

# A prototype platform for mapping heterogeneous metadata of the three domains health, environment and earth observation: the MetaMap<sup>3</sup> project

**DLR:** Jeroen Staab, Hannes Taubenböck

**UFZ:** Ronny Gey, Jan Bumberger

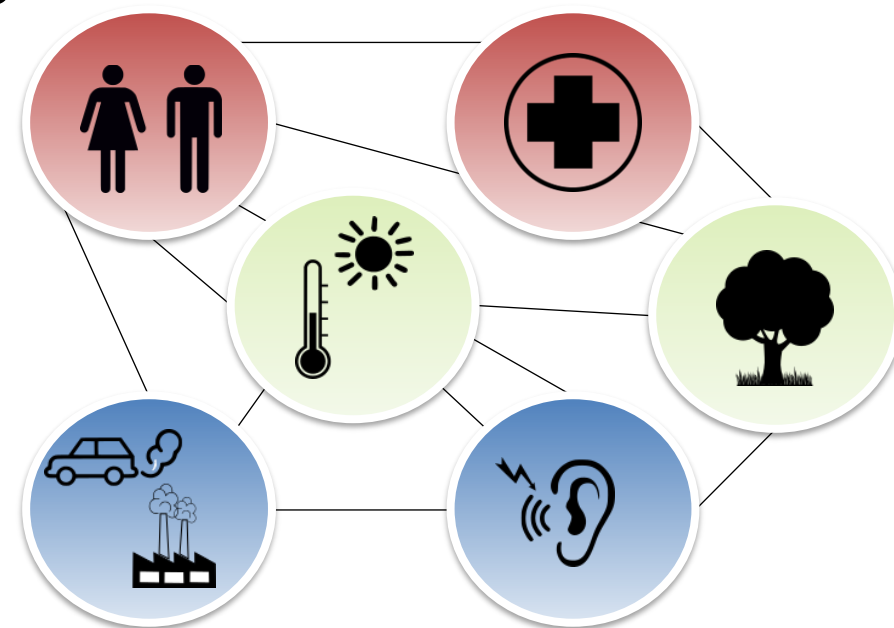
**HMGU:** Marco Dallavalle, Marie Standl, [Kathrin Wolf](#)

# Background and Motivation

- **Environment** plays an increasingly important role for **human health**
- Environmental epidemiology:
  - Quantify human exposures and assess health effects by linking **health** data from cohorts with **environmental** and **earth observation** data
- Efficient linkage is crucial, esp. with increasing amount of data
- Harmonized metadata standards that allow routinely merging are lacking
- **Health domain**: Most schemas and standards clinically oriented and cannot be directly transferred to epidemiological cohorts

# Objectives

- Generation and enrichment of interoperable metadata for exemplary data of our 3 domains
  - Health (HMGU)
  - Earth & Environment (UFZ)
  - Aeronautics, Space & Transport (DLR)
- Mapping of the respective metadata to be jointly queried and searched
- Integration into HMC framework



# AST: DLR

# DLR use case: land cover

- German-wide data for 2015
- Available at DLR's Earth Observation Center
  - NFDI4Earth, GDI-DE, INSPIRE
  - API's to Geoportal (Bund), GEOSS (NASA),...
  - Based on GeoNetwork (OpenSource & Standards)

**Already FAIR !**

The screenshot shows the EOC Catalogue interface for the 'Land Cover DE - Sentinel-2 - Germany, 2015' dataset. The page includes a description of the land cover classification, download and link options, and a spatial extent map.

**Land Cover DE - Sentinel-2 - Germany, 2015**

This land cover classification of Germany was created using Sentinel-2 imagery from the years 2015 to 2017 and LUCAS 2015 in-situ reference data (<https://ec.europa.eu/eurostat/web/lucas>). It contains seven land cover types: (1) artificial land, (2) open soil, (3) high seasonal vegetation, (4) high perennial vegetation, (5) low seasonal vegetation, (6) low perennial vegetation and (7) water with a spatial resolution of 10m x 10m. For further information, please see the following publication: <https://doi.org/10.1016/j.jag.2020.102065>

**Download and links**

Icon	Link/Action	Description
WMS icon	<a href="#">Add to map</a>	LCC_DE_2015 Land Cover DE, 2015 This dataset is published in the view service (WMS) available at <a href="https://geoservice.dlr.de/eoc/land/wms?with_layer_name=LCC_DE_2015">https://geoservice.dlr.de/eoc/land/wms?</a> with layer name LCC_DE_2015.
Link icon	<a href="#">Open link</a>	EOC Geoservice Map Context (de:lcc) EOC Geoservice Map Context <a href="https://geoservice.dlr.de/web/maps/de:lcc">https://geoservice.dlr.de/web/maps/de:lcc</a>
Link icon	<a href="#">Open link</a>	HTTP Download (Land Cover DE) HTTP Download <a href="https://download.geoservice.dlr.de/LCC_DE/files/">https://download.geoservice.dlr.de/LCC_DE/files/</a>
Link icon	<a href="#">Open link</a>	Scientific publication describing the Land Cover DE product Scientific publication <a href="https://doi.org/10.1016/j.jag.2020.102065">https://doi.org/10.1016/j.jag.2020.102065</a>

**About this resource**

**Categories** [Oberflächenbeschreibung](#)

**Keywords**

- Sentinel-2 [Q](#)
- Land cover map [Q](#)
- Germany [Q](#)
- opendata [Q](#)

**GEMET - INSPIRE themes, version 1.0**

- Land cover [Q](#)

**Spatial extent**

The spatial extent map shows the land cover data for Germany, with a yellow box indicating the area of interest.



# DLR use case: metadata standard

- ISO-19115
- XML File

```
<gmd:CI_PresentationFormCode codeListValue="mapDigital" codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/codelist/ML_gmxCodeLists.xml#CI_PresentationFormCode"/>
</gmd:presentationForm>
<gmd:otherCitationDetails>
  <gco:CharacterString>DOI: 10.15489/1ccmlap3mn39</gco:CharacterString>
</gmd:otherCitationDetails>
</gmd:CI_Citation>
</gmd:citation>
<gmd:abstract>
  <gco:CharacterString>
    This land cover classification of Germany was created using Sentinel-2 imagery from the years 2015 to 2017 and LUCAS 2015 in-situ reference data (https://ec.europa.eu/eurostat/web/lucas). It contains seven land cover types: (1) artificial land, (2) open soil, (3) high seasonal vegetation, (4) high perennial vegetation, (5) low seasonal vegetation, (6) low perennial vegetation and (7) water with a spatial resolution of 10m x 10m. For further information, please see the following publication: https://doi.org/10.1016/j.jag.2020.102065
  </gco:CharacterString>
</gmd:abstract>
</gmd:status>
<gmd:MD_ProgressCode codeListValue="completed" codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/codelist/ML_gmxCodeLists.xml#MD_ProgressCode"/>
</gmd:status>
<gmd:pointOfContact>
  <gmd:CI_ResponsibleParty>
    <gmd:individualName gco:nilReason="missing">
      <gco:CharacterString/>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>German Aerospace Center (DLR)</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:electronicMailAddress>
              <gco:CharacterString>matthias.weigand@dlr.de</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
      </gmd:CI_Contact>
    </gmd:contactInfo>
  </gmd:CI_ResponsibleParty>
</gmd:pointOfContact>
</gmd:MD_Metadata>
```

