

Grant Proposal

Author-formatted document posted on 14/06/2023

Published in a RIO article collection by decision of the collection editors.

DOI: <https://doi.org/10.3897/arphapreprints.e107872>

Climate Neutral and Smart Cities, combining data about city people and their physical environment

Bodil Agasøster, Eric Harrison, Hilde Orten

Climate Neutral and Smart Cities, combining data about city people and their physical environment

Authors

Bodil Agasøster [Orcid ID](#) , Eric Harrison [Orcid ID](#), Hilde Orten

Abstract

The main objective of the Science Project (SP) *Climate Neutral and Smart Cities* is to demonstrate that environmental data and data about people's attitudes, behavior and involvement can be combined for social, political and scientific analysis. In the project, scientists from the Social Sciences & Humanities Open Cloud (SSHOC) and Environmental Research Infrastructures (ENVRI) community work together with the aim of producing new and useful outputs for the benefit of the research community, such as indicators related to environmental indices, and methods and workflows for computing them. The environmental variables will be integrated with data from the European Social Survey for a selection of big European city regions. Data and metadata from the project will be accessible through a new prototype application that will be made available as an exploratory "labs" service from the European Open Science Cloud (EOSC) Portal and the EOSC Marketplace. This will allow easy access to cross-domain data for scientific analysis and their provenance, as well as to other deliverables from the project.

Key words

EOSC Future; Science Projects; Science Clusters; cross-domain metadata; European Social Survey; DDI-CDI

Description

Existing situation

Scientific challenges have been identified in the field of climate neutral and smart cities between the ENVRI and SSHOC project stakeholders. There are silos of thematic data holdings in the social sciences and environmental sciences fields, and awareness and interest between the communities for data holdings, resources, and cross-thematic research is not sufficiently supported.

There is insufficient semantic interoperability between domain resources. A lesson from the SSHOC project is that there is not yet sufficient European Open Science Cloud (EOSC) engagement and uptake of EOSC services by the SSH researchers and infrastructures.

Objectives

The main objective of the SP is to contribute to the research focus and goals set by the Horizon Europe program's *Climate Neutral and Smart Cities* mission by combining resources from two

clusters encompassing Citizens and Sustainable Environment. The project will demonstrate that relevant environmental data and data on citizens' values, attitudes, behavior and involvement can be combined for social, political and scientific analysis, and describe how this has been achieved.

As part of the SSHOC consortium, the European Social Survey (ESS) collects data directly related to the mission on themes such as citizens' involvement and democracy, political values and engagement, health and social care, smart economy and smart mobility. The ESS covers political and social trust, health and health inequality, attitudes towards climate change and energy, understandings and evaluations of democracy and digital communication at work and with family, amongst many other topics related to the smart agenda.

To make this data more useful to researchers, contextual data can be provided to allow for a better understanding of the core ESS data. In the past such data has sometimes been incorporated into dissemination platforms for the ESS data. This kind of contextual data has never been easy to maintain, and the way in which it has been collected has not been obvious, as it is done by other organisations for their own purposes. Further, although presented alongside the ESS data, it has never been processed into a truly integrated form, allowing a more significant use of the contextual data. Its purpose has been limited to providing a better understanding of the ESS data.

A SSHOC and ENVRI case study will be undertaken to demonstrate the interdisciplinary opportunities provided by a more-complete integration of contextual data. This will expose a valuable resource in the EOSC Portal, promoting interoperability and connectivity between clusters within EOSC by giving researchers a powerful way to use data coming from disparate organizations in different domains.

After the plan for the integration of the SP, including identification of the most relevant data and alignment of structured metadata standards, the project has harvested environmental data from sources agreed between the ESS and ENVRI partners. The environmental data is structured and processed and added to the ESS data to cater for various types of analysis with mainly social science data. The prototype application under development needs to be able to host and describe the new data and the processes performed to integrate data from the various sources.

The scope of the data holdings of the service will include environmental data such as air quality and climate indicators for a selection of big European city regions, as well as survey data from the ESS.

This SP will, for example, look at data which provide insight into how environmental factors and citizens' behavior interact and explore how environmental aspects such as policy may change based on information from survey data. The SP integration will follow, based on tests and analysis of interaction between environmental indicators and those of social behavior.

