioteknologiloven), Lov om arkiv [arkivlova], Lov om organisering av forskningsetisk arbeid (forskningsetikkloven)*

✔ *a. Yes*

1.a.1.a.1.c.12.a.1 **Which?**

✘ **This question has not been answered yet!**

1.a.1.a.1.c.13 **Are there other Terms & conditions for data access (in particular if presenting obstacles for cross-border processing of health data)?**

E.g. register data access policies (requirement of PI in the same country, moving data to other secure services)

✔ *a. Yes*

1.a.1.a.1.c.13.a.1 **Which?**

✘ **This question has not been answered yet!**

1.a.1.a.1.d.1 **Give a short description of the data and the contained personal information**

✘ **This question has not been answered yet!**

1.a.1.a.1.d.2 **Have you been in contact with a data protection official or other inspectorates?**

If your are to archive personally identifiable data, you must document that you are allowed to archive the data in accordance with current regulations. If you do not have permission to store personally identifiable data, data must be anonymised so that anonymous version can be archived before the original data is deleted.

Consider using EGA or a local branch of it.

✗ This question has not been answered yet!

1.a.1.a.1.d.3 **Where are you going to process active personally identifiable research data?**

ELIXIR Norway provides sustainable free storage to the Norwegian life sciences community in collaboration with Sigma2 on TSD upon application.

⬈ External Links: *TSD, HUNTCloud, SAFE, EUTRO*

✔ *a. TSD*

1.a.1.a.1.d.4 **Do you have a consent form the research subjects already?**

The content of informed consents needed to be valid under different laws (e.g. GDPR, the Health Research Act or other ethical legislation) might differ.

⬈ External Links: *GDPR consent, The Health Research Act - Participant consent*

✔ *a. Yes*

1.a.1.a.1.d.4.a.1 **What are the limitations of use defined in the informed consent, if any?**

Data archival and transfers have to be explicitly included in the consent. e.g. only for research on certain types of diseases, sharing only within certain geographical boundaries etc. Alternatively, state if there are no limitations.

✗ This question has not been answered yet!

1.a.1.a.1.d.4.a.2 **What is the stated intended research purpose?**

⬈ External Links: *Art. 5 GDPR - Principles relating to processing of personal data*

✗ This question has not been answered yet!

1.a.1.a.1.d.4.a.3 **Is the intended research purpose within the scope of the limitations of use that is defined in the ethics approval(s) and/or the informed consent(s)?**

✗ This question has not been answered yet!

1.a.1.a.1.d.5 **Who will be the data controllers of the personal data processed in the dataset?**

⬈ External Links: *See also*

✗ This question has not been answered yet!

1.a.1.a.1.d.6 **What is the legal basis for processing the personal data?**

GDPR - Article 6 (1) lists under what conditions the processing is considered lawful. Of these, *Consent* or *Public interest* are relevant when it comes to research. You should determine what legal basis (or bases) you have for processing the personal data in your project. Traditionally, consent has been the basis for processing personal data for research, but under the GDPR there cannot be an imbalance between the processor and the data subject for it to be considered to be freely given. In some countries the use of consent as the legal basis for processing by universities for research purposes is therefore not recommended. In those cases, public interest should probably be your legal basis. Note that if your legal basis for processing is consent, a number of requirements exists for the consent to be considered valid under the GDPR. Consents given before the GDPR might not live up to this. Also note that even if public interest is the legal basis, other laws and research ethics standards might still require you to have consent from the subjects for performing the research. Please consult with the Data Protection Officer of your organisation on which legal basis to apply to your data.

✖ **This question has not been answered yet!**

1.a.1.a.1.d.7 **What are the exemptions for the prohibition for processing of special categories of data (such as health and genetic data)?**

Processing of certain categories of personal data is not allowed unless there are exemptions in law to allow this. Among these categories ("sensitive data") are "'[...] data revealing racial or ethnic origin, [...] genetic data, [...] data concerning health'". Most types of personal data collected in biomedical research will fall under these categories. Article 9 (2) lists a number of exemptions that apply, of which consent and scientific research are most likely to be relevant for research. Please consult with your Data Protection Officer of your organisation.