## About this resource

Categories	<a href="#">Oberflächenbeschreibung</a>
Keywords	<ul style="list-style-type: none"><li>• Sentinel-2 <a href="#">Q</a></li><li>• Land cover map <a href="#">Q</a></li><li>• Germany <a href="#">Q</a></li><li>• opendata <a href="#">Q</a></li></ul>
GEMET - INSPIRE themes, version 1.0	<ul style="list-style-type: none"><li>• Land cover <a href="#">Q</a></li></ul>
Language	<ul style="list-style-type: none"><li>• English</li></ul>
Resource identifier	<ul style="list-style-type: none"><li>• <a href="https://geoservice.dlr.de/catalogue/srv/metadata/8205e1c9-14c3-47a7-ac12-8c890225680a">https://geoservice.dlr.de/catalogue/srv/metadata/8205e1c9-14c3-47a7-ac12-8c890225680a</a></li></ul>
Legal constraints	<p>License</p> <p>Nutzungsbedingungen: Lizenz, <a href="https://creativecommons.org/licenses/by-nc/4.0/">https://creativecommons.org/licenses/by-nc/4.0/</a> / Terms of use: License, <a href="https://creativecommons.org/licenses/by-nc/4.0/">https://creativecommons.org/licenses/by-nc/4.0/</a></p> <p>{ "id": "cc-by-nc/4.0", "name": "Creative Commons Namensnennung - Nicht kommerziell 4.0 International (CC BY-NC 4.0)", "url": "http://dcat-ap.de/def/licenses/cc-by-nc/4.0", "quelle": "Copyright DLR (year of production)" }</p> <p>Nutzungseinschränkungen: Das DLR ist nicht haftbar für Schäden, die sich aus der Nutzung ergeben. / Use Limitations: DLR not liable for damage resulting from use.</p> <p>Nutzungsbedingungen: Lizenz, <a href="https://creativecommons.org/licenses/by-nc/4.0/">https://creativecommons.org/licenses/by-nc/4.0/</a> / Terms of use: License, <a href="https://creativecommons.org/licenses/by-nc/4.0/">https://creativecommons.org/licenses/by-nc/4.0/</a></p>
Contact for the resource	<p>✉ German Aerospace Center (DLR)</p> <ul style="list-style-type: none"><li>• Author : <a href="mailto:matthias.weigand@dlr.de">matthias.weigand@dlr.de</a> , <a href="mailto:michael.wurm@dlr.de">michael.wurm@dlr.de</a></li><li>• Point of contact : <a href="mailto:geoservice@dlr.de">geoservice@dlr.de</a></li></ul>
Status	<ul style="list-style-type: none"><li>• Completed</li></ul>
Technical information	
Update frequency	As needed
Representation type	Grid
Scale	<ul style="list-style-type: none"><li>• 20000</li></ul>
Coordinate reference system	<ul style="list-style-type: none"><li>• EPSG:ETRS89 / LAEA Europe (EPSG:3035)</li></ul>
Lineage	Input data: Sentinel-2 multispectral imagery, LUCAS in-situ reference points, and imperviousness layers from OSM and the LoD1-de building model.

## Metadata information

[Download metadata](#)

Contact	<p>✉ German Aerospace Center (DLR)</p> <ul style="list-style-type: none"><li>• Point of contact : <a href="mailto:geoservice@dlr.de">geoservice@dlr.de</a></li></ul>
Metadata language	<ul style="list-style-type: none"><li>• English</li></ul>
Identifier	8205e1c9-14c3-47a7-ac12-8c890225680a

