Publication Policy
of the National Research Data Infrastructure for Personal Health Data (NFDI4Health) and the NFDI4Health Task Force COVID-19

Authors:

Birte Lindstädt, Aliaksandra Shutsko, ZB MED – Information Centre Life Sciences, Cologne

on behalf of the NFDI4Health Consortium and the NFDI4Health Task Force COVID-19

In close cooperation with:

Martin Golebiewski, Heidelberg Institute for Theoretical Studies (HITS), Heidelberg

Dennis-Kenji Kipker, Vanessa Lettieri, University of Bremen, Bremen

Sophie Klopfenstein, Carina Vorisek, Berlin Institute of Health at Charité (BIH), Berlin

Matthias Löbe, University of Leipzig, Leipzig

Carsten Oliver Schmidt, University of Greifswald, Greifswald

DOI: 10.4126/FRL01-006431467

Version: V1_0

Release date: 11.02.2022

1. Introduction

One goal of the NFDI4Health Task Force COVID-19[1] and the NFDI4Health[2] at whole is to improve the management of research data from health-related studies focusing on the COVID-19 pandemic and other use cases in epidemiology, public health and clinical research. The project team aims to assist the research community in making relevant German studies easier to find and to improve data sharing in the sense of the FAIR (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable) data principles[3].

To achieve this goal, a multi-level research data infrastructure is being built up, including the creation of a comprehensive inventory of German studies on COVID-19 and other use cases with structured health data from registers, administrative health databases, primary care, clinical studies, including vaccination studies, epidemiological cohorts and health reporting. Important part of this is the **German Central Health**

[1] About NFDI4Health Task Force COVID-19, https://www.nfdi4health.de/en/task-force-covid-19.html
[2] About NFDI4Health, https://www.nfdi4health.de/en/
[3] FAIR Guiding Principles for scientific data management and stewardship, https://www.go-fair.org/fair-principles/

Study Hub COVID-19[4] (**Study Hub**) of the NFDI4Health and the NFDI4Health Task Force COVID-19 allowing the search for and publication of the research data, including metadata.

This **Publication Policy** describes the requirements for the publication of research data from health-related studies in the Study Hub. Researchers or research groups wanting to make their studies discoverable and corresponding research data FAIR should fulfil these requirements. Figure 1 shows the steps needed for publication depending on the type of research data (i.e. "study description" or "study documents") according to the Publication Policy. The text of the Publication Policy can be found in Section 2 of this document.