✖ **This question has not been answered yet!**

1.a.1.a.1.d.8 **Have data processing agreements been established between the data controller(s) and any data processors?**

A data processing agreement has to contain the obligatory clauses specified in Art 28.3 of the GDPR. The agreement should also regulate the use of any sub-processors.

✖ **This question has not been answered yet!**

1.a.1.a.1.d.9 **Have Data Protection Impact Assessments (DPIA) been performed for the personal data?**

List DPIAs done and for which parts of the data. *Note:* All Nordic Data Protection Authorities have identified that most types of research projects on health or genetic data require a DPIA.

Where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data, a so called Data Protection Impact Assessment (DPIA) - Article 35. To clarify when this is necessary, the Data Protection Authorities (DPAs) in Denmark, Finland, Norway and Sweden have issued guidance of when an impact assessment is required. Large-scale processing of sensitive data such as genetic or other health related data is listed by all DPAs as requiring DPIAs. The French DPA has made a PIA tool (endorsed by several other DPAs) available that can help in performing these impact assessments. Please also consult your Data Protection Officer of your organisation.

✖ **This question has not been answered yet!**

1.a.1.a.1.d.10 **What technical and procedural safeguards have been established for processing the data?**

To ensure that the personal data that you process in the project is protected at an appropriate level, you should apply technical and procedural safeguards to ensure that the rights of the data subjects are not violated. Examples of such measures include, but are not limited to, pseudonymisation end encryption of data, the use of computing and storage environments with heightened security, and clear and documented procedures for project members to follow.

✖ **This question has not been answered yet!**

1.a.1.a.1.d.11 **What happens with the dataset after project completion?**

The GDPR states that the processing (including storing) of personal data should stop when the intended

purpose of the processing is done. There are, however, exemptions to this e.g. when the processing is done for research purposes. Also, from a research ethics point of view, research data should be kept to make it possible for others to validate published research findings and reuse data for new discoveries. This is also governed by what the data subjects have been informed about regarding how you will treat the data after project completion. The recommendation is to deposit the sensitive data in the appropriate controlled access repositories if such are available, but this requires that the data subjects are informed and have agreed to this. Other considerations

✖ **This question has not been answered yet!**

1.a.1.a.1.d.12 **Are there other relevant national legislation considerations that has to be taken into account?**

⬀ External Links: *Lov om helseregistre og behandling av helseopplysninger (helseregisterloven)*, *Lov om medisinsk og helsefaglig forskning (helseforskningsloven),* *Forskrift om organisering av medisinsk og helsefaglig forskning,* *Merknader til forskrifter til Helseforskningsloven,* *Veileder til Helseforskningsloven,* *Forskrift om befolkningsbaserte helseundersøkelser,* *Lov om helsepersonell mv (helsepersonelloven),* *Lov om pasient- og brukerrettigheter (pasient- og brukerrettighetsloven),* *Lov om legemidler mv (legemiddelloven),* *Forskrift om klinisk utprøving av legemidler til mennesker,* *Norm for helsedata,* *Lov om humanmedisinsk bruk av bioteknologi mm (bioteknologiloven),* *Lov om arkiv [arkivlova],* *Lov om organisering av forskningsetisk arbeid (forskningsetikkloven)*

✖ **This question has not been answered yet!**

1.a.1.a.1.d.13 **Are there other Terms & conditions for data access (in particular if presenting obstacles for cross-border processing of health data)?**

E.g. register data access policies (requirement of PI in the same country, moving data to other secure services)

✖ **This question has not been answered yet!**

1.a.2 **Other comments regarding processing of personal data**

✖ This question has not been answered yet!

2 **Will the RI follow any institutional policies, codes of conducts or other ethical guidelines?**

Each researcher has an independent responsibility for making sure that the research is being carried out in accordance with general scientific and ethical principles and guidelines. For an overview of general and subject-specific research ethics guidelines, see the Norwegian National Research Ethics Committees. Note that in multidisciplinary projects it may be relevant to look to guidelines for several subject areas. In addition, the Research Ethics Act applies to all research in Norway. Also, check which guidelines apply to your institution.

✖ **This question has not been answered yet!**

3 **Other ethical / legal issues.**

✖ This question has not been answered yet!

V. Processing data

In the processing phase, the data will be undergoing the mostly automated steps for processing, before the analysis and interpretation.

