

National research data infrastructure for personal health data

– The use case Task Force COVID-19 –

Iris Pigeot, Juliane Fluck

On behalf of TA5



3. Community Workshop

October 27, 2021

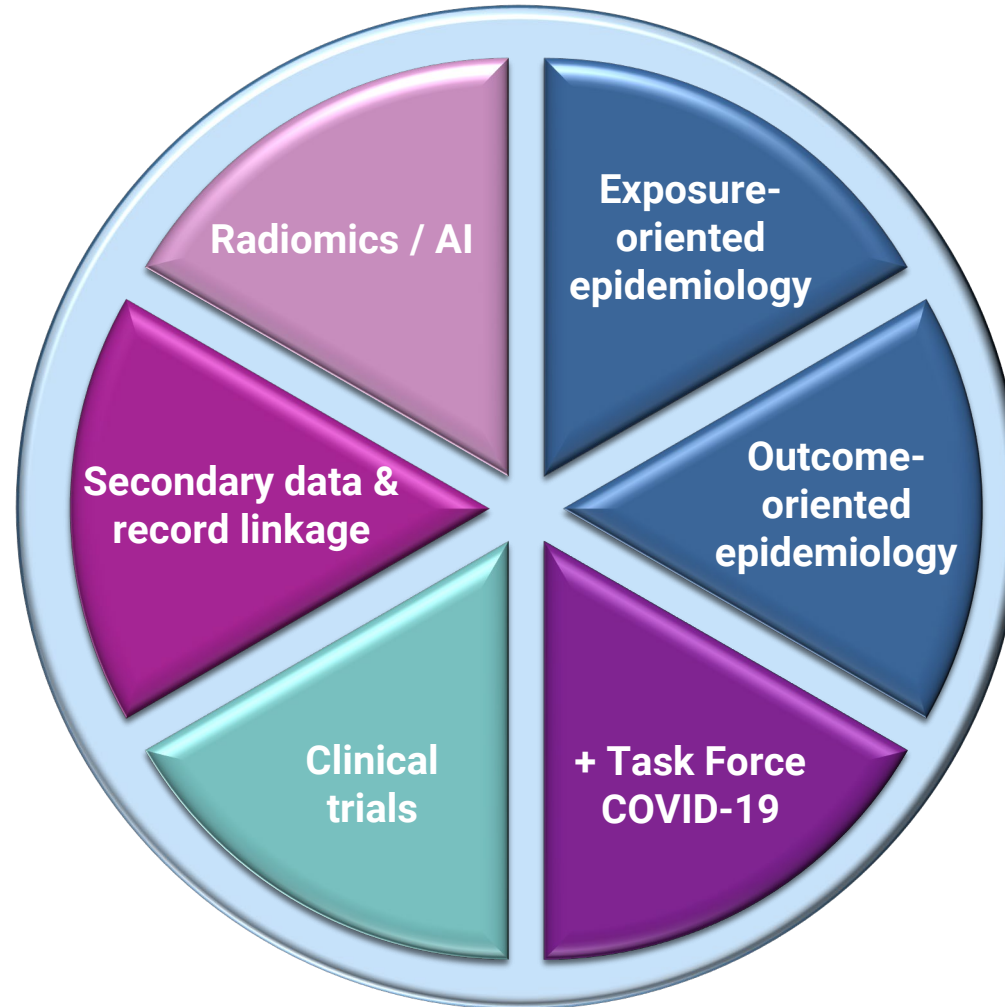


Use Cases: one important component of NFDI4Health (I)

Were selected to

- involve **various user communities** in epidemiology, public health and clinical trials
 - address **most relevant aspects** of personal health-related data and provide blueprints for upscaling
 - **implement services and standards** in major data sources in Germany
 - illustrate approaches to ensure **confidentiality** and to comply with **data protection/data privacy and ethical regulations**
- **pilot FAIRification of individual health data**
- **Task Area 5**

Use Cases: one important component on NFDI4Health (II)



The Use Case Task Force COVID-19 (I)