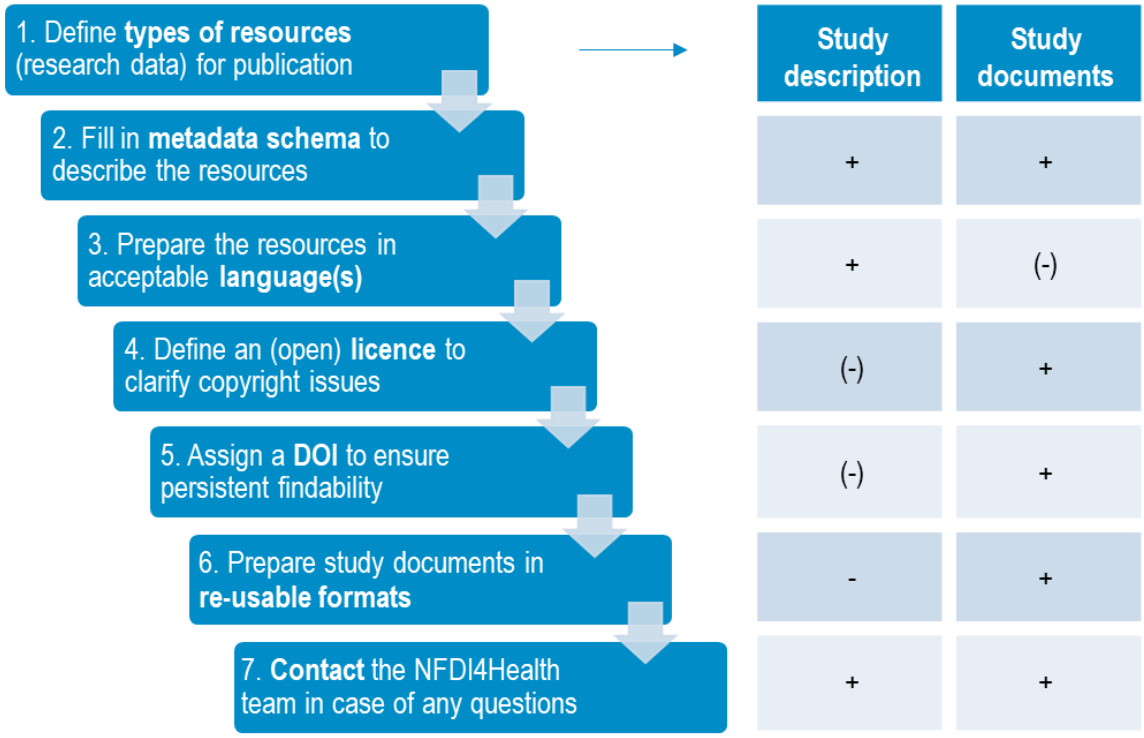

Figure 1. Steps to publish research data depending on their type in accordance with the Publication Policy of the NFDI4Health
Explanation of the symbols. "+": Requirement applies; "-": Requirement does not apply and cannot be fulfilled; "(-)": Requirement does not apply but can be fulfilled under certain conditions (see Publication Policy).

[4] German Central Health Study Hub COVID-19, https://covid19.studyhub.nfdi4health.de/

2. Publication Policy of the NFDI4Health and NFDI4Health Task Force COVID-19

One of the purposes of the Study Hub is to make data collected in health-related studies FAIR. The first step towards this is the publication of comprehensive information about the context and accessibility of the data (metadata set). In addition, study documents contain further detailed information needed for the correct interpretation of the collected data and should therefore be published as well. Accordingly, two main types of research data or resources can be found in the Study Hub.

2.1. Resource types

In the Study Hub, you can publish two main types of resources:

a) **Study descriptions** which comprise comprehensive information (metadata) about interventional or non-interventional studies from clinical, epidemiological or public health sector.

b) **Study documents** which provide information about studies or research-related content such as the design, conduct or instruments. Study documents are made available to the users of the Study Hub for download. Below, the most common types of study documents are listed:

- Study protocol (study plan, study proposal)
- Protocol amendment
- Data dictionary (data catalogue)
- Participant/Patient information sheet
- Informed consent form (template)
- Manual of operations/standard operating procedures (SOPs)
- Statistical analysis plan
- Data management plan

One special type of study documents that can be published in the Study Hub is related to the way data is collected such as:

- Questionnaire/Case report form (CRF)
- Code book
- Interview schema and themes
- Observation guide
- Discussion guide
- Other types of documents used for data collection

Please note that **participant/patient individual data** collected in studies are **not considered** as a resource type and will not be published in the Study Hub. However, information on the availability of the collected data must be provided as part of a study description.

2.2. Metadata schema

To enable a standardized collection of metadata, i.e. information describing the resources (see Section 2.1 for resource types), there is a **metadata schema** to be filled in. An overview of this can be found here[5]. The metadata fields are divided into mandatory and recommended ones. Although only some fields are mandatory, we encourage you to provide as much information as possible to improve the discoverability of your resources. In addition to bibliographic information such as title and description of a resource, you can also include the persons and organizations that are involved in the development of a resource. The results of studies published in journal articles or other text publications can also be linked. For studies, additional information should be provided about study design and accessibility of the data collected.

To submit the metadata, you can send a request at studyregistration@nfdi4health.de. Our research data publication team will be happy to provide further information or assist you in completion of the metadata form.

2.3 Language

The metadata should be provided in English, but we also accept German. The language of the resources themselves, i.e. the language in which a study in Germany is conducted or the language in which a submitted study document is written, is not limited to English and German.

2.4 Licensing

2.4.1 Licensing of metadata

Metadata, e.g. the descriptions of studies or study documents, are not subject to copyright and, therefore, do not need to be licensed and can be freely used after publication.