Report

Indications

Answered (current phase)	13 / 16
Answered	23 / 31

Metrics

Metric	Score
Accessibility	1
Good DMP Practice	0.64

Questions

1 **Will you be providing the data to the user through a shared working space ?**

Will you be using a working space that is shared between all the people working on the data in the project? Sometimes such a system is called a *Virtual Research Environment*.

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

ELIXIR Norway offers *NeLS* *a multi tiered shared storage for collaborating on data sets*

1.b.1 **Will this work space be run by dedicated specialists?**

If your work space is run and maintained by specialists, e.g. the ICT department of one of the institutes involved in the projects, this means that backup and restore as well as access management is properly addressed.

✔ *b. Yes*

1.b.2 **How will you/your users work with the data?**

There are several questions regarding the dynamics of the data in the working area, who works with it, the software that is run on it, etc.

🏷 Tags: *Science Europe DMP*

✔ *a. Explore*

1.b.2.a.1 **What kind of data will you/your users have in the work space?**

When making the work space, it helps to know whether you expect to work with very many small files, a few very large files, whether you will use a (SQL) database to store most of the data. Maybe your data is suitable for a system like Hadoop? Such information can be collected here.

✔ *genetic + quetionnaires*

1.b.2.a.2 **Do you/your users need the work space to be close to the compute capacity?**

If you have large volumes of data that are intensely and repeatedly used by the computing work flow, it may be needed to keep the storage in the same place as where the computing takes place.

📖 Data Stewardship for Open Science: *wia*

✔ *a. No*

1.b.2.a.3 **Will you/your users be working with your data in another form than the way it will be archived?**

Archival and working with data have different requirements. You want archived information to be in a form that others could read and in a format that is also understandable in a number of years. When working with the data, you need to be able to address it efficiently. If the two differ, you need to plan for conversions.

✔ *b. Yes, archival will require a conversion step*

1.b.2.a.4 **How does the storage need change over time?**

To perform capacity planning, it is important to know what the need for storage capacity at the beginning and the end of the project will be.

🏷 Tags: *Science Europe DMP*

✔ *c. Storage needs are small at the beginning of the RI runtime and will grow later*

1.b.2.a.5 **Will you need to temporarily archive data sets (e.g. to tape)?**

Usually, data sets will be archived if it is unlikely you need them in the short term, but it would be hard to create them again, and/or they are essential for reproducing your work. Archival storage of large volumes can be significantly cheaper than keeping it in the working area for an extensive period.

✔ *a. No*

1.b.2.a.6 **How will your first data come in?**

✔ *a. No special planning is needed for the initial data*

1.b.2.a.7 **How will the RI parterns/ the users access the work space?**

✖ **This question has not been answered yet!**

1.b.3 **How available/reliable should must the work space be?**

There are a number of questions that can help you to decide whether your work space will be reliable enough for your project.

🏷 Tags: *Science Europe DMP*

✖ **This question has not been answered yet!**

1.b.4 **How will access control to the work space be controlled?**

🏷 Tags: *Science Europe DMP*

✔ *c. Only RI members will have read access; only selected RI members will be able to write data*

2 **Data storage systems and file naming conventions**

It is a good idea to pre-define how data will be organised in the project work space, and to set conventions for how any data files and folders will be named.

🏷 Tags: *Science Europe DMP*

✖ **This question has not been answered yet!**

3 **Workflow development**

It is likely that you will be developing or modifying the workflow for data processing. There are a lot of aspects of this workflow that can play a role in your data management, such as the use of an existing work flow engine, the use of existing software vs development of new components, and whether every run needs human intervention or whether all data processing can be run in bulk once the work flow has been defined.

✔ *a. This has been arranged*

4 **How will you make sure to know what exactly has been run?**

✘ This question has not been answered yet!

5 **How will you validate the integrity of the results?**

✘ This question has not been answered yet!

6 **Do you need to do compute capacity planning?**

If you require substantial amounts of compute power, amounts that are not trivially absorbed in what you usually have abailable, some planning is necessary. Do you think you need to do compute capacity planning?

✘ **This question has not been answered yet!**

7 **Is the risk of information loss, leaks and vandalism acceptably low?**

There are many factors that can contribute to the risk of information loss or information leaks. They are often part of the behavior of the people that are involved in the project, but can also be steered by properly planned infrastructure.