## Spatial extent



## Temporal extent

Creation date  
2020-02-05

Period  
Sat Jun 27 2015 10:25:00 GMT+0200 ➡ Fri Sep 29 2017 10:20:00 GMT+0200

Provided by



Updated:  
vor einem Jahr

Share on social sites

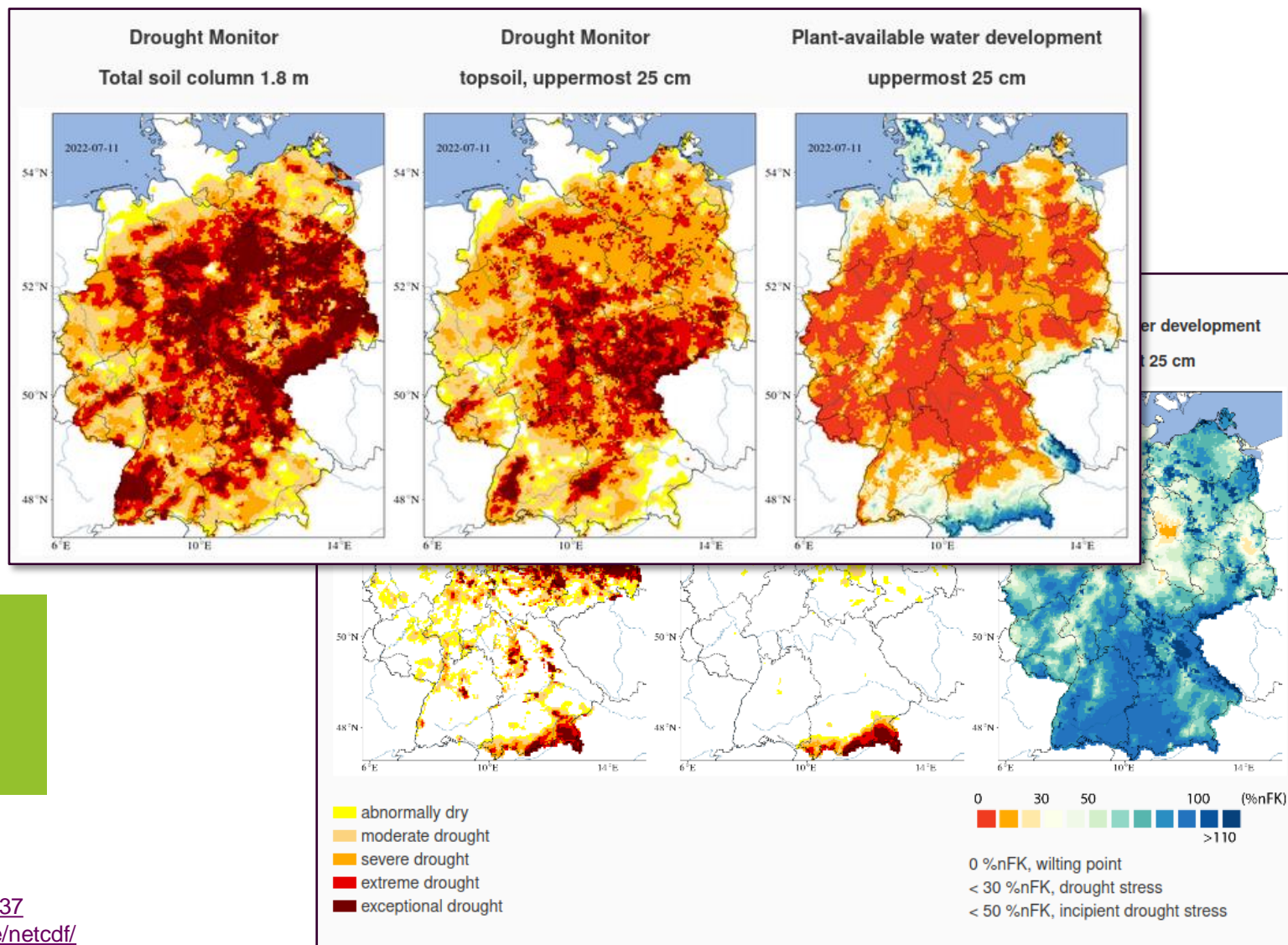


# E&E: UFZ



# UFZ use case: Drought Monitor

- Daily information on drought and soil moisture, several parameters:
- Topsoil moisture, total soil moisture, plant-available water
- Availability: **daily, 1951 - present**
- Format: **netCDF**  
(Network Common Data Form)
- Several parameters at several locations at several timepoints



FA, only partly: IR



# UFZ use case: CF Metadata Conventions

```
netcdf SM_Lall_daily_n14 {
  dimensions:
    easting = 175 ;
    northing = 225 ;
    time = UNLIMITED ; // (14 currently)

  variables:
    float SMI(time,northing,easting) ;
      SMI:long_name = "monthly soil moisture index" ;
      SMI:units = "-" ;
      SMI:coordinates = "lat lon" ;
      SMI:_FillValue = -9999.f ;
      SMI:missing_value = -9999.f ;

    double easting(easting) ;
      easting:standard_name = "projection_x_coordinate" ;
      easting:long_name = "x-coordinate in cartesian coordinates GK4" ;
      easting:units = "m" ;
      easting:axis = "X" ;

    double lat(northing,easting) ;
      lat:standard_name = "latitude" ;
      lat:long_name = "latitude" ;
      lat:units = "degrees_north" ;
      lat:_CoordinateAxisType = "Lat" ;

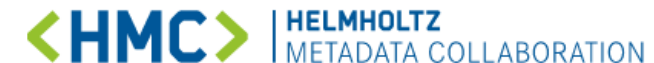
    double lon(northing,easting) ;
      lon:standard_name = "longitude" ;
      lon:long_name = "longitude" ;
      lon:units = "degrees_east" ;
      lon:_CoordinateAxisType = "Lon" ;

    double northing(northing) ;
      northing:standard_name = "projection_y_coordinate" ;
      northing:long_name = "y-coordinate in cartesian coordinates GK4" ;
      northing:units = "m" ;
      northing:axis = "Y" ;

    double time(time) ;
      time:standard_name = "time" ;
      time:long_name = "time" ;
      time:units = "days since 2016-1-30 00:00:00" ;
      time:calendar = "standard" ;
      time:axis = "T" ;

  // global attributes:
  :NCO = "netCDF Operators version 4.9.8 (Homepage = http://nco.sf.net, Code = http://-
github.com/nco/nco)" ;
  :Conventions = "CF-1.8" ;
  :institution = "Helmholtz Center for Environmental Research" ;
  :originator = "UFZ Drought Monitor" ;
  :contact = "klima@ufz.de" ;
  :crs = "EPSG:31468" ;
  :source = "UFZ Drought Monitor / Helmholtz Centre for Environmental Research" ;
  :creation_date = "2022-07-13" ;
}
```

- Conventions for CF (Climate and Forecast) metadata
- Designed for processing and sharing of files created with **netCDF** API
- **Standard-like status** (e.g. in atmospheric and coastal research)
- **No** established **convention** for netCDF headers at UFZ (in progress)
- Geolocation not fully written into header
- Convert into xml format using e.g. **R geometa**



# HEALTH: HMGU

# HMGU use case: Birth Cohorts

## GINIplus

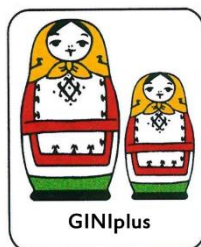
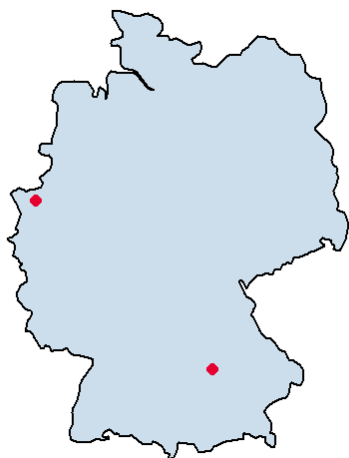
**Population-based**

N=5991

**Recruitment**

1995 – 1998

Munich and Wesel



## LISA

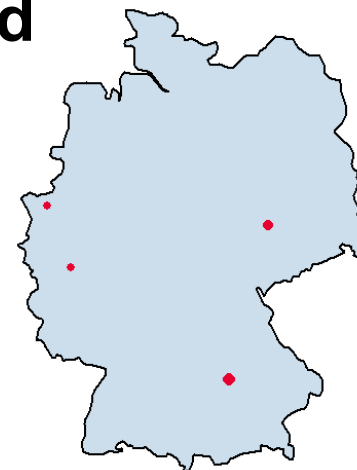
**Population-based**

N=3094

**Recruitment**

1997-1999

Munich, Wesel,  
Leipzig, Bad  
Honorf





# HMGU use case: GINIplus

partly: FA, but not: IR

Metadata stored in .xlsx  
(questionnaire + measurements):  
115 Excel sheets, N of v

Follow-up timepoints:

Birth	
<i>Intervention</i>	
München:	1,165
Wesel:	1,087
<i>Observation</i>	
München:	1,784
Wesel:	1,955
<b>Total:</b>	<b>5,991</b>

1 year	2 years	3 years	4 years
Total: 4,671	Total: 4,354	Total: 4,194	Total: 4,011

F: via website  
A: contact PI

→ Improve FAIRness!