- Funded by **Deutsche Forschungsgemeinschaft (German Research Foundation)** as additional use case of NFDI4Health
 - **Duration:** 07.2020 – 06.2023; **total funding volume:** € 1.065.084
 - **Coordination:** Leibniz Institute for Prevention Research and Epidemiology – BIPS
- Establishment of **nationwide information infrastructure** for COVID-19 research
- Focus beyond clinical patient data on **consequences of the pandemic outbreak** for public health
- Participating **institutions:**



Schmidt, Fluck, Golebiewski, Grabenhenrich, Hahn, Kirsten, Klammt Löbe, Sax, Thun, Pigeot (2021) Die NFDI4Health - Task Force COVID-19. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 64:1084-1092

The Use Case Task Force COVID-19 (II)

Additional use case of NFDI4Health, but

- Serves as a kind of **microcosm**
 - Provides **unique opportunity** to learn about potential pitfalls and challenges
 - **Lessons learned** to facilitate necessary steps in NFDI4Health
 - Set-up of services and standards to be tested against other use cases and expanded to cover needs of NFDI4Health
- **“Blue print”** for overall NFDI4Health

Objective I: Improved FAIRness of COVID-19 research data collections

- Inventory of German studies on COVID-19 from epidemiology, public health and clinical trials
- Centralised data catalogue with search function
- Improved access
- Measures for improved data exchange
- Uniform assessments of data quality

Objective II: Promotion of timely implementation of harmonised research related to COVID-19

- Guidelines, training materials and standards e.g. for data management, survey instruments and use of routine data
- Standardised analysis routines for data exploration
- Services for (partially) automated analysis of medical image data

Objective III: Record linkage of various COVID-19 data records

- Concept for [linking various relevant data sources](#)
 - Hospital data
 - Outpatient prescription data
 - Sequencing data
 - Image data
 - Sociodemographic and psychosocial data
- Including [data protection concept](#) and [template for informed consent](#)

Objective IV: Standardisation, search portal and information extraction

- **Standardisation of information to be collected for data publication**
 - Development of [publication guidelines](#) for standardised publication of COVID-19 studies
 - Development of [metadata scheme](#) for collection, detection and publication of COVID-19 studies
 - Collection of terminologies and [development of terminology service](#)
- **Establishment of search portal** for personal health data on COVID-19
 - [Implementation of search functionalities](#) based on developed metadata standards and terminologies
- **Information extraction and text mining**
 - [Enrichment of developed terminology resources](#) with synonyms for optimal use on search portal, for metadata annotations or text mining approaches

First results in numbers

- **Study hub**
 - Studies: 855
- **Data catalogue**
 - **Identified studies¹: 233**
 - Studies that meet the selection criteria of Task Force COVID-19: 105
 - Studies contacted: 96
 - Studies with semantic annotation of survey items: 21
 - **Instruments: 23**
 - **Questionnaire items: 3.506**
- **Text mining-based preprint viewer**
 - COVID-19 related preprints: 37.861
 - **Availability of peer-reviewed versions:**
 - No peer-reviewed article: 29.940
 - Found by Pre2Pub algorithm: 4.232
 - Linked through Crossref/preprint server: 3.689
 - **Semantic classes:**
 - Medical Subject Headings - Diseases: 34.630
 - HUGO Gene Nomenclature Committee: 7.198
 - SARS-CoV-2 proteins: 4.107
 - SARS-CoV-2 virus variants: 235
- **CT analysis tool for COVID-19**
 - Image quality features: 28
 - COVID-19 scores: 2

