Integrating Activities for Advanced Communities

D4.2- Report on current data flows

Project No.730938– INTERACT

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

Start date of project: 2016/10/01 Duration: 48 months

Due date of deliverable: 2017/10/01 Actual Submission date: 2017/10/01

Lead partner for deliverable: AWI

Author: Boris Radosavljevic, Anna Irrgang, Boris Biskaborn, and Øystein Godøy

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)	

List of Contents

Publishable Executive Summary **3**
Introduction **5**
Survey Method **6**
Results **6**
1.1 Questions and answers 7
Q4. Are you the responsible data manager for your station? 7
Q5. Are you reporting to a monitoring framework(s)? 7
Q6. If you answered yes on the previous question, or report to a non WMO-network, which network(s) are you reporting to? 8
Q7. Does your station have a clearly defined data policy? 11
Q8. Does your station have a clearly defined data management plan? 11
Q9. Are your data archived in a data center? 12
Q10. If you answered yes on the previous question, which data center(s) do you archive your data in? 12
Q11. Are your data described using standardized discovery metadata? 14
Q13. Are your discovery metadata searchable online? 15
Q14A. Do you maintain a catalog of your data? (e.g. an online catalog, an excel sheet, a database etc.), and if so, Q14B which online catalog(s)? 16
Q15. Are your data accessible through the discovery metadata? (i.e. does your metadata contain a URL to where data can be obtained) 17
Q16. Are your data available in a standardized format? 17
1.2 Summary of poll 18
1.3 Spatial analysis of answers 19
Discussion **22**
1.4 INTERACT Station Data Management in terms of FAIR 22
1.4.1 Is INTERACT data Findable? 22
1.4.2 Is INTERACT data Interoperable? 22
1.4.3 Is INTERACT data Reusable? 22
1.5 Survey design 23
1.6 Concluding comments on survey outcomes 23
Summary and Conclusions **24**
References **24**

Publishable Executive Summary

Research stations in the INTERACT network have acquired a considerable wealth of environmental data through ongoing monitoring efforts and projects. The relevance of research data stewardship and open access to well-curated and re-usable research data has been rising significantly over the past decade (European Commission, 2016; "G8 Open Data Charter and Technical Annex", 2013). Preserving research data for long-term use, including the storage in adequate repositories has been identified as a key issue by the scientific community, research agencies, and the public (Elger et al., 2016). A survey was conducted in order to assess data management practices, identify gaps and bottle necks in data flows at INTERACT. We received a representative sample of 64 stations, or 78% of stations in the network. Data management practices followed by stations in the INTERACT network represent two end members with roughly the same proportion. Either the stations' data management practices are adequate, or significant improvement is necessary. Significant gaps and bottlenecks in data flows were identified. Considerable lack of information on data and metadata standards and management exists in the network. A high number of stations are unclear on responsibilities and data management practices, and in many cases, data are not archived properly, i.e. long-term data security is not warranted.

Therefore,

- Information material should be provided by WP4 Data Forum
  - To improve understanding of data flows and standardization
  - To help individual stations implement sustainable data management practices in accordance with the data management plan
- Comprehensive guidelines on data management within INTERACT should be provided
  - An inventory of data types should be carried out to facilitate standardization of metadata standards and controlled vocabularies
  - An inventory of suitable repositories should be made
  - The data management plan should be amended accordingly

Furthermore, to facilitate data management aligned with FAIR principles, INTERACT should

- encourage usage of searchable data archives

- encourage utilization of data archives that provide unique identifiers
- survey the data archive landscape to ensure interoperability for metadata
- encourage usage of self-explaining file formats

Introduction

The relevance of research data stewardship and open access to well-curated and re-usable research data has been rising significantly over the past decade (European Commission, 2016; "G8 Open Data Charter and Technical Annex," 2013). Preserving research data for long-term use, including the storage in adequate repositories has been identified as a key issue by the scientific community, research agencies, and the public (Elger et al., 2016). However, the current state of data sharing and re-use requires considerable improvement.

High latitude stations participating in the INTERACT network have acquired a considerable wealth of environmental data through ongoing monitoring efforts and projects. The regions where these stations operate are especially sensitive to climate change. Therefore, data from INTERACT stations is highly relevant to the scientific community and the public. Integrating INTERACT data into relevant international data management frameworks like WMO Information System (serves for example GAW and, GCW), ICSU World Data System, and the Open Research Data Pilot (OpenAIRE) necessitates the development of common standards and a more unified approach to station metadata and data management at INTERACT stations. This key prerequisite represents a major task for INTERACT in the funding period of 2016-2020. The overarching goal for INTERACT data management is aligned with the FAIR Guiding Principles for Scientific Data Management: Data should be **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable (Box 1, (Wilkinson et al., 2016).

The first step included assessing current data management practices in order to identify gaps and bottlenecks in data flows (D4.2), as well as open access to data. For this purpose, we circulated a survey among station managers. The survey guided the development of a data management plan (D4.1). To tailor the plan to the needs of the community, the data management plan will be presented for discussion in the course of the general assembly in Svalbard, October 2017, where the INTERACT Data Forum (T4.2) will be established.

The Data Forum will foster dialog among station managers, internal and external data management specialists. It will provide a platform for collaboration, a place to get information and help. The paragraphs below show a detailed break-down of the survey responses followed by a discussion, summary, and recommendations.

