The EOSC Future Science Project ‘Climate Neutral and Smart Cities’: Integrating data from different clusters

Hilde Orten, Sikt - Norwegian Agency for Shared Services in Education and Research

EOSC Future webinar May 16 2023

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement 101017536

T6.3 Science Project 9, Climate Neutral and Smart Cities

Objectives:

- Demonstrate that relevant **environmental data** and **data on citizens' values, attitudes, behavior and involvement** can be combined for social, political and scientific analysis

- Data available via the EOSC Platform

Cross-cluster partners:

- SSHOC: ESS Eric (City University of London; Sikt), CESSDA (SND)
- ENVRI Community: IAGOS
- Project independent experts

Problem statement

- High level: Man in his/her environmental context

- More specifically for this project:

  Do observable indicators related to climate change or local air pollution, have effect on attitudes to climate change or affect on well-being indicators?

Science Project 9, Climate Neutral and Smart Cities

Three pilars

Indicator production and data integration

ESS Data
Climate data
Air quality data

Cross-Domain metadata

DDI Cross-Domain integration
use together with DDI-Lifecycle

Dissemination

ESS Context lab
EOSC Portal

Indicator production and data integration

Climatedata from ERA5

Name	Long Name	Type
adapto…	adaptor.mars.i…	Local File
fg10	10 metre wind …	Geo2D
lati…	latitude	—
lon…	longitude	1D
t2m	2 metre temper…	Geo2D
time	time	1D
tp	Total precipitation	Geo2D

Air quality data from EEA

Sample	AirPolli	AirPollutantCode	Aveı	Concentration	UnitO
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	103.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	92.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	102.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	100.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	86.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	64.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	55.0000000000	µg/m3
SAM.CZ_ABRAA_PM10_40270	PM10	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/5	hour	48.0000000000	µg/m3

Indicator production

Data integration

Data from the European Social Survey for a selection of European cities

D19 **CARD 37** You may have heard the idea that the world's[73] climate is changing due to[74] increases in temperature over the past 100 years. What is your personal opinion on this? Do you think the world's climate is changing? Choose your answer from this card.

Definitely changing	1	
Probably changing	2	**GO TO D21**
Probably not changing	3	
Definitely not changing	4	**ASK D20**
(Refusal)	7	**GO TO D21**
(Don't know)	8	

ASK IF DEFINITELY NOT CHANGING AT D19 (code 4)
D20 **CARD 38** How much have you thought about climate change before today?

Not at all	1	
Very little	2	
Some	3	
A lot	4	**GO TO D30**
A great deal	5	
(Refusal)	7	
(Don't know)	8	

Example indicator production

aqiwdpm10	Worst air quality index level PM10, date	This variable captures the worst air quality index level for PM10 (Particles less than 10 µm (PM10),background stations by date and region