🏷 Tags: *Science Europe DMP*

✔ *a. Explore*

> **7.a.1** **Do RI members store data or software on computers in the lab or external hard drives connected to those computers?**
>
> When assessing the risk, take into account who has access to the lab, who has (physical) access to the computer hardware itself. Also consider whether data on those systems is properly backed up
>
> 🏷 Tags: *Science Europe DMP*
>
> ✔ *a. No*
>
> **7.a.2** **Do RI members carry data with them?**
>
> Does anyone carry project data on laptops, USB sticks or other external media?
>
> 🏷 Tags: *Science Europe DMP*
>
> ✔ *a. No*
>
> **7.a.3** **Do RI members store project data in cloud accounts?**
>
> Think about services like Dropbox, but also about Google Drive, Apple iCloud accounts, or Microsoft's Office365
>
> ✔ *a. No*
>
> **7.a.4** **Do RI members send project data or reports per e-mail or other messaging services?**
>
> ✔ *a. No*
>
> **7.a.5** **Do all data centers where RI data is stored carry sufficient certifications?**
>
> 🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

7.a.6 **Are all RI web services addressed via secure http (https://)?**

🏷 Tags: *Science Europe DMP*

✖ **This question has not been answered yet!**

7.a.7 **Have RI members been instructed about the risks (generic and specific to the project)?**

RI members may need to know about passwords (not sharing accounts, using different passwords for each service, and two factor authentication), about security for data they carry (encryption, backups), data stored in their own labs and in personal cloud accounts, and about the use of open WiFi and https

🏷 Tags: *Science Europe DMP*

✖ **This question has not been answered yet!**

7.a.8 **Did you consider the possible impact to the RI or organization if information is lost?**

🏷 Tags: *Science Europe DMP*

✔ *d. Yes; we will need to work on this.*

7.a.9 **Did you consider the possible impact to the RI or organization if information leaks?**

🏷 Tags: *Science Europe DMP*

✔ *d. Yes; we will need to work on this.*

7.a.10 **Did you consider the possible impact to the RI or organization if information is vandalized?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes; the effect is small*

7.a.11 **Are personal data sufficiently protected?**

🏷 Tags: *Science Europe DMP*

✔ *d. Yes, all data will be anonymized as early as possible*

Please note that GDPR law in Europe specifies that data is only anonymous as long as nobody in the world has enough information to re-identify the subject.

8 **Do you have a contingency plan?**

What will you do if the compute facility is down?

✔ *a. We will wait until the problem is fixed*

9 **Will you version datasets?**

SEEK which is used in FAIRDOMHub and can be used together with NeLS supports versioning by default.

NeLS can also be used with Git Large File Storage (LFS)

🔗 External Links: *FAIRDOMHub, SEEK, NeLS, Git Large File Storage (LFS)*

✔ *a. Yes*

VI. Interpreting data

The interpretation of the data consists of the last steps of processing (often with manual interventions), visualisation, and data integration. In this chapter many questions about data interoperability will come up.

Report

Indications

Answered (current phase)	0 / 1
Answered	0 / 8

Metrics

No metrics for this chapter.

Questions

1 How will you be doing the integration of different data sources?

✘ **This question has not been answered yet!**

2 Will you/your users be using common or exchangeable units?

✘ This question has not been answered yet!

3 Will you/your users be using common ontologies?

✘ This question has not been answered yet!

4 Will there be potential issues with statistical normalization?

✘ This question has not been answered yet!

5 Will you/your users be integrating different data sources to get more samples or more data points?

✘ This question has not been answered yet!

6 Will you/your users be integrating different data sources in order to get more information for each sample or data point?

✘ This question has not been answered yet!

7 Do you/your users have all tools to couple the necessary data types?

✘ This question has not been answered yet!

8 Will you/your users be doing (automated) knowledge discovery?

▤ Data Stewardship for Open Science: *bzu*

✘ This question has not been answered yet!

VII. Preserving data

VII. Preserving data

In this chapter, issues regarding data publication and long term archiving are addressed.

Report

Indications

Answered (current phase)	0 / 5
Answered	0 / 14

Metrics

No metrics for this chapter.