GINIplus, n=5,991

Birth 1 2 3 4 6 10

including  
physical examination:

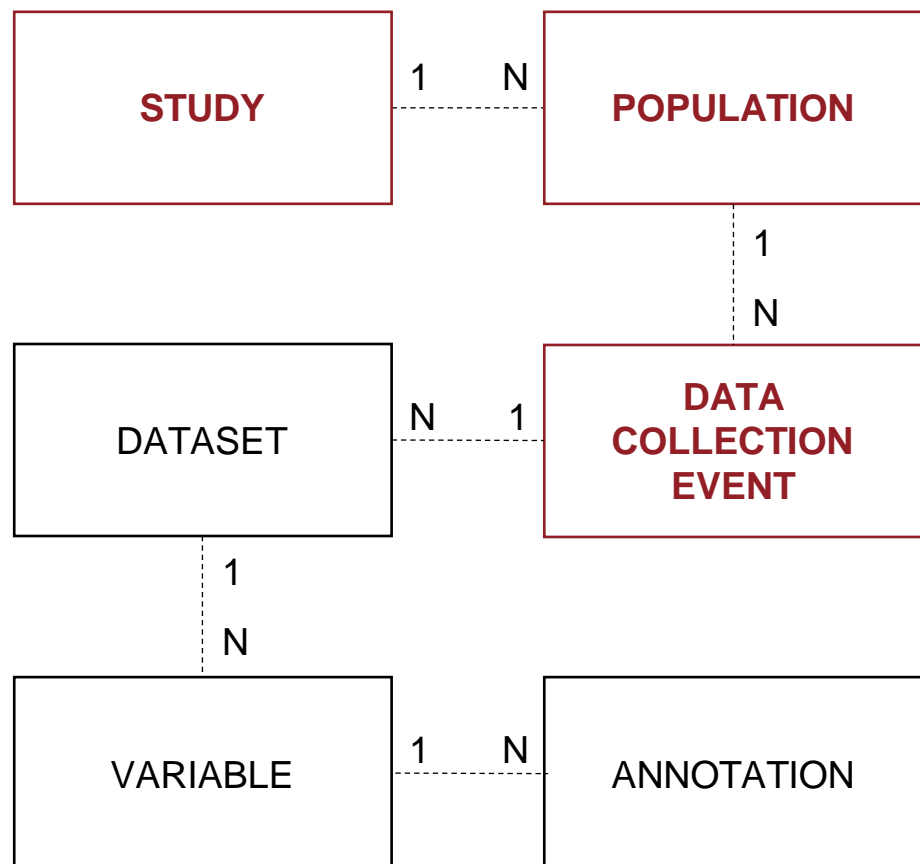
6 years  
Total : 3,878

15 years  
Total: 3,198

10 years  
Total : 3,317




25 years  
Ongoing

# EPI cohorts: Identify standards & services



Adapted from Bergeron J et al.  
<https://doi.org/10.1371/journal.pone.0200926>

- High Granularity of cohort metadata
- 1:N relationship of data and its metadata

Granularity	Catalog / platform	Standard / schema
Study Population Data collection event Dataset Variable	<b>NFDI4Health StudyHub</b>	NFDI4Health Metadata Schema DOI: 10.4126/FRL01-006431357  Maelstrom 
Study Data collection event Dataset Variable	<b>Geonetwork</b>	ISO19115   Maelstrom 

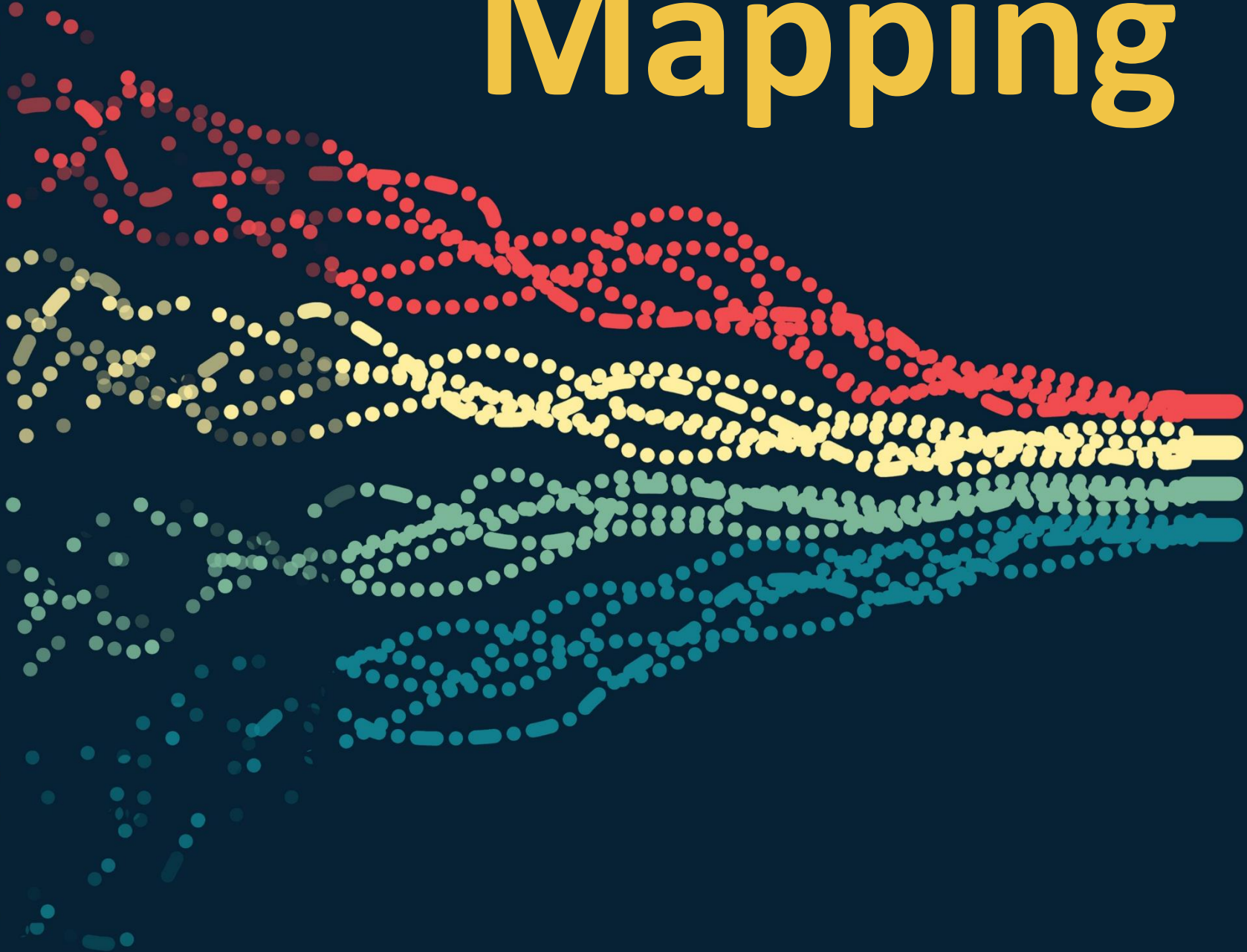
# Mapping

0 1 0 0 1 0 0 0 1 1 1 1  
1 0 1 1 0 1 0 0 1 0 1 0 1 0 0 0 0 1 0  
1 0 0 0 0 0 1 0 1 0 1 0 1 1 0 1 1 1  
0 0 1 1 1 0 1 1 1 0 1 0 0 0 1 1 0 1 0 0  
1 1 0 0 0 1 1 0 0 1 1 1 0 0 0 0 1 1 0 1