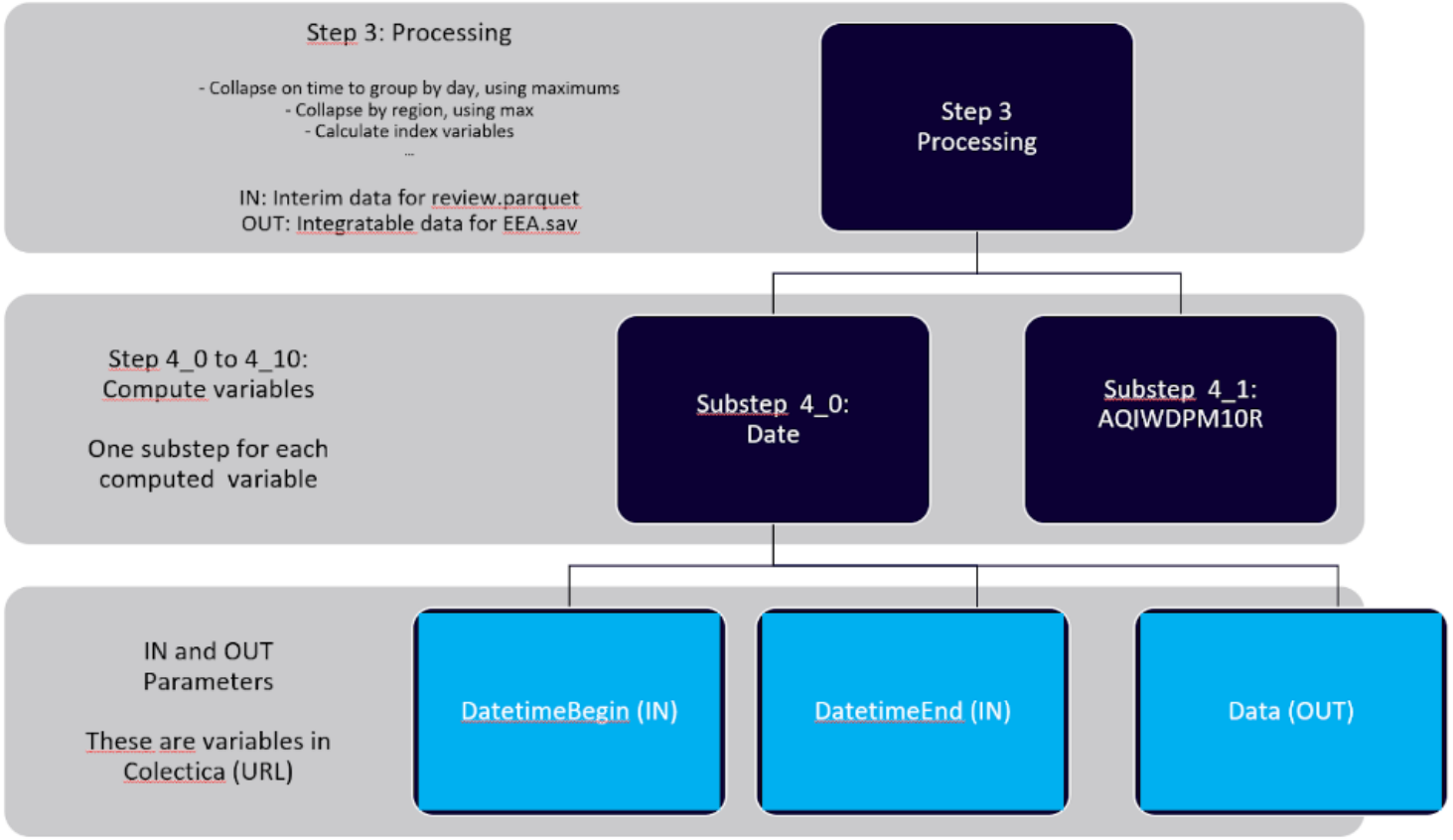
Computation rules

Compute target variable 'aqiwdpm10' starting from variables 'Concentration' and 'AirPollutant = PM10' for each background station. Compute where 'AirQualityStation' has values for the pollutant. Find 'Concentration' value at the 99th percentile for the pollutant. Create EEA Air Quality Index where PM10: 0 to 20 eq 'Good' represented by value '0'; 20 to 40 eq 'Fair' represented by value '1'; 40 to 50 eq 'Moderate' represented by value '2'; 50 to 100 eq 'Poor' represented by value '3'; 100 to 150 eq 'Very Poor' represented by value '4'; 150 to 1200 eq 'Extremely poor' represented by value '5'. The worst air quality level measured on a specific date on one of the background stations in the region provides values to the variable.

Cross-Domain metadata

Document process steps in a machine readable way to enhance transparency and FAIR practices using open standards

Document data integration steps in DDI-Cross Domain Integration (CDI)