Links and interactions between the SSHOC and ENVRI Clusters are thus described through data to be employed and services for their analysis. In addition, an important link will be the communities from both sides to working successfully together. Real interaction between the science clusters is demonstrated. Interfaces between the two science clusters and their "trading zones" have been identified and explorative analysis have been undertaken to identify useful indicators. Based on this analysis, relevant working hypotheses are being formulated and tested. The exploration of exactly how such cross-cluster collaborations can be carried out is seen as an important aspect of the work: while there is broad agreement that such collaboration is desirable, the specifics of how best to do it are less clear.

Common investment in these zones is realized by means of essential ingredients such as open access (FAIR) data, reproducible analysis, mobilized communities etc. This common investment has great potential to generate novel research and innovation. The science projects offer a model of this type of interaction: common understanding of hypotheses, which require FAIR data and metadata, analytical services and communities to work on.

For environmental data to be relevant for researchers and policy/decision makers, relevant data will be transformed into indicators. Main challenges related to this are:

1. To define and specify what is good data quality (document the provenance of the data, make data interoperable and reusable, develop integrated indicators, and operationalize them into variables for analysis);
2. When needed and possible, various metadata standards and controlled vocabularies will be aligned so that environmental data can be mapped, using metadata standards such as the DDI Cross Domain Integration (DDI-CDI) and others. ESS uses DDI Lifecycle (DDI-L), and other standards such as NetCDF are used by some of the contextual data sources). This approach has already been demonstrated in *The Role of DDI-CDI in EOSC: Possible Uses and Applications*, a report for the European Open Science Cloud Co-Creation Project, and
3. To explore options for harvesting, transformation, merging and processing of contextual environmental data.

ENVRI brings in environmental data expertise, as well as their experiences from the ENVRI Dashboard.

While relevant data sources will first and foremost be existing data, the project will be keeping an eye on future data flows from ongoing and future projects that may be added to the application later.

The SP efforts and technical implementation resources will be made available from the ESS website through the distributed EOSC Portal, and input to outreach and learning will be produced. This will be done through an *ESS Labs* service which allows interested parties to explore this project in a hands-on fashion, which allows them to understand and leverage the approaches as appropriate to their own organisations and clusters.

Compliance to criteria

Eligibility

The partners are two SSH European Research Infrastructure Consortia, CESSDA ERIC and ESS ERIC, including research domain archives as Linked Third Parties, and the ENVRI FAIR cluster.

Contribution to EOSC

A contribution to EOSC is to illustrate how cross-domain data integration can be fully described, using the DDI-CDI metadata standard, which has already demonstrated its potential uses and applications for aligning various data types in a report for the European Open Science Cloud Co-Creation Project.

DDI-CDI focuses on a uniform approach to describing the data lineage and the process of data integration at a detailed level, so that a researcher can understand exactly which data

and processes were used in the creation of integrated variables available with the core ESS data. The DDI-CDI can tie together a range of data in different formats allowing the various types of data to be connected and understood to support transformation and processing for integrated use. DDI-CDI can be aligned with other DDI specifications (DDI-Codebook, DDI-Lifecycle) to support integration of external data in systems which use DDI, or can reference data and metadata structured in other standard formats such as NetCDF. DDI-CDI is explicitly designed to work with many popular generic technology standards to allow for easy integration into systems which support them. It offers an extension to the suite of DDI work products which will help those in the Social, Behavioral and Economic domains and outside of them integrate the expanding range of data required by today's research.

Quality

The DDI-CDI standard facilitates FAIR compliance and facilitates machine actionability of data from different domains at a detailed level. It enables tracking of the provenance and record data transformations. The science project is exploring implementation of these features. This approach is in line with other initiatives in the FAIR space, such as the WorldFAIR project being conducted by CODATA, RDA, and other organisations looking at cross-domain FAIR implementation. DDI-CDI is among the domain-independent standards the project recommends.

Relevance

This project falls within the core of the EU focus areas, in that it involves climate change and air quality indicator variables to be integrated with survey data about attitudes to climate change and other data about citizen's attitudes, values and behaviors, and make the combined data available to researchers for analysis.

Implementation, Plan of work:

Tasks

The tasks identified for the project are described below.

