
FACE-IT

A Data Management Plan created using dmproadmap

Creators: Robert Schlegel, Jean-Pierre Gattuso, First Name Kai Bischof, Simon Jungblut

Affiliation: Institut de la Mer de Villefranche

Funder: European Commission (Horizon 2020)

Template: Horizon 2020 DMP

ORCID iD: 0000-0002-4497-1920

Grant number: 869154

Project abstract:

The overarching objective of FACE-IT is to enable adaptive co-management of social-ecological fjord systems in the Arctic in the face of rapid cryosphere and biodiversity changes. The project will identify ways to manage the impacts of climate change on the cryosphere and marine biodiversity, and the interaction with other drivers of change. FACE-IT will contribute to IPCC assessments as well as key Sustainable Development Goals. The concept of FACE-IT rests on a comparison of selected Arctic fjord systems at different stage of cryosphere loss in Greenland, Svalbard and Finnmark, Northern Norway. The underlying two-pronged hypothesis is that the biodiversity of Arctic coastal zones is changing in accordance with the rates of cryosphere changes, and that these changes affect local communities, food production, livelihoods and other ecosystem services. FACE-IT approaches European Arctic fjords as local social-ecological systems. It gathers a strong interdisciplinary team of internationally recognised experts from both natural and social sciences. FACE-IT is organized in eight interdisciplinary work packages focusing on the drivers of change (WP1), their effects on biodiversity (WP2), ecosystem functioning (WP3), food provision and indigenous livelihoods

(WP4), nature-based tourism (WP5), the co- production of knowledge to identify governance strategies for adaptive co-management (WP6), and public outreach and policy input (WP7), project management (WP8), and ethics requirements (WP9). It includes the participation of Arctic stakeholders to ensure that Indigenous and local knowledges, perceptions and concerns about ongoing changes are taken into account in defining innovative and adaptive co-management approaches towards a more sustainable future. In this way FACE-IT will deliver significant contributions towards the implementation of the new integrated EU policy for the Arctic.

Last modified: 10-03-2021

FACE-IT - Initial DMP

1. Data summary

Provide a summary of the data addressing the following issues:

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

The overarching objective of FACE-IT is to enable adaptive co-management of social-ecological fjord systems in the Arctic in the face of rapid cryosphere and biodiversity changes. The purpose of the data collection for this project is to help facilitate these intended research goals. The data generated by the researchers in this project will likewise be used for said goals.

As there are no work packages (WPs) in this project with the express aim of investigating the physical changes in the Arctic, even though this underpins all of the work packages, this project relies heavily on the retrieval and re-use of existing physical data. To that end a report on the availability of these physical data (drivers of change) has been set as a 6 month deliverable in the project (D1.1 "Report on key drivers of changes in Arctic biodiversity") and a meta-database of the data to be used by and generated for this project is set as a 12 month deliverable (D1.2 "Meta-database"). Note that these deliverables are not limited to focussing on physical drivers alone. The types of drivers that have been identified as important are: cryosphere, physical, carbonate chemistry, biological, and societal. Ecological drivers are classified here as biological. These data will be collected from multiple established data sources including but not limited to: [NPDC](#), [SIOS](#), [NOAA](#), [PANGAEA](#), [EMODnet](#), [Copernicus](#), [SAON](#), [NSF](#), and [INTAROS](#).

The data collected/generated for this project fall into three broad categories: 1) Gridded data produced through modelling of the European Arctic and site specific sub-regions; 2) station data collected *in situ* at specific sites; and 3) social science data collected from interviews with stake holders. These data will be generated in a range of formats but will be aggregated into either CSV or NetCDF file formats if not already generated as such. It remains to be seen how this will be accomplished for the social science data.

The size of the data required to initiate the research outlined for this project is not yet known. A [meta-database](#) of the needed data has been established and it is estimated that less than 300 GB of data will need to be downloaded. Many research groups within the WPs have already downloaded much of the data they need. The modelling data that will be generated for this project is estimated to exceed multiple TB. These data will be stored on the server on which they are created and will be accessible to FACE-IT members via OPeNDAP file sharing. The *in situ* station and social science data generated for this project are not expected to exceed 100 GB combined and will be hosted on secure file

sharing data servers housed exclusively within the EU ([Nextcloud](#) and [pCloud](#)) for rapid data sharing within the FACE-IT project. All final data created for FACE-IT will be stored in perpetuity on PANGAEA (see below).