0 1 0 0 1 0 0 0 0 1 1 1 1  
1 1 1 1 1 1 0 1 1 1 1 0 0 1 1  
1 0 0 0 0 1 0 1 0 0 1 0 1 1 0 1 1 1 1  
0 0 1 1 1 0 1 1 1 0 1 0 0 0 1 1 0 1 0 0  
1 1 0 0 0 1 1 0 0 1 1 1 0 0 0 0 1 1 0 1

0 1 0 0 1 0 0 0 0 1 1 1 1  
1 1 1 1 1 1 0 1 1 1 1 1 0 0 1  
1 0 0 0 0 1 0 1 0 0 1 0 1 1 0 1 1 1  
1 0 1 0 1 1 0 1 0 1 1 0 0 1 0 1 0 0 1

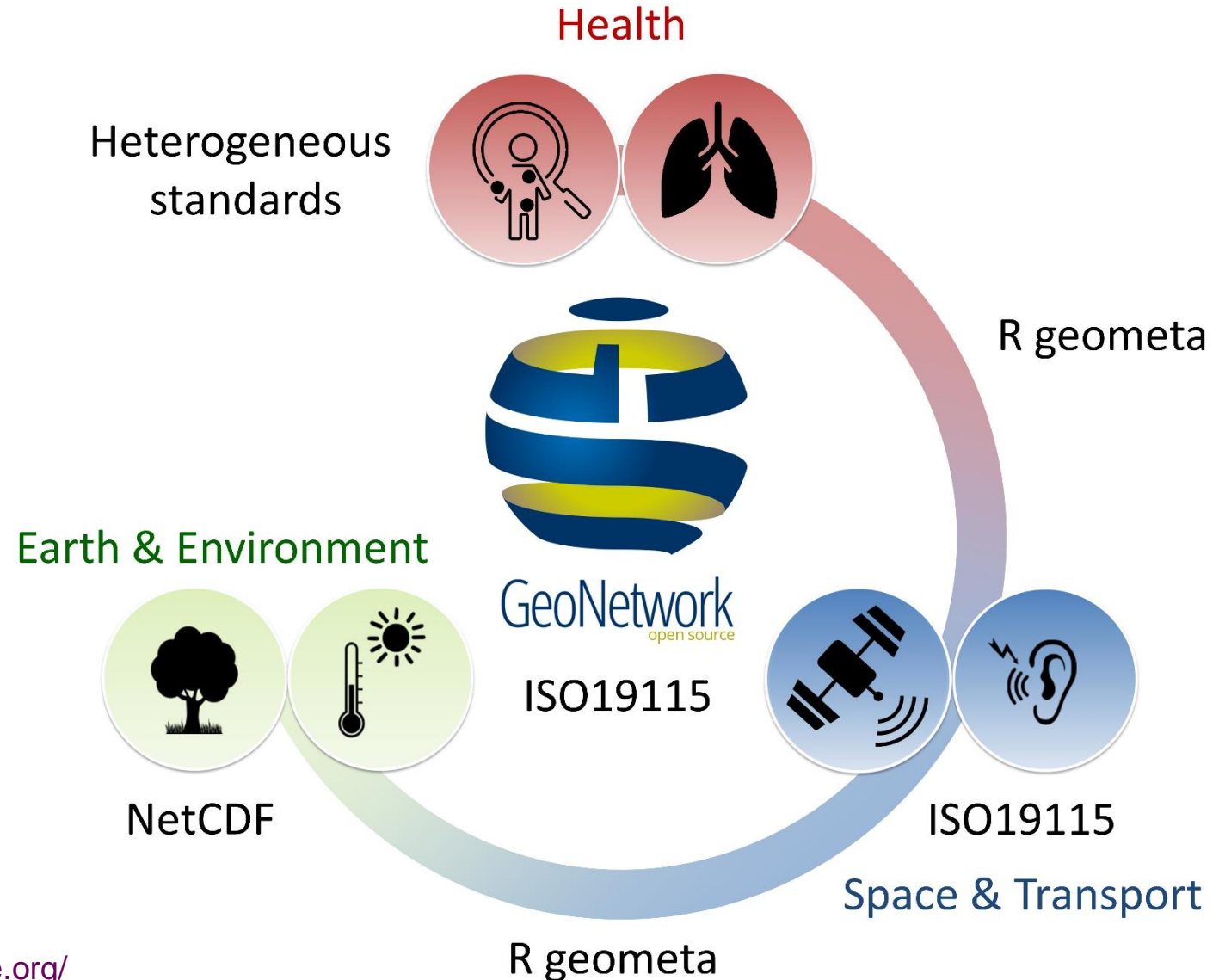
0 1 0 0 1 0 0 0 0 1 1 1 1  
1 0 1 1 0 1 0 0 1 0 1 0 1 1 0 0 0 1 1  
1 0 0 0 0 0 1 0 1 0 1 0 1 1 0 1 1 1  
1 0 1 0 1 1 1 0 0 0 1 1 0 1 0 1 1 1  
1 1 0 0 0 1 1 0 0 1 1 1 0 0 0 1 1 0 0 1





# Mapping strategy


- Main mapping criteria: **spatial and temporal coverage**
- Mapping platform: **GeoNetwork\***
  - Catalog application to manage spatially referenced resources
  - Provides powerful *metadata* editing and search functions
  - Interoperable via REST API
  - Interactive web map viewer
  - Used in Spatial Data Infrastructure initiatives worldwide

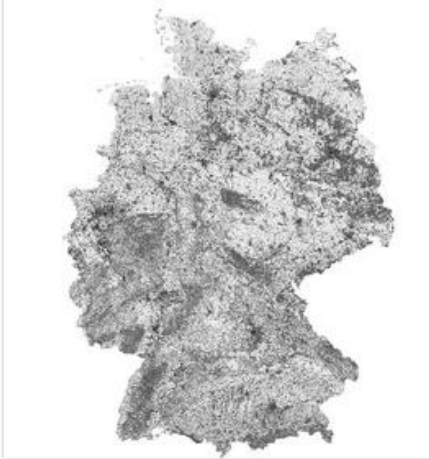



- Instance hosted on HMGU test server

→ **Cyberattack 15.03.2023!**



**Latest news** | Most popular | Comments

**Land Use Classification Map of Germany (WMS)**  


**Land Cover DE - Sentinel-2 - Germany,...**  


**GINIplus - COVID-19 Questionnaire**  


**Validation**

Show messages:  

Schema validation

0 Error

Datacite (DOI)

4 / 6

Geonetwork Rules

1 / 1

ISO Rules

16 / 16

URL Validation

3 / 3

- Records can be validated
- Catalog administrators can configure validation levels



Search **6** data sets, services and maps, ...