Questions

1 **Will you /your usersbe archiving data (using so-called 'cold storage') for long term preservation already during the RI runtime/project?**

Much of the raw data you have will need to be archived for your own later use somewhere. This is often done off-line on tape, not on the disks of the compute facility. Please note that this does not refer to the data publication.

📓 Data Stewardship for Open Science: *kjp*

✖ **This question has not been answered yet!**

2 **Specify details of data types which will be produced at your RI**

It is useful to think about a data types as some collection of data that will be ending up in the same place.

🏷 Tags: *maDMP, Science Europe DMP*

✖ This question has not been answered yet!

3 **Will any of the repositories you use charge you/your users for their services?**

🏷 Tags: *Science Europe DMP*

✖ This question has not been answered yet!

4 **Did you budget for the time and effort it will take to help user to prepare the data for publication?**

🏷 Tags: *Science Europe DMP*

✖ This question has not been answered yet!

5 **Will you be making sure that blocks of data deposited by you or by the users in different repositories can be recognized as belonging to the same study?**

✖ This question has not been answered yet!

6 **Are there any recurring fees to keep data or documents available?**

Are you using any commercially licensed products to keep data, software or documents available, for which a regular fee must be paid?

✖ **This question has not been answered yet!**

7 **Will you be archiving your data after the RI runtime in 'cold storage'?**

Will you be storing (in cold storage) copies of your own data for a longer period after the project has ended? Possibly as a continuation of archival as part of data storage strategy during the project? Data archival is distinct from data publishing, an archive is usually limited in who can access the data.

⊟ Data Stewardship for Open Science: *fxe*

✖ **This question has not been answered yet!**

8 **Will you also publish data if the results of your study are negative/inconclusive or unpublishable?**

Even if you do not obtain the results you had foreseen from your own study, the data can still be valuable for reuse in another context. Also, publishing the data can avoid that someone else collects a similar data set with a similar negative result.

✖ **This question has not been answered yet!**

9 **Specify a list of software packages you will be publishing**

Specify a short name for each software package.

✖ This question has not been answered yet!

10 **How will you be making sure there is good provenance of the data (and analysis)?**

Data analysis is normally done manually on a step-by-step basis. It is essential to make sure all steps are properly documented, otherwise results will not be reproducible.

🏷 Tags: *Science Europe DMP*

✖ This question has not been answered yet!

11 **Will reference data be created?**

Will any of the data that you will be creating form a reference data set for future research (by others)?

⊟ Data Stewardship for Open Science: *rbz*

✖ **This question has not been answered yet!**

12 **How will you document your/the user data?**

For reusability, the data should be well documented. In this section of the questionnaire you can specify what kinds of documentation you will be providing.

🏷 Tags: *Science Europe DMP*

✖ This question has not been answered yet!

13 **Will you do systems biology modeling (for users)?**

✖ This question has not been answered yet!

14 **Will you do structural modeling?**

✖ This question has not been answered yet!

VIII. Giving access to data

This chapter deals with the information needed by people who will re-use your data, and with the access conditions

they will need to follow.

Report

Indications

Answered (current phase)	14 / 16
Answered	25 / 28

Metrics

Metric	Score
Findability	0.3
Accessibility	1
Good DMP Practice	0.75
Openness	0.3

Questions

1 **Will you be working with the philosophy 'as open as possible' for your data/your users data?**

❧ Tags: *Science Europe DMP*
📋 Data Stewardship for Open Science: *jvm*

✔ *b. Yes*

2 **Are there potential copyright and Intellectual Property Rights (IPR) issues?**

✔ *a. Yes*

   2.a.1 **How will you manage copyright and Intellectual Property Rights (IPR) issues?**

   ✔ *Approved applications only*

3 **Can all of your data at your RI become completely open immediately?**

❧ Tags: *maDMP, Science Europe DMP*

✔ *a. No*

   3.a.1 **Are there legal reasons why (some of your) data can not be completely open?**

   ❧ Tags: *maDMP, Science Europe DMP*

   ✔ *b. Yes*

      3.a.1.b.1 **Are there privacy reasons why data can not be open?**

      ❧ Tags: *maDMP*

      ✔ *b. Yes*

         3.a.1.b.1.b.1 **Are there restrictions on where the data need to be stored?**

         ❧ Tags: *maDMP*

         ✔ *c. Yes, they must stay in the same country*

            3.a.1.b.1.b.1.c.1 **Are you going to use a national platform for your data?**

✔ *c. HUNT DB*

3.a.1.b.1.b.2 **Could pseudonymization be used to make the data more openly available?**

Legally, pseudonymous data (which means that someone has the key to reverse the process) is still considered privacy sensitive information. However, the EU is working on special cases where the data can still be opened as long as the key availability is sufficiently limited.