Example in DDI-CDI provenance application prototype

SP9 ESS Labs » SP9 Process Description

European Social Survey

Quick search

Go

Table of Contents

- Integrate climate and air quality data with ESS

SP9 Process Description

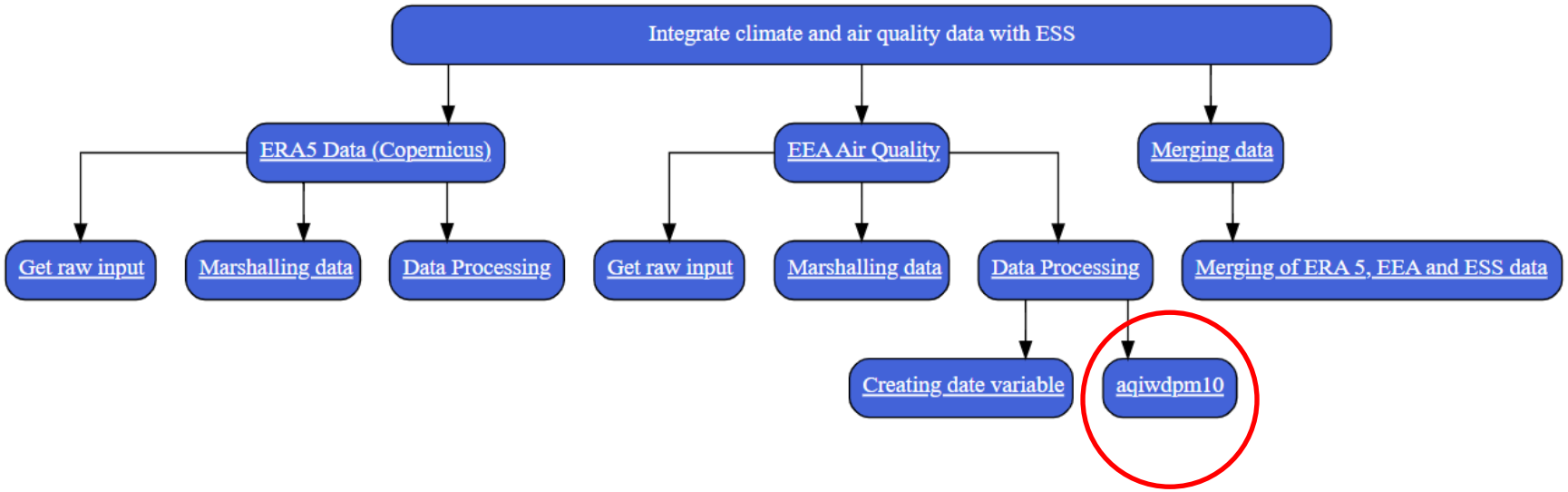
Prototype

Main Sequence of the process

ProcessingAgent: Processing Agent 1 Purpose: Purpose of the processing agent Production environment: Sikt - Norwegian Agency for Shared Services in Education and Research acting as a participant of SP9

- Integrate climate and air quality data with ESS
  - ERA5 Data (Copernicus)
    - Get raw input
    - Marshalling data
    - Data Processing
  - EEA Air Quality
    - Get raw input
    - Marshalling data
    - Data Processing
      - Creating date variable
      - aqiwdpm10
  - Merging data
    - Merging of ERA 5, EEA and ESS data

Note: Move the mouse cursor over a name to see more information. Click on a name to go to the corresponding page.



Example in DDI-CDI provenance application prototype, ctd.

European Social Survey

Quick search

Go

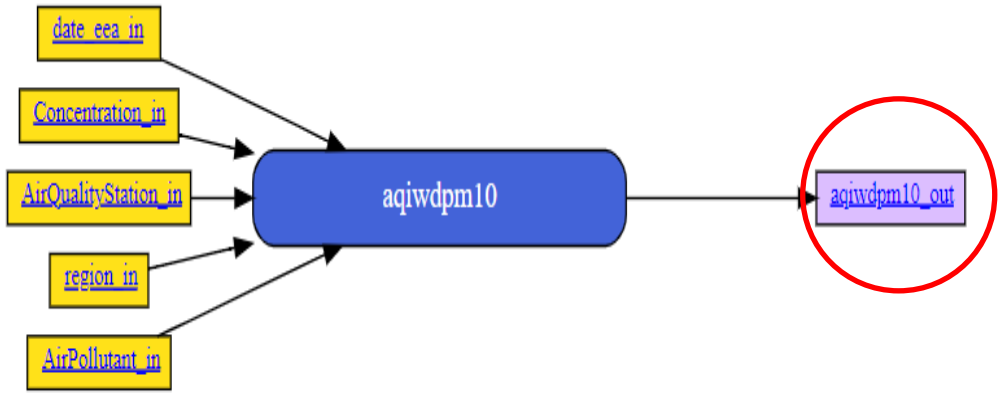
Table of Contents

- Integrate climate and air quality data with ESS
  - ERA5 Data (Copernicus)
  - EEA Air Quality
    - Get raw input
    - Marshalling data
    - Data Processing
      - Creating date variable
      - aqiwdpm10
  - Merging data

