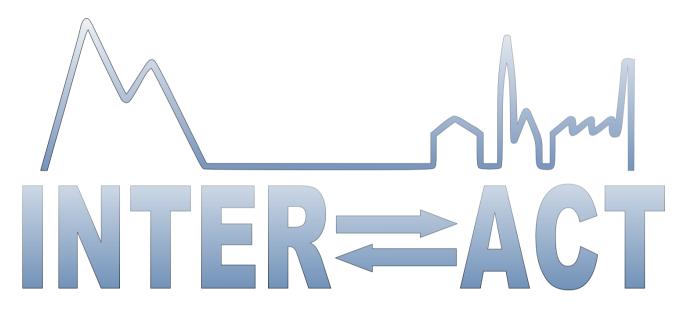# Integrating Activities for Advanced Communities

## D5.2- Virtual Access Assessment Report 2

Project No.730938– INTERACT

**H2020-INFRAIA-2016-2017/H2020-INFRAIA-2016-1**

Start date of project: 2016/10/01

Due date of deliverable: 2020/30/11 (M48)

Duration: 48 months

Actual Submission date: 2020/11/30

Lead partner for milestone: UOULU

Author: Kirsi Latola

| | Dissemination Level | |
|---|---|---|
| **PU** | Public | X |
| **PP** | Restricted to other programme participants (including the Commission Services) | |
| **RE** | Restricted to a group specified by the Consortium (including the Commission Services) | |
| **CO** | Confidential, only for members of the Consortium (including the Commission Services) | |

## Publishable Executive Summary

INTERACT is providing free on-line access to metadata and datasets collected and maintained by research infrastructures. In some cases, the datasets were already available on-line when the project started. However, at many sites this was not the case, and therefore INTERACT Virtual Access is the first time these datasets are openly accessible. At the end of month 48 of the project, 23 research stations had established open access to their data and provide VA through INTERACT VA Single-Entry Point. Virtual Access offered during the second half of the project period was assessed by the Virtual Access Board, which concludes that the current status of VA is a good starting point, and good progress has been made during this period. Especially, it is evident that a structured approach to interoperability at the discovery level has been initiated and that this will be further pursued in INTERACT 3. Interoperability at the data level is still pending, but interoperability at the discovery level is most important in the current situation. Furthermore, VAB appreciates the effort to support stations in achieving FAIR VA and the effort on data rescue that has been described where also the stations provide positive feedback.

## Background

Virtual Access (VA) is free on-line access to metadata and datasets collected and maintained by INTERACT infrastructures. In total, 23 INTERACT stations were funded to offer VA, and six planned to provide it as in-kind contribution. In the end of the RP3 and WP11 (VA) on M 48, twenty-three research stations had established open access to their data and provide VA through INTERACT VA Single-Entry Point during INTERACT 2. INTERACT VA is openly accessible via online single entry point at the INTERACT website (http://www.eu-interact.org/transnational-access/virtual-access/), where the VA is managed by the partners themselves.

The Virtual Access offered by INTERACT was assessed by the Virtual Access Board using the periodic report for RP3 (M37-48), which was prepared just prior to the Virtual Access Assessment as the main resource for evaluation. The composition of Virtual Access Board was set jointly with the WP4 Leader and it comprises of persons working with data management and thus with full understanding of the opportunities, benefits and challenges of providing on-line data.

The members of the Virtual Access Board are:
• Øystein Godøy, Norwegian Meteorological Institute, Norway, Chair (WP4 leader)
• Markus Fiebig, Norwegian Institute for air research, Norway, GAW expert through ET-WDC
• Kirsten Elger, Library and Information Services, GFZ German Research Centre for Geosciences, Potsdam, Germany

## Virtual Access Board's assessment report

The INTERACT Virtual Access Board (VAB) has evaluated the material provided. This includes the periodic report for WP 11 and the Virtual Access Single-Entry Point website at https://eu-interact.org/accessing-the-arctic/virtual-access/. These materials provide a good overview of the current status of Virtual Access within INTERACT.

It is still quite evident that the approach to virtual access differs much between the stations that offer it. Harmonisation in the reporting mechanisms (including definitions of terms like visitors/users, sessions etc.) between stations would still be beneficial for comparison of the various stations, also in order to generate KPI measures for the VA programme within INTERACT 3.

According to the periodic report, "Virtual Access (VA) means free access to stations' data and databases.". This calls for an evaluation of the VA with respect to the FAIR guiding principles which are promoted by the EU in this context and actively addressed through activities such as the European Open Science Cloud (EOSC) and OpenAire. VAB has evaluated INTERACT in terms of interoperability between data centres which support the ideas of EOSC and OpenAire. The current status of VA is a good starting point, and since the last report it is evident that a structured approach to interoperability at the discovery level has been initiated and that this will be further pursued in INTERACT 3. Interoperability at the data level is still pending, but VAB agrees that

interoperability at the discovery level is most important in the current situation. This along with a review and subsequent integration of components in the INTERACT Information and Communications Technology framework is promising for the future INTERACT VA. Furthermore, VAB appreciates the effort to support stations in achieving FAIR VA and the effort on data rescue that has been described where also the stations provide positive feedback. In the current implementation, VA is primarily supporting the FAIR concept of Accessible data, but good work is done on Findable data, a work that the VA in INTERACT 3 will benefit from. There are however major challenges on the concepts of Interoperable and Reusable data. A very important first step would be to guide stations on appropriate licenses for the data. This information should be visible both on the INTERACT VA entry point and at the stations.

VAB would thus repeat the recommendations from the previous review and add a new recommendation to make INTERACT VA to:

1.      Emphasising the principles of free and open data, including the IASC principle of ethically open data.

2.      Consolidating the current overview of the virtual access.

3.      Promoting interoperability at the discovery and data level at the stations through collaborations with existing archives/data repositories that are adhering to the FAIR guiding principles (ideally repositories that are using DOIs and are providing standardised metadata accessible via standardised interfaces). This also includes strategic collaboration of the stations with domain (e.g. GBIF) and/or variable oriented data repositories (e.g. ACTRIS, ICOS etc.) on reporting mechanisms, and would include exchange of metadata and data through standardised interfaces in standardised formats.

4.      Emphasising the need for identified and traceable data through application of Persistent Identifiers (e.g. DOI) combined with future utilisation of provenance structures developed through relevant communities in RDA and EU projects (e.g. ENVRIplus and ENVRI-FAIR). This can be achieved by collaboration with the relevant Research Infrastructures for those variables concerned and by collaborating with appropriate data repositories (e.g. PANGAEA) for other data types.

5.      Emphasising the need to clearly identify licenses for all data, both through information provided at the entry point (recommending a fallback license like Creative Commons - Attribution) and specifically at each station and dataset.

6.      Emphasising the need for adding clear information on how to cite the data at each data access point in addition to the request to acknowledge the programme and stations (for data with assigned DOI the citation information is included in the metadata).

That being said, VAB also acknowledges and strongly recommends to continue the data rescue activities described by several stations. These efforts should at any time complement the interoperability efforts by the stations.