¹ Active search for further studies was discontinued at the beginning of 2021

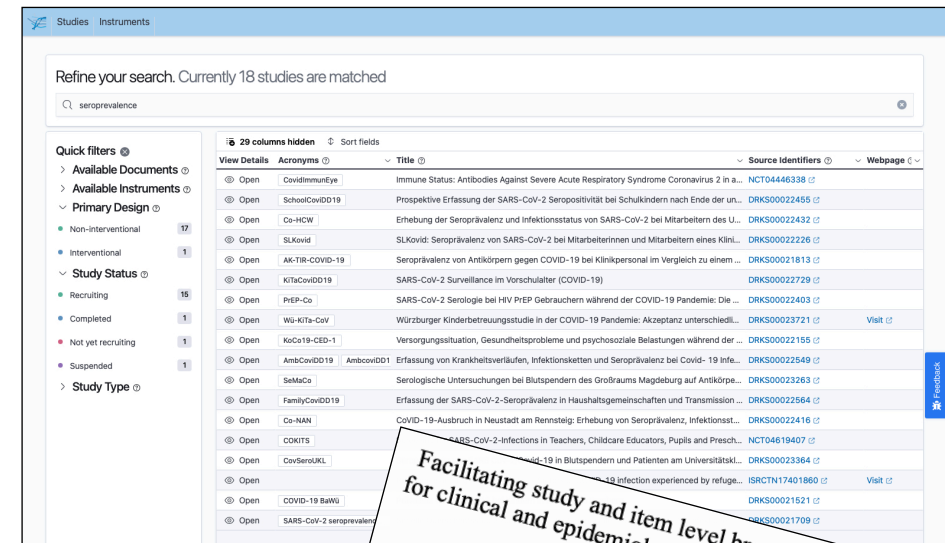
Results: Publication guidelines for German Central Health Study Hub COVID-19

- Definition of [data types/data resources](#):
 - Study metadata
 - Study documents (e.g. study protocol, data dictionary, instruments like templates of questionnaires or case report forms)
- Reference [metadata scheme](#) developed for recording COVID-19 studies and study documents
- Assignment of [persistent identifier: DOI](#) (Digital Object Identifier) for [study documents](#)
- [Licenses](#) to enable subsequent use of content: preferably [Creative Commons licenses](#)
- [File formats](#): [machine-readable](#) preferred, e.g. .xml, .csv, .xls, .json, .rdf; human-readable also accepted.
- [Language\(s\)](#): English preferred, German accepted

➤ **"Blue print"** for NFDI4Health

Results: German Central Health Study Hub COVID-19

- Access: <https://covid19.studyhub.nfdi4health.de>
- Advanced search and filter functions for:
 - Metadata of studies, instruments and documents
 - "Questionnaire items" of instruments
- Hierarchical display of studies and their documents and instruments
- Version management of documents possible
- Integrates studies from
 - Registries such as DRKS, WHO/ITRCP, ClinicalTrials.gov
 - Data cataloguing by the Task Force
- Weekly updates



View Details	Acronyms	Title	Source Identifiers	Webpage
Open	CovidImmuneEye	Immune Status: Antibodies Against Severe Acute Respiratory Syndrome Coronavirus 2 in a...	NCT04446338	
Open	SchoolCovid19	Prospektive Erfassung der SARS-CoV-2 Seropositivität bei Schulkindern nach Ende der un...	DRKS00022456	
Open	Co-HCW	Erhebung der Seroprävalenz und Infektionsstatus von SARS-CoV-2 bei Mitarbeitern des U...	DRKS00022432	
Open	SLKovid	SLKovid: Seroprävalenz von SARS-CoV-2 bei Mitarbeiterinnen und Mitarbeitern eines Klin...	DRKS00022226	
Open	AK-TIR-COVID-19	Seroprävalenz von Antikörpern gegen COVID-19 bei Klinikpersonal im Vergleich zu einem ...	DRKS00021813	
Open	KiTaCovid19	SARS-CoV-2 Surveillance im Vorschulalter (COVID-19)	DRKS00022729	
Open	PfEP-Co	SARS-CoV-2 Serologie bei HIV PfEP Gebräuchern während der COVID-19 Pandemie: Die ...	DRKS00022403	
Open	Wu-KiTa-Cov	Würzburger Kinderbetreuungsstudie in der COVID-19 Pandemie: Akzeptanz unterschiedli...	DRKS00023721	Visit
Open	KoCo19-CED-1	Versorgungssituation, Gesundheitsprobleme und psychosoziale Belastungen während der ...	DRKS00022155	
Open	AmbCovid19	Erfassung von Krankheitsverläufen, Infektionsketten und Seroprävalenz bei Covid-19 Infe...	DRKS00022549	
Open	SeMaCo	Serologische Untersuchungen bei Blutspendern des Großraums Magdeburg auf Antikörper...	DRKS00023263	
Open	FamilyCovid19	Erfassung der SARS-CoV-2-Seroprävalenz in Haushaltsgemeinschaften und Transmission...	DRKS00022564	
Open	Co-NAN	COVID-19-Ausbruch in Neustadt am Remsteig: Erhebung von Seroprävalenz, Infektionsst...	DRKS00022416	
Open	COKITS	SARS-CoV-2-Infektionen in Teachers, Childcare Educators, Pupils and Presch...	NCT04619407	
Open	CovSerouKL	SARS-CoV-2-Infektionen in Blutspendern und Patienten am Universitätsk...	DRKS00023364	
Open	COVID-19 BaWu	COVID-19 infection experienced by refuge...	ISRCTN17401860	Visit
Open	SARS-CoV-2 seroprevalence		DRKS00021521	
Open			DRKS00021709	

