

RDM Advanced webinar series FAQ

This is a frequently asked questions page collected in the RDM Advanced webinar series sessions during March-April 2021. The questions and answers are categorized by the main RDM themes.

- [Documentation of data and metadata](#)
- [Law & Ethics](#)
- [Storing solutions](#)
- [Publishing and archiving your data](#)
- [Resources, roles and responsibilities](#)
- [DMP in general](#)
- [Miscellaneous questions](#)

Documentation of data and metadata

Law & Ethics

What does it mean that the research is carried on in public interest? Isn't science usually conducted for the common good? Hence, is it or is it not recommended to ask for consent for data procession for instance in an interview study?

There is a difference between *legal basis for personal data processing* and *asking consent for attending a study in ethical sense*. These are two independent issues and typically confused with each other. If using "research is carried out in public interest" as the legal basis for your personal data processing, you only need to inform a research participant about the research but you do not need his or her signature on a paper, in other words written agreement that they are willing to participate in your study.

To use public interest as a legal basis, is more flexible. Your participants cannot withdraw from the study, from the study later. To ask ethical permission to attend the study (that is inform people) is a different thing. You always have to inform people that they involve in your research. You can design a data collection setup where you do not tell people that they are attending a research. There is plenty of cases where this is justified. In these cases, however, you need to have an ethical review board comments for you research setup to check that this kind of setting is necessary in order to conduct your project.

Data will be collected in Africa and Finland by researchers employed at HU and by researchers employed at the partner institutions. Data will be jointly investigated by teams in all participating institutions. Data will be stored in a repository based in Finland. How can I handle data protection and assign controller roles? What help can research services provide regarding agreements with the international partners?

According to the GDPR, an institution processing sensitive and personal data (>250 employees) has to appoint a Data Protection Officer (DPO) - who is the DPO at Univ. Helsinki?

University of Helsinki Data Protection Officer is Lotta Ylä-Sulkava.

Are there any templates of University of Helsinki informed consent form on Flamma?

Unfortunately no(t yet), but you can find instructions for drafting a consent form in Ethical review board's webpage: <https://www.helsinki.fi/en/research/services-researchers/ethical-review-research/humanities-social-sciences-and-behavioural-sciences>

"When designing informed consent forms and questionnaires, a number of aspects should be taken into consideration prior to requesting an ethical review, as such aspects may have an impact on the research subjects' understanding of the study and their attitude towards it. In the information letter to the subjects and in the consent form, researchers should use language that is intelligible to the target group. (...) Information letters and consent forms should be in the subjects' native language or in a language that they understand fluently. (...) When drafting information letters and forms, use polite, respectful language."

I am working with human microbiome datasets (genetic sequences from microbes sampled from human subjects). Therefore I do have information about the human subject (which is clearly personal data), but also genetic information from the microbial communities. It is unclear to me if the sequencing datasets should be considered personal datasets since they are not human genomic data.

Is it sufficient to state in my article manuscript that my research follows University of Helsinki ethical guidelines strictly if my research project per se doesn't require an ethical review?

It would be more informative to readers to know, how did you manage personal information in your project than referring to ethical guidelines, which are not very concrete and should be self-evident to follow them in research.

If I consider my research data are useful for future scientific studies as secondary data and the data are archived at FSD after my study, so if a Non-EU based researcher wishes to access my research data, is this the transfer of data outside of EU?

In many disciplines in the humanities, it is vital that the personal data, including possible sensitive data is and will be available for later research. What do you recommend in regards to this kind of data - especially in the archiving phase?

I am planning to do a data register-based study, containing sensitive information. Does this study require an ethical review by the committee? How do I move forward as I cannot ensure the consent of every person in the register base? How do the “informed about research and access to data by participants rights” work with register-based studies?

You usually don't need to inform the data subject when doing register based study because it would be impossible if it would “cause unreasonable burden”. In these situations it is possible to add your privacy notice to your projects website for example (if you have one).

What if after collecting the data I see different opportunities in it, for instance new analysis methods or concepts, is it reuse of the data? Another option is to be a bit general in defining the subject but that is unethical - or to what level should the accuracy be brought?

From legal perspective, the data can be used for “further use” if the “further use” is research carried out in public interest. The further use should however be somehow “compatible” with the primary use and we evaluate the situation with these kind of criteria: “is the further use similar” (for example same research theme) “what was informed to data subjects” “what could the data subjects reasonably expect their data to be used” “what effects the further processing could cause to data subjects” and make the decision for the possibility to use the data based on this evaluation.

What about publishing metadata? Is an informed consent needed i.e. should we routinely include the description of the intended data in the information sheets?

Is a public event publicly available data? Do I have to ask for informed consent if I want to record the event?