The data generated for this project will be useful to a wide range of future research in the Arctic, particularly that focusing on fjord systems. This is due in part to the multidisciplinary nature of the project. It is also because many of the *in situ* station data generated will be continuations of existing time series, whose continued maintenance is imperative to the understanding of changes in Arctic fjord systems. The model data generated for this project will provide a high resolution spatial understanding for where and when the European Arctic will be changing in a number of ways that will inform future research outside of this project. The interviews taken for this project will also provide historic records to future social science research in a changed future Arctic.

2. FAIR data

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

All data generated for FACE-IT will be stored on the [PANGAEA](#) data portal for long-term maintenance. This online repository is certified by multiple sources as a trusted data repository and contains a consistent internal metadata standard, ensuring the findability of the data stored there. It is one of the industry standards for hosting natural science data and there are multiple methods of searching through the database available for users within the HTML based user interface of the website and via the users local terminal using an [R package](#). In addition to findability through consistent metadata and search keywords, every dataset uploaded to PANGAEA is also given a DOI (digital object identifier).

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**

- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

All data generated by FACE-IT partners will be made internally available within three months of the finalisation of the data by the creating party. For model and other automatically generated data this means within three months of the creation of the data. For data that will require lab work (e.g. phytoplankton counts), this will mean three months from when these data are finally processed and could be used by the creating party. Much of the *in situ* sampled station data will be able to be made internally available within three months of collection. The biological station data may require multiple months, perhaps as much as a year for genetic samples, before they are ready to be made internally available. The social science data generated from interviews etc. will contain sensitive information regarding individual participants that will not be made internally or externally available. Some version of the interview results with anonymised participant information will be made publicly available before the conclusion of the project. For data that are not to be made publicly available, the reduced or anonymised data hosted publicly on PANGAEA will contain the necessary contact information to gain access to the restricted data if necessary.

With the exception of the sensitive social science data, all data generated by FACE-IT partners will be made publicly available and be uploaded to PANGAEA, either as CSV or NetCDF files, within two years of when they become internally available. Through the data hosting process on PANGAEA it will be ensured that all necessary metadata will accompany the data. No special software will be necessary to access these data, but serious analysis of them will require the use of any of the common open source command line programming languages (e.g. R, python, MATLAB, etc.). These software packages will not be included with the data as it is assumed that any researchers interested in these data will already have access to the aforementioned software and be proficient in their use.

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

Interoperability of the data will be ensured by adhering to the PANGAEA metadata vocabulary and by using established file formats. After data are created by the FACE-IT partners and made internally available, they will be taken up by the FACE-IT data scientist who will process them into the versions that will be made publicly available, meeting the standards outlined above. Because FACE-IT is a multidisciplinary project, all data generated will be processed to have a standard vocabulary that works well across disciplines.

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

Data generated by FACE-IT partners will be licensed under [creative commons license 3.0](#) unless explicitly required otherwise by third-party sources. This is in keeping with the standard for open data on PANGAEA, which is where all data generated for the FACE-IT project will be hosted in perpetuity. Data that are intended to be public will be made available for re-use outside of the FACE-IT project no more than 24 months after they become internally available. Regardless of the period of embargo, data will be uploaded to PANGAEA as soon as they are complete, but under a built-in embargo option on the service, which will be automatically lifted once the allotted time has passed.

By uploading data to PANGAEA it will ensure the re-usability of these data, including third-parties. This is ensured due to the rigorous quality review process that every dataset uploaded to PANGAEA must undergo. Funding has been set aside to cover the cost of the services provided by PANGAEA (see section 3). An editor from PANGAEA has been assigned to work with the data scientist for the FACE-IT project on each data upload to ensure that all data are FAIR. In theory, the data will remain re-usable *ad infinitum*. Data hosting methods are sure to change in the future, but as a trusted data hosting resource it can be assumed that PANGAEA will keep up with developments in data hosting.

3. Allocation of resources

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

The cost of ensuring that the data generated for FACE-IT are made FAIR are covered by the funds earmarked for data storage on the [PANGAEA](#) data hosting portal at €20,000. This data storage portal is certified by: ICSU World Data System, WMO Information System, and CoreTrustSeal.