Facilitating study and item level browsing for clinical and epidemiological COVID-19 studies

Carsten Oliver SCHMIDT^{1,2}, Johannes DARMS³, Aliaksandra SHUTSKO⁴, Matthias LOEBE⁵, Rajini NAGRANI⁶, Bastian SEIFERT⁷, Birte LINDSTÄDT⁸, Martin GOLEBIEWSKI⁹, Sofiya KOLEVA¹⁰, Theresa BENDER¹¹, Christian Robert BAUER¹², Ulrich SAX¹³, Xiaoming HU¹⁴, Michael LIESER¹⁵, Vivien JUNKER¹⁶, Sophie KLOPFENSTEIN¹⁷, Atinkir ZELEKE¹⁸, Dagmar WALTERMATH¹⁹, Iris PIGEOT²⁰, Juliane FLUCK²¹, on behalf of the NFDI4Health Task Force COVID-19

¹University Medicine of Greifswald, ²ZB MED – Information Centre for Life Sciences, ³University of Leipzig, ⁴Leibniz Institute for Prevention Research and Epidemiology - BIPS, ⁵Heidelberg Institute for Theoretical Studies, ⁶McGill University Health Centre, ⁷University Medical Center/Gütingen, ⁸Berlin Institute of Health

Abstract. COVID-19 poses a major challenge to individuals and societies around the world. Yet, it is difficult to obtain a good overview of studies across different medical fields of research such as clinical trials, epidemiology, and public health. Here, we describe a consensus metadata model to facilitate structured searches of COVID-19 studies and resources along with its implementation in three linked complementary web-based platforms. A relational database serves as central study metadata hub that secures compatibility with common trial registries (e.g. ICTR) and standards like HL7 FHIR, CDISC ODM, and DataCite). The Central Search Hub was developed as a single-page application, the other two components study platforms have different features concerning cohort browsing, item browsing, and access to documents and other study resources to meet divergent user needs. These we want to promote transparent and harmonized COVID-19 research.

Keywords. Browsing metadata, COVID-19, FAIR data, metadata standards

Schmidt et al., on behalf of the NFDI4Health Task Force COVID-19 (2021) Facilitating study and item level browsing for clinical and epidemiological COVID-19 studies. *Stud Health Technol Inform* 281:794-798

Results: Data catalogue

- Epidemiological survey instruments on COVID-19 in Germany
- Access: <https://mica.covid19.studyhub.nfdi4health.de/>

Below is a list of studies and resources related to COVID-19, which contain survey items and item banks. More COVID-19 studies without semantic annotation of survey items can be accessed through: <https://covid19.studyhub.nfdi4health.de/>.

Studies / Resources

- CILS4EU**
Children of Immigrants Longitudinal Survey in Four European Countries
The Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) project focuses on the integration of children of immigrants in four European countries: Germany, the ...
[Read more](#)
- COPE**
Studie zum Umgang der Allgemeinbevölkerung mit COVID-19/SARS-CoV-2
The study investigates the psychological burden of the general population in a large German sample (n= 23,000) in terms of generalized anxiety, depression, distress, locus of control, personality ...
[Read more](#)
- CORONA-MONITORING lokal**
Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in Deutschland - Studienprotokoll von CORONA-MONITORING lokal
The study CORONA-MONITORING lokal is a population-based seroepidemiological observational study, in which repeated cross-sectional examinations with independent samples in each case are to be ...
[Read more](#)
- COVID-19 CRF**
CORE COVID-19 Case Report Form
ISARIC's (International Severe Acute Respiratory and emerging Infection Consortium) purpose is to prevent illness and deaths from infectious disease outbreaks. ISARIC is a global federation of ...
[Read more](#)

Below please find a list of COVID-19 instruments.
After selecting an instrument, a brief description is provided along a graphical overview on targeted areas. When selecting "Questionnaire items", a searchable list of items appears.