🏷 Tags: *maDMP*

✔ *a. No*

3.a.1.b.1.b.3 **Could anonymization be used to make the data more openly available?**

Different anonymization techniques exist. Disadvantage of anonymization is that data integration becomes virtually impossible, but it may be the only way to open up your data for other research

🏷 Tags: *maDMP*

✔ *a. No*

3.a.1.b.1.b.4 **Could you use data aggregation to make the data openly available?**

Aggregated data, where typically at least 15 individuals are in any data point, are considered sufficiently anonymous. This is an alternative way of making data openly available for future research

🏷 Tags: *maDMP*

✔ *a. No*

3.a.1.b.2 **Are there IP reasons why data can not be open?**

✔ *b. Yes*

3.a.1.b.2.b.1 **Is it clear who owns data and documents?**

✔ *b. Yes*

3.a.1.b.2.b.1.b.1 **Who will own the intellectual property rights (copyrights) of the data that you will collect or create?**

✔ *NIPH*

3.a.1.b.2.b.2 **Will someone be given decision power to move documents or data to a new place after the project has finished?**

In one case in the past, all documents that had been assembled by a project in a documentation system had to be deleted because not a single person could decide to move them to a new platform when the documentation system was going off-line.

✔ *b. Yes*

3.a.1.b.3 **Will you/your users be allowing authenticated access to the data?**

🏷 Tags: *Science Europe DMP*

✔ *b. Yes*

3.a.1.b.3.b.1 **Where will the data be stored?**

↗ External Links: *The European Genome-phenome Archive (EGA)*

✔ *b. In a national or institutional repository that arranges restricted access*

3.a.1.b.3.b.2 **Who will take care of authentication of potential users?**

✔ *a. We will use username/password authentication, possibly augmented with two-factor security*

3.a.1.b.3.b.3 **Who will take care of authorization of potential users?**

🏷 Tags: *Science Europe DMP*

✔ *c. We will make use of an existing data access committee*

    3.a.1.b.3.b.3.c.1 **Which existing Data Access Committee?**

    🏷 Tags: *Science Europe DMP*

    ✔ *SEe https://www.fhi.no/en/studies/moba/for-forskere-artikler/genetic-data-from-the-norwegian-mother-and-child-cohort-study-mobagenetics/*

3.a.1.b.3.b.4 **Are the criteria for application to access the data openly available (e.g . are there well described conditions for access (i.e. a machine readable license)?**

✔ *b. Yes*

3.a.1.b.3.b.5 **Has auditing for the re-use been arranged?**

✖ This question has not been answered yet!

3.a.2 **Are there business reasons why (some of) the data at your RI can not be completely open?**

🏷 Tags: *Science Europe DMP*

✔ *c. Yes, other business reasons*

    3.a.2.c.1 **What other business reasons are there not to open all data immediately?**

    🏷 Tags: *Science Europe DMP*

    ✔ *Illegal to share in the first place*

3.a.3 **Are there other reasons why (some of) the data at your RI can not be completely open?**

🏷 Tags: *Science Europe DMP*

✔ *c. Yes, other reasons*

    3.a.3.c.1 **What other reasons are there not to open all data immediately?**

    🏷 Tags: *Science Europe DMP*

    ✖ **This question has not been answered yet!**

3.a.4 **Will you use a limited embargo?**

4 **Will there be valorization or translational returns of the data generated at your RI?**

✘ **This question has not been answered yet!**

Data Management Plan generated by Data Stewardship Wizard <https://ds-wizard.org>

Tags: *Science Europe DMP*

✔ *b. No, data will be released only as soon as restrictions are falling away*

4 **Will there be valorization or translational returns of the data generated at your RI?**

✘ **This question has not been answered yet!**