# Catalog starting page

Browse by

INSPIRE themes

Topic



Land cover



1



Land use



1

Type of resources



Dataset



3



Service



1



## Type of resources ▼

☒ Dataset (3)

☐ Service (1)

## Spatial representation type ▼

☐ Grid (1)

☐ Vector (1)

## Available in ▼

☐ Download service

☐ View service (1)

## Keywords ▼

☐ Germany (2)

☐ Opendata (2)

☐ Land cover (1)

☐ Land cover map (1)

☐ Sentinel-2 (1)

☐ Soil Moisture Index (1)

☐ Birth cohort (1)

☐ Covid-19 (1)

☐ Epidemiology (1)

☐ Questionnaire (1)

## Scales ▶


## Years ▶

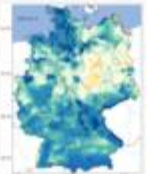
## Organizations ▼

☐ German Aerospace Center (DLR) (1)

☐ Helmholtz Centre for Environmental Research (1)


3 Results


☐

**SMI total soil monthly 1951-01 - 2021-12**



The UFZ Drought Monitor provides daily information on drought and soil moisture throughout Germany. This information is based on simulations using the mHM meso-scale hydrological model developed at the UFZ ([www.ufz.de/mhm](http://www.ufz.de/mhm)). Updated daily, the maps show you the drought status of the soil in its entirety and the topsoil, which reacts more quickly to


Helmholtz Centre for Environmental Research



☐

**Land Cover DE - Sentinel-2 - Germany, 2015**



This land cover classification of Germany was created using Sentinel-2 imagery from the years 2015 to 2017 and LUCAS 2015 in-situ reference data (<https://ec.europa.eu/eurostat/web/lucas>). It contains seven land cover types: (1) artificial land, (2) open soil, (3) high seasonal vegetation, (4) high perennial vegetation, (5) low seasonal vegetation, ...

German Aerospace Center (DLR)

☐

**GINIplus COVID-19 questionnaire**

This resource describes the spatial metadata of the Questionnaire of the COVID-19 pandemic of the GINIplus study. GINIplus is the acronym of the German Infant Study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development.

# UFZ use case: drought monitor

SMI total soil monthly 1951-01 - 2021-12

The UFZ Drought Monitor provides daily information on drought and soil moisture throughout Germany. This information is based on simulations using the mHM meso-scale hydrological model developed at the UFZ ([www.ufz.de/mhm](http://www.ufz.de/mhm)). Updated daily, the maps show you the drought status of the soil in its entirety and the topsoil, which reacts more quickly to recent precipitation, and the water available to plants in the soil.

On going

## Download and links



### Scientific publication

Scientific publication describing SMI <https://doi.org/10.1088/1748-9326/11/7/074002>

Open link

## About this resource

### Theme

- Soil Moisture Index

### Keywords

- Soil Moisture Index
- Germany
- Opendata

### Language

- English

### Contact for the resource

**Helmholtz Centre for Environmental Research**

- Author, Author, Point of Contact:  
✉ [klima@ufz.de](mailto:klima@ufz.de)

### Status

On going

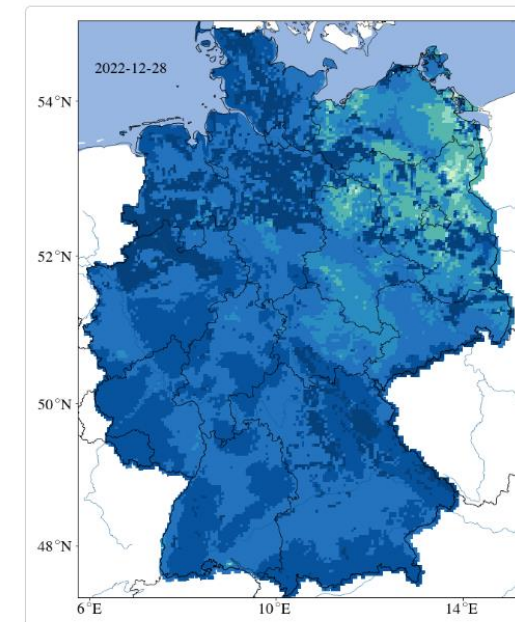
## Technical information

## Metadata information



Download metadata

Overview



Spatial extent



Resource events

# HMGU use case: GINIplus COVID-19 questionnaire

## GINIplus COVID-19 questionnaire

This resource describes the spatial metadata of the Questionnaire of the COVID-19 pandemic of the GINIplus study. GINIplus is the acronym of the German Infant Study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development.

On going

### About this resource

Categories	
<a href="#">Health</a> <a href="#">Environment</a>	
General	<ul style="list-style-type: none"> <li>Birth cohort <a href="#">Q</a></li> <li>Covid-19 <a href="#">Q</a></li> <li>Questionnaire <a href="#">Q</a></li> <li>Epidemiology <a href="#">Q</a></li> </ul>
Language	<ul style="list-style-type: none"> <li>English</li> </ul>
Resource identifier	<ul style="list-style-type: none"> <li>identifier</li> </ul>
Credit	GINIplus MetaMap3
Status	On going

### Technical information

Representation type	Vector
Coordinate reference system	<ul style="list-style-type: none"> <li>25832</li> </ul>

### Metadata information

[Download metadata](#)

Contact	<ul style="list-style-type: none"> <li><a href="#">Leibniz-Institut für umweltmedizinische Forschung</a></li> <li>Auf'm Hennekamp 50, Düsseldorf, 40225, Germany</li> <li>Point of Contact:               <ul style="list-style-type: none"> <li><a href="#">Elke Link (Data manager)</a></li> <li>+49 211 3389-286</li> </ul> </li> </ul>
---------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Spatial extent



### Resource events

Creation	01-01-2015 02:00
----------	------------------

[Provided by](#)