Instruments

- CILS4COVID**
COVID-19 Add-On (CILS4COVID) Field Questionnaire
[Read more](#)
- Cope with Corona**
Studie zum Umgang der Allgemeinbevölkerung mit COVID-19/SARS-CoV-2 (Corona-Virus) / Cope with Corona
[Read more](#)
- CORONA-MONITORING lokal**
Kurzfragebogen
[Read more](#)
- COVID-19 CRF**
COVID-19 core case report form
[Read more](#)
- Covid Mental health**
Die COVID-19 Umfrage Deutschland: Seelische Gesundheit in Krisenzeiten
[Read more](#)
- CRISIS**
Adult Self-Report Baseline Current Form
[Read more](#)
- CRISIS**
Youth Self-Report Baseline Form: Current Form
[Read more](#)
- DynaCORE-C**
DynaCORE Questionnaire-C
[Read more](#)
- Essen-Covid**
Fragebogen für Corona-gefährdete Bürgerinnen und Bürger der Stadt Essen
[Read more](#)

NFDI4Health - COVID-19 Questionnaires

The NFDI4Health Taskforce COVID-19 is a platform for centralized data [cataloging](#) of epidemiological studies on COVID-19 conducted in Germany.

Beta Version 0.9

This website provides semantic search functionalities for selected COVID-19 survey instruments and item banks of relevance for epidemiologic and public health studies. The focus is on German studies. Information on these and other studies can be accessed through <http://covid19.studyhub.nfdi4health.de/>. This site also provides instruments for download in their original formatting.

21 Studies / Resources [More info](#)

23 Instruments [More info](#)

3.506 Questionnaire items [More info](#)

Search strategies:

A. Obtain an overview on studies and resources
Please use the "More Info" button in the "Studies/Resources" field. Resources other than studies are sites that provide guidance on items of relevance for COVID-19 surveys.

B. Search for a particular COVID-19 instrument
Please use the "More Info" button in the "Instruments" field. Subsequently, a list of selected instruments is presented. After selecting an instrument, a brief description is provided along a graphical overview on targeted areas. When selecting "Questionnaire items", a searchable list of items appears.

C. Obtain an overview on items that target a defined target area
Please use the "More Info" button in the "Questionnaire items" field. All items of all questionnaires are displayed. By using the left column "Search criteria", properties of items may be selected.

[Search](#)

Results: Study document and metadata catalogue

- COVID-19 study metadata capturing, as well as document and instrument browsing based on the widely used FAIRDOME SEEK system for scientific data management
- Access (test implementation): <https://seek.covid19.studyhub.nfdi4health.de/>



Home / Documents Index
Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in Deutschland - Studienprotokoll von CORONA-MONITORING lokal

Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in Deutschland - Studienprotokoll von CORONA-MONITORING lokal Version 1

Request Contact | Unsubscribe | Download | View content | Add | Act

Derzeit erkranken viele Menschen in Deutschland an der Atemwegserkrankung COVID-19. Das neuartige Coronavirus SARS-CoV-2 löst diese Erkrankungen aus. Mit der Studie CORONA-MONITORING lokal wird in vier von COVID-19 besonders betroffenen Regionen in Deutschland (Bayern, Straubing (Bayern)) sowie Berlin-Mitte, insgesamt sollen 8.000 Erwachsene teilnehmen können aus wissenschaftlichen Gründen ausschließlich Menschen, die vom Robert Koch-Institut nach den Vorgaben des Bundesdatenschutzgesetzes durchgeführt.