1. Relevance: Planning the SP data and metadata integration based on identification of relevant data sources for environmental data (expected to be sensor and reanalysis data) and expected impact (all partners).
2. Prepare implementation: Explore technical frameworks and solutions.
3. Interoperability: Explore possibilities for alignment of metadata standards by the EOSC Interoperability Framework and the DDI-CDI model (ESS ERIC/Sikt) and Consortium of European Social Science Data Archives (CESSDA) ERIC/Svensk Nationell Datatjänst (SND) with ENVRI).
4. Reusability: Transforming and structuring environmental data and metadata into re-usable data in a prototype application (ESS ERIC/Sikt), CESSDA/SND, ENVRI)
5. Prototype development: Develop a prototype application and populate it with new data, expressed in the form of integrated variables suitable for use by researchers across domains (ESS ERIC/Sikt).
6. Analysis: Test and analyse data from the ESS together with environmental indicators developed in the project (ESS ERIC)
7. Outreach: Input from the use case to EOSC Future outreach and learning (ESS ERIC/Sikt)
8. Deliverables: Preparation of documentation and reports related to the EOSC Futures Work Package 6 deliverables (all partners)

Milestones based on tasks - or rather, on EOS future project level milestones are outlined below:

1. Plans for SP integration and implementation (M16)
2. Data from the project prepared (M24)
3. Metadata standards aligned (M26)
4. ESS data dissemination tool prototype upgrade including tools, documentation, and Services Registry populated with the new components (M28)
5. Input to outreach and learning (M28)
6. Science project integrated based on tests and data analysis (M29)
7. Documentation and reports delivered to the Work Package lead (M29)
8. Making the prototype application ESS Labs available as a service through EOSC portal and catalogs (M30).

The SP will contribute to the achievement of the WP6 deliverables: D6.3 Deliverables on the Results of the Scientific Projects as described in the DoA (M30).

In addition, the SP is expected to contribute to risk reduction in the following areas:

- Low uptake by external user communities: Interactions by SP9 project members with RDA and CODATA - notably with the WordFair Project - are highlighting the importance of the outcomes of this Science Project.
- Limited adoption of interoperability guidelines: Members of SP9 are active in the EOSC Semantic Interoperability Task Force, GO FAIR, the FAIR Digital Object Forum, and the Cross-Domain Interoperability Framework effort. These groups are focused on standards and best practices for data sharing and will provide the frameworks for these into the future. Notably, the Reuse and Reusability paper will provide recommendations for metadata standards and their application in the context of multi-disciplinary science.
- Lack of engagement by stakeholders: This Science Project has reached out to non-project partner stakeholders in the environmental community such as the Climate and Environmental Research Institute NILU and the Meteorologisk institutt (MET Norway) to help advise on data related issues. Engagement of these stakeholders enhances the visibility of scientific outcomes and awareness of EOSC generally.

Use of resources

The project is exploring usages of Core services, the Helpdesk in particular. In a project which is exploratory in nature, it is important to provide a consistent mechanism for handling the questions which will arise - the ESS Labs service is likely to produce many such questions, and the full benefit will not be realised if these are not followed up consistently. The Helpdesk service is thus seen as an important way of ensuring the desired impact is realised.

What the SP brings to EOSC Future platform

The ESS Labs service will provide a prototype of FAIR compliant data from repositories with environmental data (for example the European Environmental Agency (EEA) and Copernicus ERA5), as well as data from the European Social Survey on citizens' values,

attitudes, behavior and involvement in a prototype application made available as a service available from the EOSC Portal.

Target groups include researchers, academics, students, first from the social and behavioral sciences, and later environmental managers, policy makers stakeholders. Application domains: cross-domain scientific collaboration, research project development and decision makers.

Impact

Strategic

Contributes to the establishment/further development of EOSC by upgrading the ESS platform with sustainable tools approaches for multi-disciplinary data integration, and by linking data and research communities/users. Demonstrates the feasibility and relevance of cross-domain cooperation and analysis.

Scientific /User communities

The SP is an interdisciplinary project, but in the short-term social scientists are likely to be the primary user group. Impact will then be on the level of an infrastructure mainly targeting social sciences, by demonstrating that environmental data and data from other disciplines can be integrated with social science individual and contextual data, thus offering richer data and permitting research into novel research questions. This will establish an opportunity/platform for cross-domain research and innovation.

The science project will bridge and mobilize two research communities and contribute to the integration of research infrastructure data and services or the "composability" between the environment and social sciences domains and demonstrate how EOSC can be incorporated into the routine provisioning of research infrastructure resources.