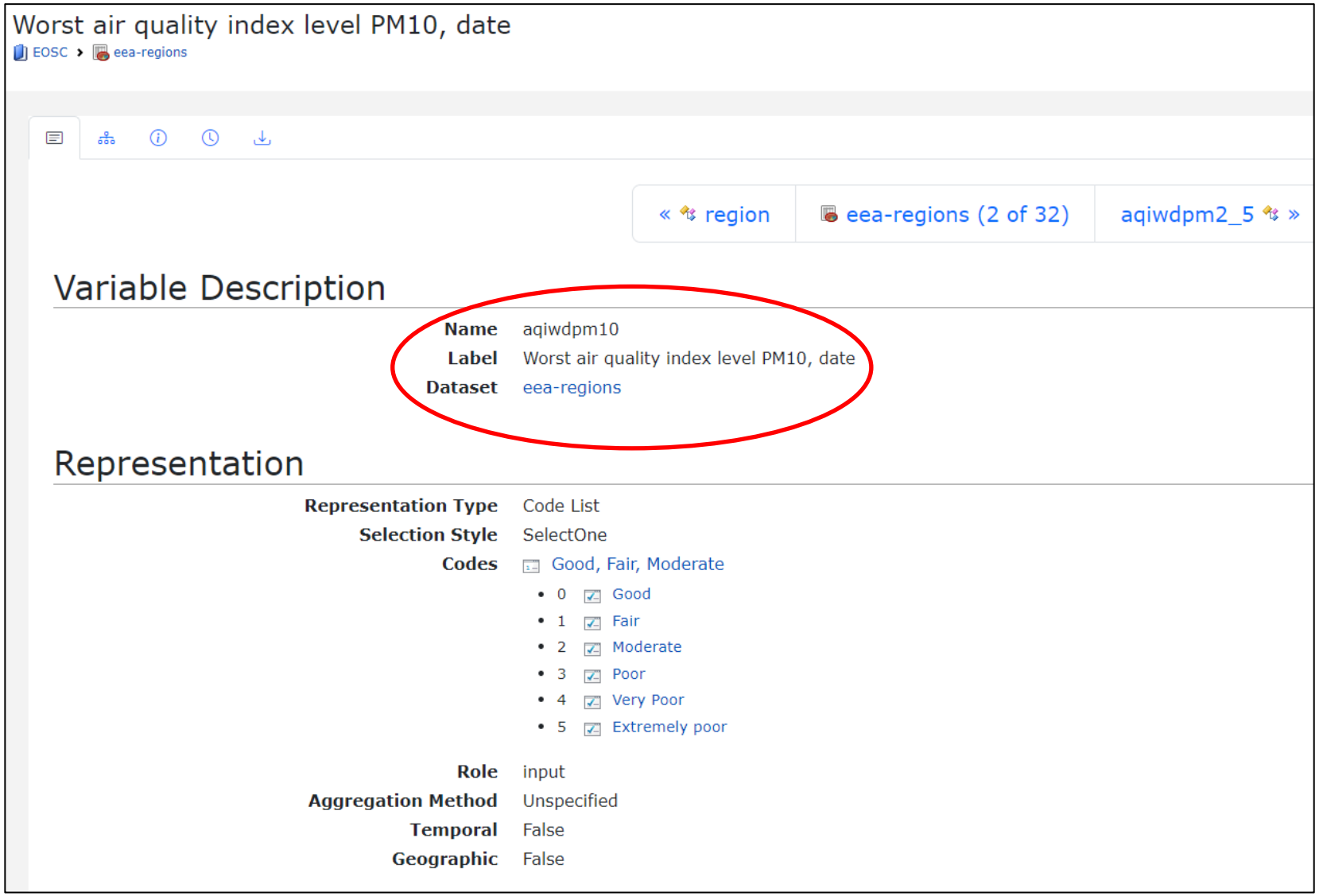
aqiwdpm10

Description: Compute target variable 'aqiwdpm10' starting from variables 'Concentration' and 'AirPollutant = PM10' for each background station. Compute where 'AirQualityStation' has values for the pollutant. Find max 'Concentration' value for the pollutant for the date. Create EEA Air Quality Index where PM10: 0 to 20 eq 'Good' represented by value '0'; 20 to 40 eq 'Fair' represented by value '1'; 40 to 50 eq 'Moderate' represented by value '2'; 50 to 100 eq 'Poor' represented by value '3'; 100 to 150 eq 'Very Poor' represented by value '4'; 150 to 1200 eq 'Extremely poor' represented by value '5'. Add missing values where data are missing. The worst air quality level measured on a specific date on one of the background stations in the region provides values to the variable.

This step uses a script written in Python 3.9.



Example in DDI-CDI provenance application prototype, ctd.