2.4.2 Licensing of study documents

Documents are copyrighted works. To foster reusability of the published documents – which is an important point of FAIR data principles – authors of study documents may allow third parties to use them (i.e. make a copy, redistribute, edit, post online) by signing a license agreement with them. However,

[5] NFDI4Health Task Force COVID-19 Metadata schema V2_0, https://doi.org/10.4126/frl01-006431357

concluding individual contracts is time-consuming. Therefore, open standard licenses have been developed to simplify the distribution of a work. The best known and most widely used standard licenses are the Creative Commons licenses (CC licenses).

CC licenses are made up of different modules (license modules) which can be combined with each other to compile an individual license according to requirements and to grant the corresponding use. By specifying the license modules, it is immediately clear how the work is allowed to be used. Further information on licensing, including explanation of license modules and selection of a suitable license, can be found in the flyer[6].

Some combinations of CC license modules largely restrict the re-use of resources which is why we strongly recommend determining a less restrictive license for your works. Below you can find a list of suggested licenses:

- CC BY 4.0 (Creative Commons Attribution 4.0 International Public License)
- CC BY-NC 4.0 (Creative Commons Attribution-NonCommercial 4.0 International Public License)
- CC BY-SA 4.0 (Creative Commons Attribution-ShareAlike 4.0 International Public License)
- CC BY-NC-SA 4.0 (Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Public License)

Although CC licenses are preferred, other license models are also possible.

2.4.2.1 Labelling study documents with a license

For the terms and conditions of the selected license to take effect, your work must be properly marked. This is often achieved by adding a statement that your material is licensed under the chosen CC license and attaching the link to the CC webpage where the standard text of the license agreement can be found. As study documents are available for download and, therefore, can be used offline, a simple indication of the license on an individual webpage of the resource in the Study Hub is not sufficient: The license statement and the link to the license text should be inserted directly in the files.

For example, if you decide to license your study document under the CC BY 4.0 license, the following text should be included:

[6] Flyer "Publish study documents and survey instruments quickly and easily", https://www.nfdi4health.de/en/service/media-library/item/111-flyer-publish-study-documents-and-survey-instruments-quickly-and-easily.html

"This work is licensed under the Creative Commons Attribution 4.0 International Public License (CC BY 4.0). To view a copy of the license, visit https://creativecommons.org/licenses/by/4.0/legalcode."

The Creative Commons help page[7] provides further important information on marking works with a CC license.

2.4.2.2 Important notes

When selecting a CC license, some important notes should be taken into account:

- CC licenses are irrevocable: A later withdrawal of the rights granted by a CC license is not possible.
- Only works free of third-party rights can be licensed: Only the copyright holder or someone with expressed permission from the copyright holder can apply a CC license to a work.

If a work cannot be licensed, you can still publish it in the Study Hub by granting the NFDI4Health and the NFDI4Health Task Force COVID-19 the rights to use it. In this case, the parts of a study document that have been taken from works of third parties must be marked in an appropriate manner, for example, by referencing the source of information.

If you have any further questions or concerns about licensing your works, please let us know via e-mail to studyregistration@nfdi4health.de.

2.5 Persistent identifier

Any publication of study documents can be minted with a Digital Object Identifier (DOI). DOI is a persistent identifier used to cite and link any objects — physical, digital or abstract — and guarantee their persistent findability. As a precondition for DOI assignment, study documents must be stored in the Study Hub. Also, the person submitting a resource has to be authorized to mint a DOI, which means that the agreement of all parties holding intellectual property rights is provided and all authors, or creators, of the resource are specified.

[7] Frequently Asked Questions: How do I apply a Creative Commons license to my material? https://creativecommons.org/faq/#how-do-i-apply-a-creative-commons-license-to-my-material

2.6 File formats of the study documents

To make study documents findable and interoperable, provided file formats should be open (non-proprietary), machine-readable and -actionable. Formats like .xml, .txt, .csv, .json, PDF(A) or .rdf are preferred.

We also accept non-machine-readable formats like common PDF and proprietary formats like .xlsx or .docx. However, this will reduce the findability, interoperability and reusability of the study documents.

3. Contact information

If you have any questions and/or concerns about publication of your resources in the Study Hub or need any support, feel free to contact our research data publication team by sending an e-mail to the following address: studyregistration@nfdi4health.de.

4. Acknowledgement

This work was done as part of the NFDI4Health Task Force COVID-19 and NFDI4Health Consortium (www.nfdi4health.de). We gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project numbers 451265285 and 442326535.

5. License