What if a person requests to be removed from the study after a publication is published with their citation (qualitative study)?

You would have a quite good position to say that sorry but this is not possible at this point and you could argue this by “freedom of speech” or “academic expression”. Of course if you have collected the data from the data subject, you should have informed about this when the participant gave his/her opinions and asked a consent to publish the citations with names.

If I make copies of published media articles (newspapers etc.) for research purposes, can I archive the corpus I collected at the end of the project? How does copyright work in that case?

Storing solutions

How secure are those cloud storages to store and share sensitive data?

Don't store sensitive data in cloud services, use University of Helsinki IT services.

If you need to transfer confidential data via third parties, always encrypt the files using strong (20-character, genuinely random) password, and deliver the password separately and safely (for instance, as a GSM text message). NB! There is no guarantee that encryption methods that are considered state of the art at the moment are uncrackable in the future, nor a way to ensure or know that no leaks occur at the middleman. Therefore, process the data in advance by pseudonymization, narrowing down the content, or other suitable means to moderate your worst case scenario risk level.

What kinds of data can be encrypted? Where can I find more information on that?

Any kind of files and folders can be encrypted. More about information security: <https://helpdesk.it.helsinki.fi/en/information-security-and-cloud-services/information-security/information-security>

Can you get access to Umpio storage data also when abroad, e.g. carrying out long-term fieldwork?

You can access Umpio wherever you have a working Internet connection. In some locations connection speed may be much more limited than you've become used to in Finland, though, thus making transfer of large files time consuming. Outside EU/EEA, you may also want to check if a locally purchased prepaid SIM provides more affordable high speed Internet usage, as data roaming can be very costly.

Certain countries that restrict heavily citizens' access to Internet for claimed political, moral, national security etc. reasons, may set further limitations, although at the moment we're not aware of such that would affect Umpio access from anywhere.

What is the cost of storage solutions? What do I need to budget for this?

Basically the price list is here but for very large storage we need to discuss case-by-case: <https://flamma.helsinki.fi/en/group/it-ja-puhelin/it-centers-service-price-list#menu4>

Data storage and sharing table: <https://wiki.helsinki.fi/x/kgV5FQ>

Publishing and archiving your data

At what stage of the project should we contact a suitable repository to inquire about storing the dataset? before collecting the data, after collecting the data..

It would be better to check the requirements of the repository before collecting the data and in some cases also to contact. It depends about the subject. See e.g. <https://www.fsd.tuni.fi/en/services/depositing-data/>

If you have pseudonymized data and you want to open it fully/partially, do you always have to anonymize it? What if the data is not sensitive but just personal data, is anonymization still always mandatory?

I'm planning to publish the non-sensitive parts of my data in FSD, is it recommended to publish my metadata in some other repositories in addition to FSD as well?

Etsin enables you to find research datasets from all fields of science. Etsin contains information about the datasets and metadata in the national Finnish Fairdata services. We also currently harvest information from the Language Bank of Finland, the Finnish Social Science Data archive and the Finnish Environmental Institute, and new sources will be included.

What if I have biological samples collected from animals and I would like to make them available for other researchers - what is a good way to do this? Should I publish the metadata somewhere and let the possibly interested researchers contact me?

This is a good idea. You can publish it in Etsin, for example or in a repository related your field of science (use re3data.org to find it). Or, if you can open the whole dataset and have prepared it properly, you can just publish the dataset in a preferably curated data repository in your field of science.

4 years is a long time in research and your analysis methods, including data produced from biological samples, may change during the course of the project and thus data depositing/archiving may change during the project. How to prepare for this in the RDM plan?

<https://www.kielipankki.fi/corpora/>

Library license guide: <https://libraryguides.helsinki.fi/oa/eng/license>

<https://blogs.helsinki.fi/thinkopen/datarepo2021-summary/>

Resources, roles and responsibilities

Which costs are costs for research itself or which are for data management? It is quite difficult to separate them.

About the responsibility, does a PhD candidate have the responsibility to manage the PhD research project data in a good way even though her /his supervisor did not mention anything about the data management?

DMP in general

It is clear that the DMP needs to be updated etc, but do you have tips as how to put this in the plan that you present e.g. to the funder? Enough details is necessary for the RDM have credit in the eye of the funder, but flexible enough to allow changes in due course?

In a PhD thesis, would it be good to include a small part about data management?

Miscellaneous questions

Is the planned DMP training session for Academy-funded project managers only or can others participate as well? Would be highly appreciated!

We have dedicated workshops for those who have received the positive funding decision - and a couple of open workshops as well. Also, if your group needs you can order tailor-made sessions from us as well. For tailored workshops or seminars, please, send a request to Dataproduct@helsinki.fi