From around 5-6 years after the infrastructure is established, medium to long-term effects will show scientific value. Two-way effects, from environmental indicators to citizens, and from attitudinal data to policymaking and policy implementation, can be analyzed and the research output exploited.

Societal / Economic

New, FAIR-compliant data can contribute to better-informed policies – based on citizens' values, attitudes and behavior – can ensure more environment-friendly behavior and higher societal uptake of innovative policies in these areas.

The linking of cross-domain data (about for instance environmental quality) with existing attitudinal data should increase awareness of how behaviors are both shaped by, and contribute to, that context. The ESS Topline publication on *Public Attitudes to Climate Change, ESS8, 2016* (23 countries) focuses on data on attitudes to climate change, energy security and energy preferences. It refers to Barasi's (2017) study¹ which shows that while Europeans recognize climate change as a problem, they do not yet seem motivated to substantial behavioral change.

¹ Barasi, L. (2017): The climate majority. London: The New Internationalist.

From the data produced by the SP, in the longer term, decision makers can see the impact of environment and changes of living in that environment on citizens' perceptions and well-being. This can result in better-informed and more relevant policies and implementation.

EU Policies

Climate neutral and smart cities is one of the European research and innovation missions and an integral part of the Horizon Europe Framework Programme. The SP aims to contribute to the research focus and goals set by the mission by combining resources from two clusters encompassing Citizens and Sustainable Environment.

Engagement Plan

Target groups

Researchers and students from the Social Sciences (short term goal), Policy Makers, Environmental Scientists (medium to long term goal). A secondary target are technology implementers concerned with best practice in enabling research across disciplines in line with the FAIR principles.

SP key concept

The SP contributes to the hot topic of the impacts of environmental factors on citizen's behaviors and society.

The SP will be applied by: 1) Integrating data from different sources for analysis; 2) Demonstrating that survey data from the ESS and environmental sciences can be analyzed together in a meaningful way; 3) Developing a prototype application for making data available to users; 4) Making services available from the EOSC Portal; and 5) Engaging the relevant scientific communities.

Dissemination measures

- Scientific publications
 - Scientific publication of the project template of the SP (this document). In addition, research papers covering the following topics will be published:
 - a paper describing the data and methods used in the project;
 - a workflow and processes paper;
 - a reuse and reusability paper describing metadata use and possibilities metadata;
 - as well as a scientific paper describing the preliminary analysis results.

- Conferences

European DDI User Conference ([EDDI 2022](#)), EOSC Symposium 2022, European Geoscience Union General Assembly ([EGU2023](#)), Workshop in the margins of RDA, Gothenburg 2023: [DDI-CDI: Optimising Your Data Description for Integration and Reuse](#).

- Webinars

Presentation at joint [EOSC Future/INFRAEOSC webinar](#) focusing on EOSC use cases, May 2023. Launch webinar autumn 2023.

- Demonstrations to other communities

The SP could be demonstrated to other projects like [ICOS Cities](#), [WorldFAIR](#), [FAIR IMPACT](#), as well as the [EOSC Semantic Interoperability Task Force](#), the RDA community and others.

- Education and training events

Collaboration with partners regarding tutorials tailored to different audiences.

- Networking

[RDA 20th Plenary Meeting](#) and other.

Acknowledgements

Many thanks to Archana Bidargaddi, Arofan Gregory (Consultant), Benjamin Beuster, Hanna Thome Grieg, Åse Jorun Holthe-Tveit, Knut Kalgraff Skjåk, Eirik Stavestrand, Joachim Wackerow, Consultant (Sikt - ESS ERIC), Hannah Clark (In-service Aircraft for a Global Observing System (IAGOS) - ENVRI), Iris Alfredsson, Ilse Laze, David Rayner (SND - CESSDA ERIC), Irena Vipavc Brvar and Maja Dolinar (Slovenian Social Science Data Archive (ADP) - CESSDA ERIC) for contributing their time and competence to the project.

Many thanks also to Britt Ann Kårstad Høiskar, Miha Markelj, Sverre Solberg, Kjetil Tørseth - [NILU](#) and Øystein Godøy, Lara Ferighi, Hans Olav Hygen – [Norwegian Centre for Climate Services](#) (NCCS) for providing expert advice about environmental data.

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

References

Barasi, L. (2017): The climate majority. London: The New Internationalist.