SEEK ID: <https://seek.studyhub.nfdi4health.de/documents/15?version=1>

Filename: 134_RKI Corona-Monitoring lokal Questionnaire.pdf

Format: PDF document

Size: 33.2 KB

Selected: Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in... (Docu Tree) | Split | Fullscreen

COVID-19 Studies in Germany

Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in Deutschland

Studienprotokoll von CORONA-MONITORING lokal

Seroepidemiologische Studie zur Verbreitung von SARS-CoV-2 in der Bevölkerung an besonders betroffenen Orten in Deutschland - Stu

Activity

Views: 59 Downloads: 0

Created: 17th Dec 2020 at 13:48
Last updated: 26th Feb 2021 at 12:57
Last used: 22nd Mar 2021 at 11:21

Tags

Germany questionnaire stud instrument templates

Home / Documents Index / Children of Immigrants Longitudinal Survey in Four European Countries - Germany. COVID-19 Add-On (CILS4COVID). Field Questionnaire

Children of Immigrants Longitudinal Survey in Four European Countries - Germany. COVID-19 Add-On (CILS4COVID). Field Questionnaire Version 2

This document contains the field questionnaires used in the CILS4COVID survey. Since the CILS4COVID survey included a survey experiment, there are two versions of the questionnaire that only differ in their question order. Both versions of the postal questionnaire are presented separately in the document. The web questionnaire contains both versions and documents the filters used to assign participants to one of the two experimental groups.

SEEK ID: <https://seek.studyhub.nfdi4health.de/documents/13?version=2>

Filename: CILS4COVID_quest_N4H.pdf

Format: PDF document

Size: 709 KB

Selected: Children of Immigrants Longitudinal Survey in Four European Countries - Germany. COVID-19 Add-On (CILS4COVID). Field... (Docu Tree) | Split | Fullscreen

Description: This document contains the field questionnaires used in the CILS4COVID survey. Since the CILS4COVID survey included a...

SEEK ID: <https://seek.studyhub.nfdi4health.de/documents/13>

Children of Immigrants Longitudinal Survey in Four European Countries - Germany. COVID-19 Add-On (CILS4COVID). Field Questionnaire

Browse + Create Help Search here... Search

New Study

The fields marked with ** are required before you can save the resource.
The fields marked with * are required before you can submit the resource.
You are able to save the resource first with essential data and complete it later with details.

Basic Information: -

Resource Roles: -

Associated Resources: -

Study Design General: -

Study primary design *

✓ Please select...