The management of the data will be carried out by the data scientist for WP1, whose responsibility it is to first identify and provide the existing data required by members of the FACE-IT consortium for their related research purposes. The data generated by researchers are then transferred to and aggregated by the data scientist where they will

be made internally available via the [Nextcloud](#) and [pCloud](#) file transfer services. The finalised. data will then be processed to meet the standards for hosting on PANGAEA before being uploaded there.

There are no unforeseen costs in the long-term preservation of the data as the data hosting services at pCloud and PANGAEA have already been paid for in perpetuity. The value of storing data on PANGAEA is immense as it will ensure continued access to and use of the data in future research.

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

Data created for FACE-IT will initially be stored on the local machines of the respective investigators. These data will then be transferred as soon as finalised to the WP1 data scientist via SFTP where they will be aggregated and backed up on a local HDD as well as in the cloud services [Nextcloud](#) and [pCloud](#) for rapid dissemination within the FACE-IT consortium as desired. The Nextcloud service is a self-hosted file transfer protocol that is being hosted on a server instance at the University of Bremen. The pCloud cloud service is a secure lifetime service with servers exclusively within the EU. Both services have data hosting and sharing restricted by password protection. Sensitive data relating to the personal information of subjects involved in interviews etc. will not leave the local machines of the investigators who collect the data.

The long-term/post project storage of data will be on [PANGAEA](#). This data hosting portal is the natural science industry standard for long-term data management and adheres to the FAIR principles of open access data.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

The data generated for FACE-IT adhere to the guidance outlined in the [IASC Statement of Principles and Practices for Arctic Data management](#). For the physical, cryosphere, carbonate chemistry, biological, and ecological data generated there are no further ethical considerations as there are no human subjects in the experiments that generate these data. For the social science in FACE-IT, all activities shall meet the national legal and ethical requirements of the Arctic countries in which they are carried out. Below are the ethics requirements specifically for: 1) research related to humans, 2) protection of personal data, and 3) third countries.

1) Humans: The procedures and criteria that will be used to identify and recruit research participants for the community observing systems of FACE-IT are described below.

a) Procedures: The recruitment will be undertaken by the project in cooperation with representatives of the local communities. Participation will be entirely voluntary. The community members that the project will invite to participate in project workshops, discussions, sharing of knowledge, observations and experience are seen as 'co-creators' (and not as 'objects' of research).

b) Criteria: The project together with representatives of the local communities will identify who among the local communities may be interested in participating. The criteria will be the knowledge and experiences of the community members and their interest in participating. The aim is to obtain the participation of both men and women and different age classes, as they tend to use different natural resources and have knowledge about different topics.

2) Protection of personal data: Below we describe the justification for collection and processing of personal data, the procedures involved, and the information sheet.

a) Justification for collection and processing of personal data: Information on gender, age classes, and the experiences and interest of community members in community-based observing will be used to contact and recruit participants in relation to the community-based observing systems. Aside from this, personal information is not used in the project. No further personal data are collected.

b) Procedures: The project together with representatives of the local communities will identify who among the local communities may be interested in participating. The criteria will be the knowledge and experiences of the community members and their interest in participating. The aim is to obtain the participation of both men and women and different age classes (as they tend to use different natural resources and have knowledge about different topics). Information on gender, age, and their experiences and interest in community-based observing is not stored. The project will comply with the EU directive on data protection and with any updates it might receive during the lifetime of the project.

c) Templates of the information sheet/informed consent: The information sheet/informed consent form that will be communicated to the participants is included in Appendix 1. Note that as per the grant agreement this form is in Norwegian, rather than English, because it was decided to provide this type of material in local languages. If required, this document will be translated to Greenlandic and Danish.

d) The participants will be compensated for the time they use on community-based observing. In each area, every participant will be treated equally favourable and they will obtain the same daily compensation, irrespective of personal characteristics, e.g. religion, gender or age.

3) Third countries: The ethical standards and guidelines of Horizon2020 will be rigorously applied, regardless of the country in which the research is carried out. For the activities in Svalbard, the project will need a research permission from the local authorities. The project will obtain the research permission from the local authorities in Svalbard before the start of the research. The project will keep these permissions available at any time if needed for the European Commission representatives or for ethics reviewers.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

No procedures in addition to the Horizon2020 guidelines for data management, as outlined in this document, are currently proposed for the FACE-IT project.