[Updated:](#)

a month ago

No ratings ★

[Share on social sites](#)

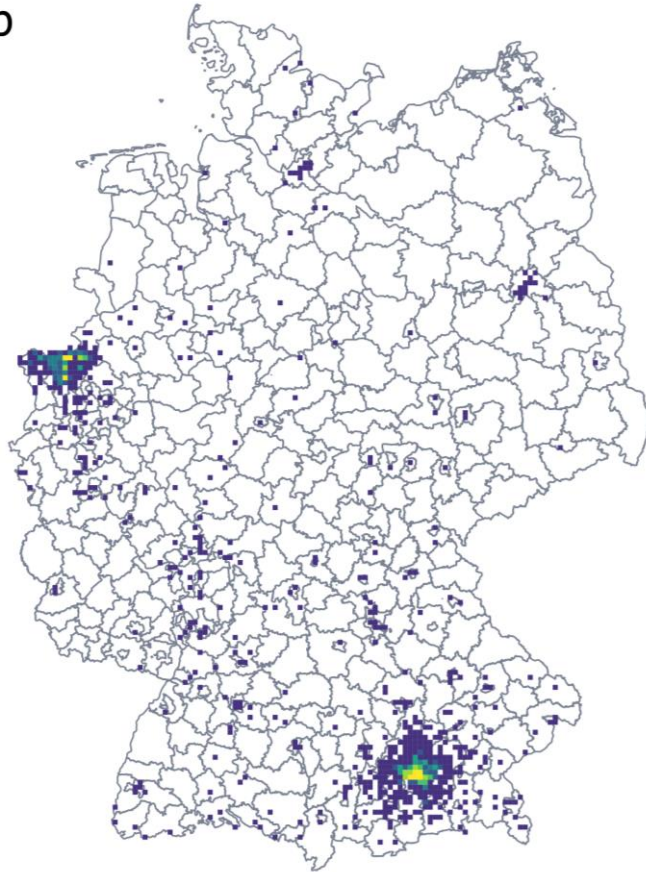




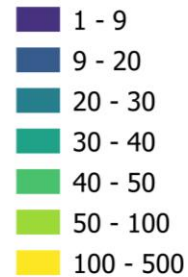
# GINI & LISA change of residences

→ Add to  GeoNetwork  
opensource

Follow-up  
20 years

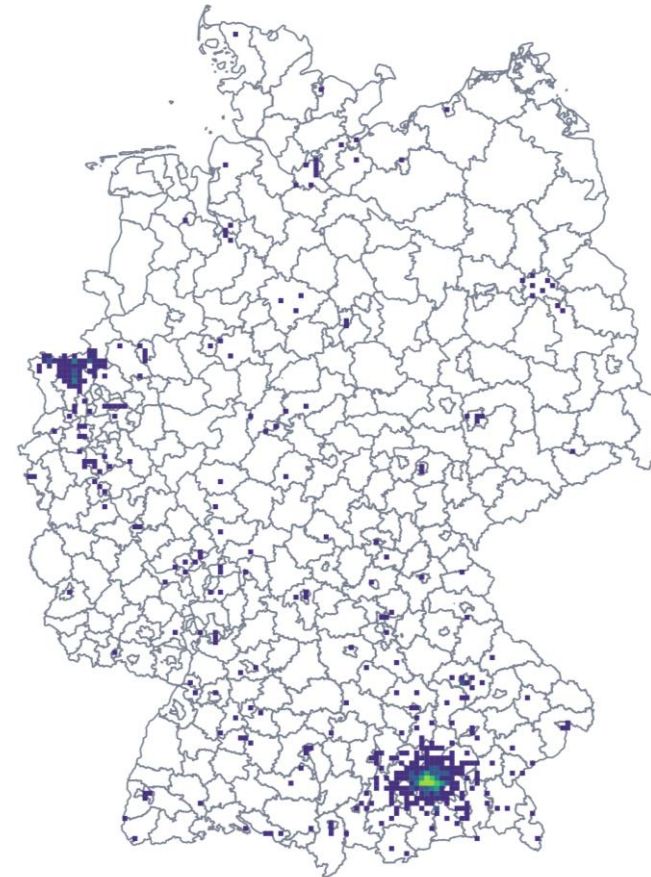


Number of GINI participants

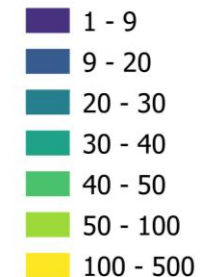


Districts

0 100 200 km



Number of LISA participants



Districts

0 100 200 km

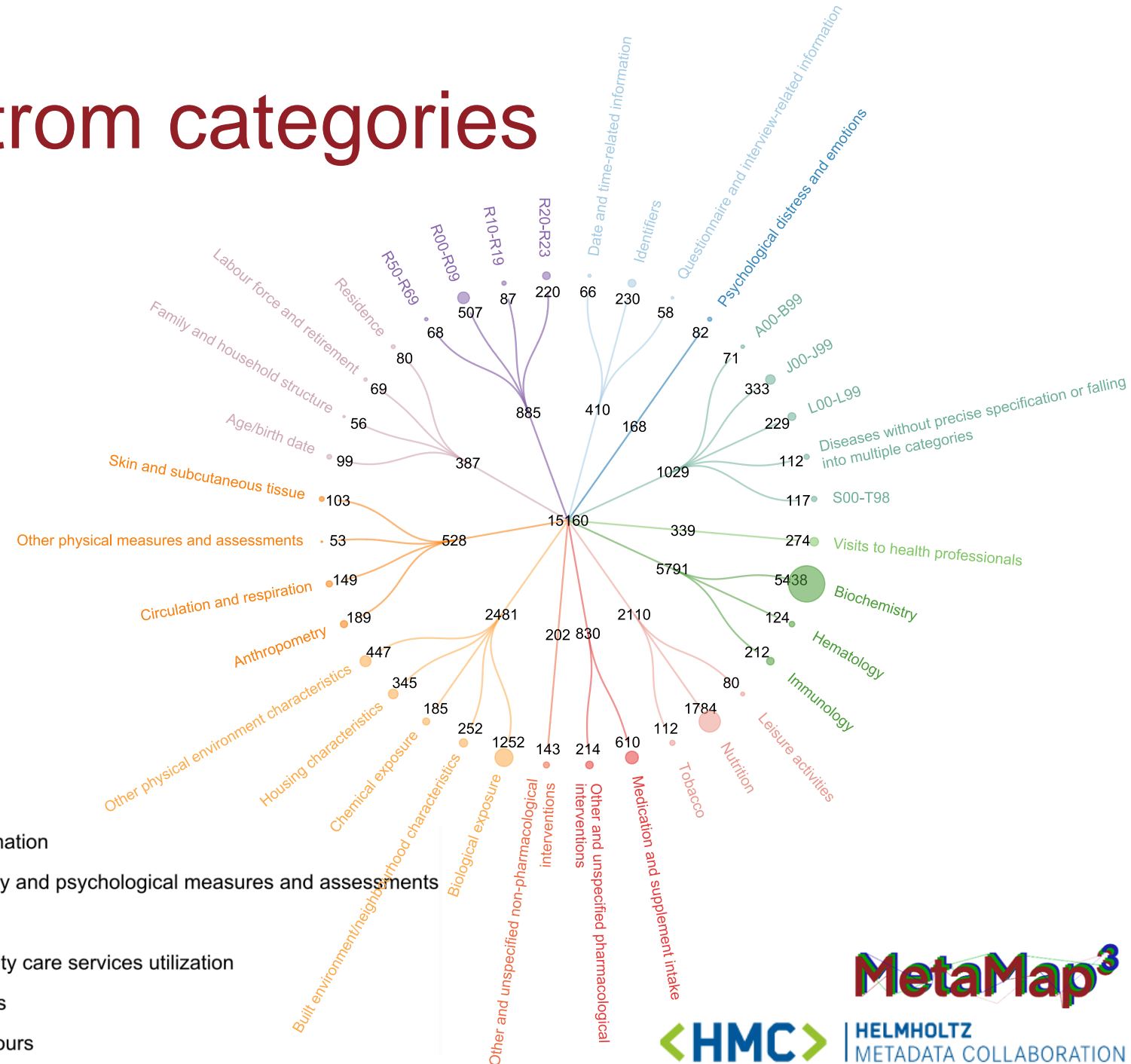


# GINI & LISA: Maelstrom categories

- ~15,000 variables classified based on the Maelstrom Research Taxonomies
- Dendrogram of the LISA cohort variables from baseline to 20-years follow-up



- |                                                  |                                                                     |
|--------------------------------------------------|---------------------------------------------------------------------|
| • Medication and supplements                     | • Administrative information                                        |