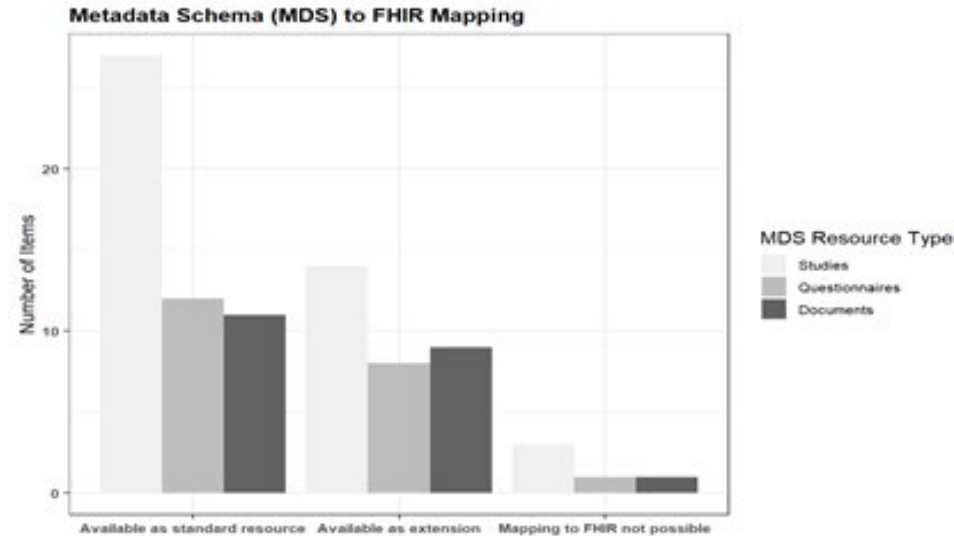
Non-interventional

Interventional

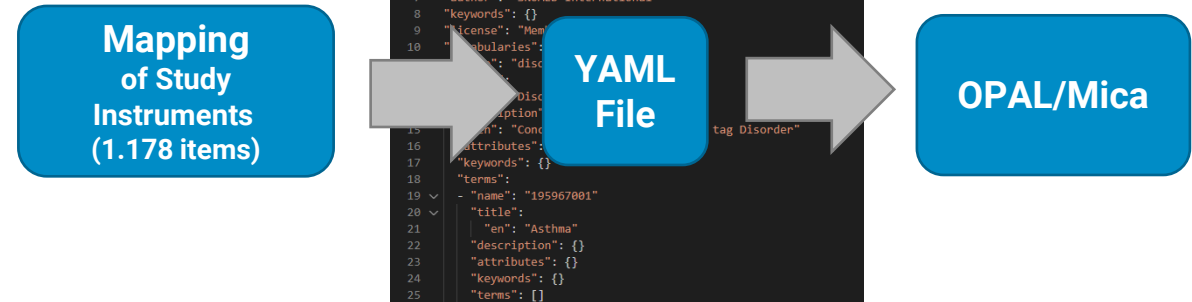
Study type description

Results: Interoperability – Implementing international standards

FHIR mapping of metadata schema



Mapping and implementing SNOMED CT in OPAL/Mica



- Majority of metadata items can be mapped with FHIR, but FHIR extensions necessary
- Currently, mapping of SNOMED CT conducted for eight instruments in OPAL/Mica
- SNOMED CT successfully implemented in Mica/OPAL using YAML format and python scripts
 - SNOMED CT hierarchy and grammar cannot be implemented in current Mica/OPAL format
 - On the long run: Mica/OPAL to be adapted for optimal implementation of international terminologies

YAML = YAML Ain't Markup Language | FHIR = Fast Healthcare Interoperability Resources

Results: Text mining-based Preprint Viewer

- Access: <https://preview.zbmed.de/>
- Advanced search and filter functions for abstracts
- Direct links to full texts
- Export functions for retrieved results
- Integrates preprints from the following document servers
 - medRxiv and bioRxiv | ChemRxiv | ResearchSquare | arXiv | Preprints.org
- Daily updates

The screenshot displays the preVIEW: COVID-19 web application. At the top, there is a navigation bar with 'preVIEW: COVID-19', 'ABOUT', 'TUTORIAL', and 'CONTACT'. Below this, a ZB MED logo is visible. The main content area features a search bar with a 'Multi' dropdown and a search button. A 'Publication Date' histogram shows the distribution of documents from April 2020 to May 2021. A search result is highlighted, showing the title 'COVID-19 preVIEW: Semantic Search to Explore COVID-19 Research Preprints' and the authors: Lisa LANGNICKEL, Roman BAUM, Johannes DARMS, Sumit MADAN, and Juliane FLUCK. The abstract text is visible, discussing the rapid availability of COVID-19 information and the design of the preVIEW search engine. The interface also includes options to 'enable feedback mode', 'Expand all abstracts', and 'Highlight semantic concepts'.

Langnickel, Baum, Darms, Madan, Fluck (2021) COVID-19 preVIEW: Semantic Search to Explore COVID-19 Research Preprints. Stud Health Technol Inform 281:78-82.

Results: Record Linkage

- Overview of different linking approaches with and without unique identifiers
 - In Germany: no unique identifier for medical research
 - Various personal identifiers used in medical research, partly even none precluding record linkage
 - Recommendation for personal identifiers: Core data set of Medical Informatics Initiative
- Good practice data linkage (GPD): recommendations for linkage processes
 - Application of GPD in current studies, e.g., NAKO, SHIP, LIFE, ...
 - MII concept in line with GPD
- Ongoing: Privacy preserving rules