Metadata allowing reuse and reproducibility

Exploring possibilities DDI-Cross-Domain Integration (DDI-CDI)

- Provenance
- Description of data from different structures
- Catalogue details
- Variable cascade
- Use together with other standards – for example DDI-Lifecycle

SSHOC
social sciences & humanities open cloud

ENVRI
community

Dissemination

- ESS Labs – prototype application with this project as its first instance.
- Starting from ESS Data Portal technology
- Add DDI-CDI based provenance components
- EOSC Portal
  - Add ESS Labs as a service
  - European Social Survey ESS allready onboarded

SSHOC social sciences & humanities open cloud

ENVRI community

Thanks to

- Eric Harrison (ESS HQ City University of London)
- Eirik Stavestrand, Hanna Thome Grieg, Benjamin Beuster, Archana Bidargaddi (Sikt)
- Joachim Wackerow, Arofan Gregory (Consultants for Sikt)
- Iris Alfredsson, David Rayner, Ilse Laze (SND)
- Hannah Clark (IAGOS)
- Experst from NILU and the Norwegian Meteorological Institute

- Thanks to EOSC Future and the WP6 lead for making this possible

SSHOC
social sciences & humanities open cloud

ENVRI community

Thank you very much for your attention!

hilde.orten@sikt.no

Appendix

Identified environmental indicators and indices

Concept	Indicator	Indices (computed measures)
Climate change (observable properties)	Temperature	Temperature anomaly
	Precipitation	Extreme precipitation days
	Wind strength	Storm – extreme wind
Air quality (classic indicators)	Inhalable particular matter (PM10, PM2,5)	Air Quality • Day of interview • Week before the interview • Month etc.
	Nitrogen Dioxide (NO2)	
	Sulfur Dioxide (SO2)	
	Ozone (ground –level) (O3)	

SSHOC social sciences & humanities open cloud

ENVRI community

Data sources for Air Quality and Climate Data

Data source	Data source name	Reference (URL)	Comment
Air quality	European Environmental Agency (EEA)	https://www.eea.europa.eu/	Air quality indicator data, by hour
Climate indicators	Copernicus ERA5	https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5	Re-analyses climate data, by hour

SSHOC social sciences & humanities open cloud

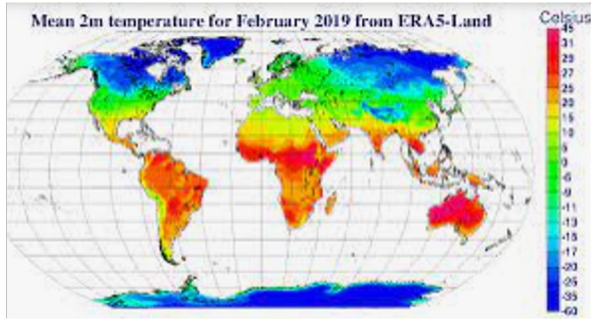
ENVRI community

Geographic regions to be covered*

Urban region NUTS level + ESS var Domicile = 'big city', 'suburb' or 'outskirt of big city'	N in ESS Round 8 2016
Stockholm (SE11 Stockholms län)	246
Berlin (DE3 Berlin)	125
Praha (CZ010 Hlavní mesto Praha)	277
Budapest (HU110 Budapest)	291
Wien (AT13 Wien)	398
Madrid (ES30 Comunidad de Madrid)	167
Bruxelles/ Brussels (BE10 Région de Bruxelles-Capitale /Brussels Hoofdstedelijk Gewest)	182
London (UKI London)	143
Oslo (NO01 Oslo og Akershus)?	251

* Explore inclusion of data from Paris and Munich from ICOS Cities?

Methods for data integration



- Reduce amount of data
  - Climate data cover hourly measures for the whole world since many decades past
- Geomapping
  - NUTS polygons, grids and measurement station geopoints need to be combined
  - Mapping to common projection
  - Use Geostat population data to weight measures within areas
- Time
  - Relate all indicators to the timing of the interview
  - Compute variable ‘Date’ to relate events and indices to each other
  - Include data from back in time for normalisation purposes
- Indicator production and measures
  - Input from climate and air quality experts to learn about best practices.

SSHOC social sciences & humanities open cloud

ENVRI community

Expected data users

1) Social and behavioral scientists (immediate term)

2) Policy makers (medium or long term)

3) Researchers from other domains (Environmental scientists)(medium or long term)