| • Non-pharmacological interventions              | • Cognition, personality and psychological measures and assessments |
| • Physical environment                           | • Diseases (ICD-10)                                                 |
| • Physical measures and assessments              | • Health and community care services utilization                    |
| • Socio-demographic and economic characteristics | • Laboratory measures                                               |
| • Symptoms and signs (ICD-10)                    | • Lifestyle and behaviours                                          |



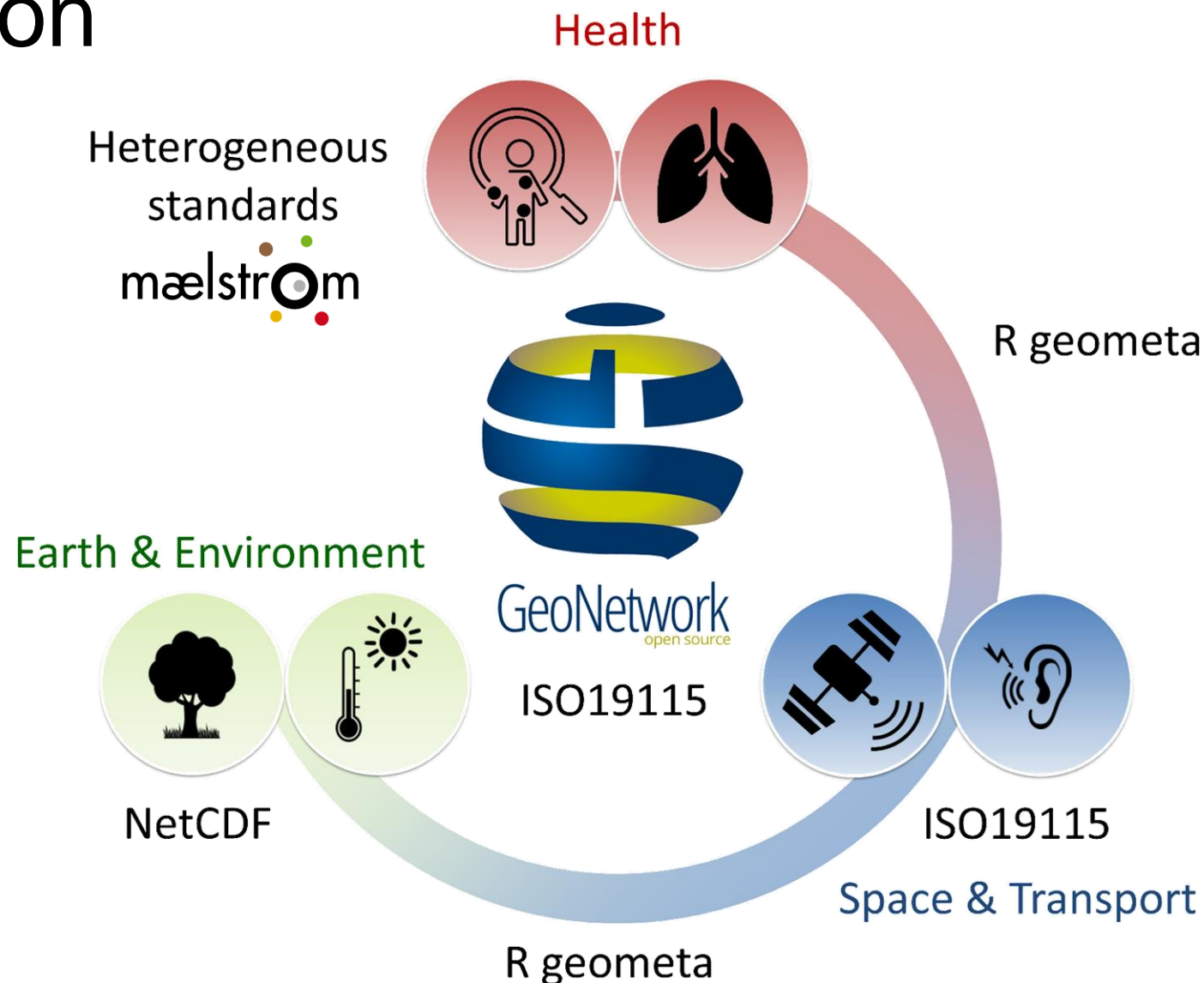
# Summary & Conclusion

- Prototype for mapping set up
- Joint standard: ISO19115
- Use cases converted and uploaded

→ **Jointly querying and searching now possible**

- Health metadata enriched
- Uploaded into NFDI4health

→ **FAIRness improved**





# Outlook

- Set up new VM with Geonetwork instance
- Add further datasets
- Utilize filter
- Make available to HMC

**We are open for additional (meta-)data!**



Thanks for your attention!

Questions???

MetaMap<sup>3</sup>