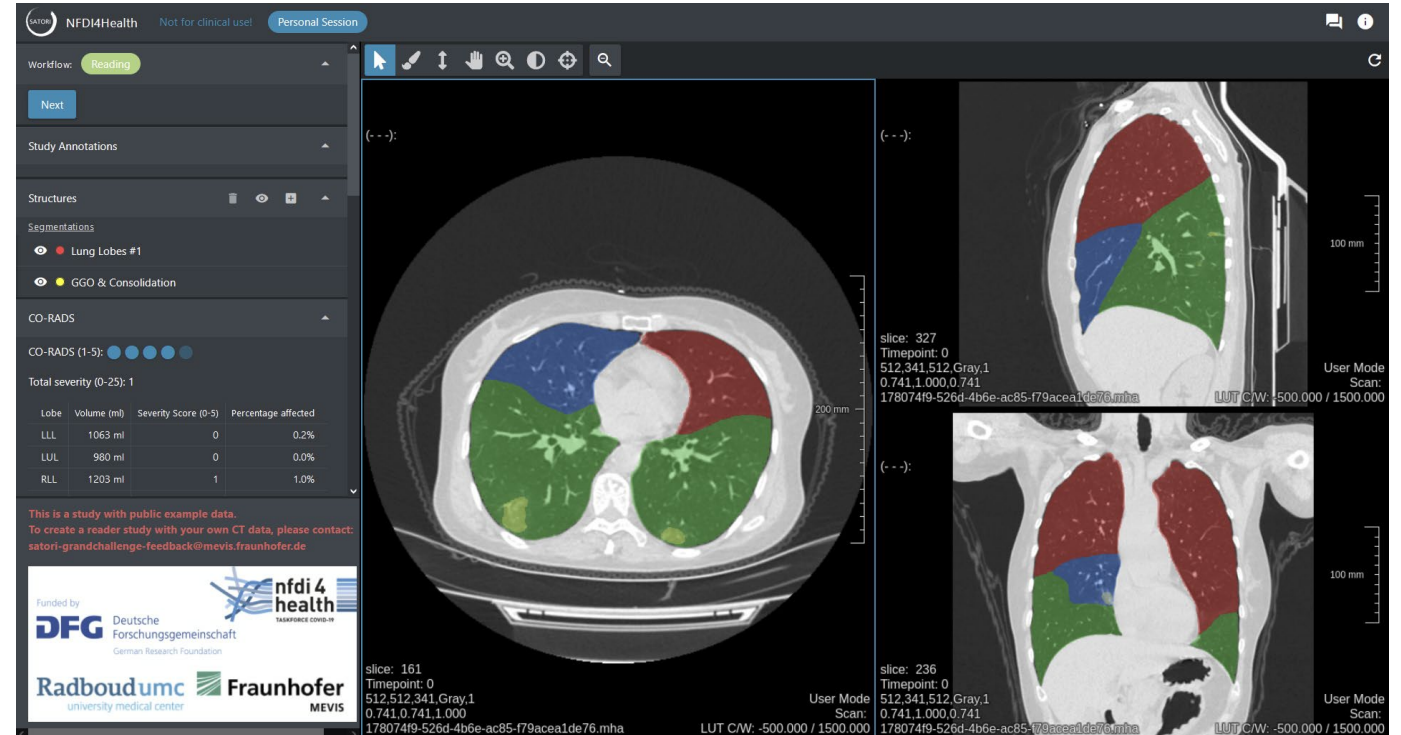
Results: Providing an AI-based image analysis service for CT data

- Publically available fully automatic CT analysis tool for COVID-19

- Automatic segmentation
- Calculation of disease-specific scores
- Interactive viewing components

- Information gain:

- Probability of COVID-19
- Severity
- Lung volume
- Image quality



<https://grand-challenge.org/reader-studies/satori-nfdi4health-test/>

**Thank you very much for
your attention!**



Contact:

Prof. Dr. Iris Pigeot

pigeot@leibniz-bips.de

**Leibniz Institute
for Prevention Research and
Epidemiology – BIPS**
Achterstr. 30, 28359 Bremen

Prof. Dr. Juliane Fluck

fluck@zbmed.de

**ZB MED - Information Centre
for Life Sciences**
Gleueler Straße 60, 50931 Köln

www.nfdi4health.de



Gefördert durch
DFG Deutsche
Forschungsgemeinschaft