The **FAIR** Guiding Principles

To be **Findable**:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be **Accessible**:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1 the protocol is open, free, and universally implementable
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be **Interoperable**:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be **Reusable**:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

Wilkinson et al 2016, doi: 10.1038/sdata.2016.18

Survey Method

The survey consisted of 16 multiple choice and free text questions. The questions addressed categories of the FAIR principles and were prepared by the group WP4, reviewed by representatives of the Station Manager Forum, and one external expert was consulted. The survey was mailed to station managers, but either the data manager or the station manager could reply.

Results

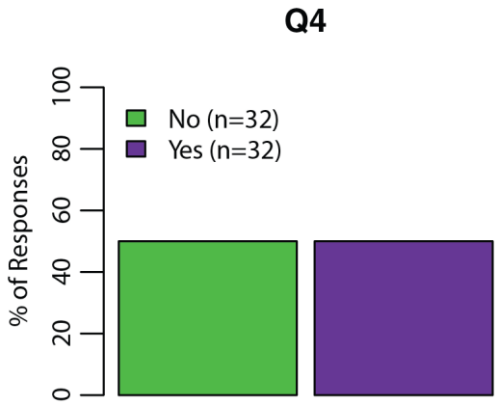
Sixty four (n=64) stations participated in the survey. At the time this document was written, INTERACT has grown to include 82 member and observer stations. The presented results are therefore based on 78% of all INTERACT and Observer Stations. The first three questions asked for the respondent's research station email address and the name and email address of the responsible data manager for the station. Apart from a short elaboration in Section 2.1 Q4, these questions are not relevant to this report. Other questions aimed to assess gaps and bottlenecks in data flows and curation, as well as to evaluate the extent current data management practices are aligned with the FAIR data principles.

1.1 Questions and answers

Q4. Are you the responsible data manager for your station?

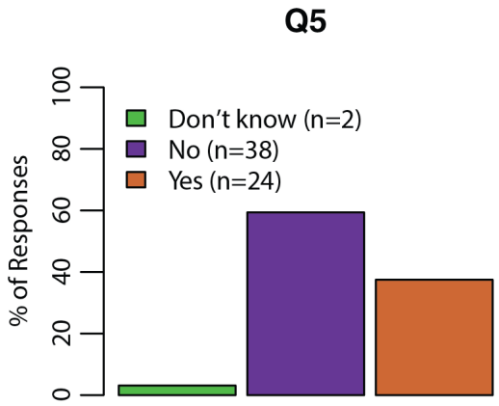
This question was answered by both data managers and others by an equal proportion. However, the survey was emailed to station managers with the request to have a data manager reply. The second question (not shown) inquired about the name of the data manager. Here is a summary:

- Seven (n=7) do not have a designated data manager.
- Three (n=3) stations indicated data collected is project owned
- Six (n=6) stations have multiple persons responsible for data management
- Forty eight (n=48) stations entered a name for the responsible data manager of the station.



Q5. Are you reporting to a monitoring framework(s)?

The purpose of this question was to identify participation in monitoring frameworks are e.g. various activities under the World Meteorological Organization (e.g. Global Atmosphere Watch, Global Cryosphere Watch, WMO Hydrological Observing System, Global Observing System, Global Terrestrial Observing System, Global Ocean Observing System, Global Climate Observing System etc.).



According to the responses, 59.4 % of stations are participating in monitoring frameworks, 37.5 % are not participating, while 3.1 % responses could not answer the question with certainty.

Q6. If you answered yes on the previous question, or report to a non WMO-network, which network(s) are you reporting to?

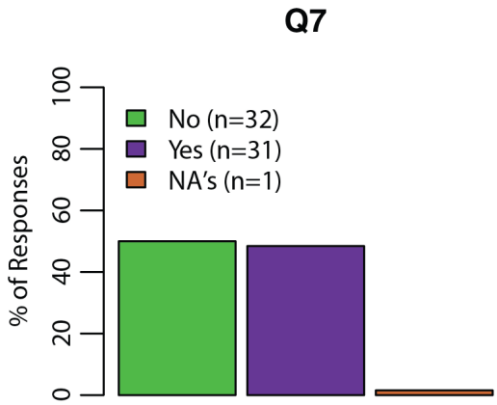
Fourteen respondents specified which observing networks their station reports to. Most represented are WMO networks such as Global Atmosphere Watch (GAW) and Global Cryosphere Watch (GCW).

Station name	**Network**
Abisko Scientific Research Station	Swedish Meteorological and Hydrological Institute, Geological Survey of Sweden, Swedish National Phenology Network, IVL Swedish Environmental Research Institute, Swedish University of Agricultural Sciences.
Actru station	GAW
ECN Cairngorms	UK, Environmental Change Network
Greenland Institute of Natural Resources	Several including e.g. NAFO, ICES, G-E-M.dk, GBIF, DMI, NASA, ASP, Promice, AMAP
Hintereis	WGMS
Hornsund Polish Polar Station in Spitsbergen	WMO, INTERMAGNET, WGMS, AERONET, NASA
Hyytiälä	GAW
Igarka	CALM

Geocryology Lab	
Karholl Aurora Observatory	GAW
Kevo Subarctic Research Institute	International Seismological Centre
Khibiny educational and research station	National Hydrometeorological network (Russia)
Kluane Lake Research Station	GCW
Nunavut Research Institute	Weather Network(lightning detection), Geological Survey of Canada (Geomagnetic Readings)
Pallas-Sodankylä Research Station	GAW, GCW, GRUAN: GCOS Reference Upper Air Network, GEWEX: Global Energy and Water Cycle Experiment,NDACC: Network for the Detection of Atmospheric Composition Change, EU-ICOS: Integrated Carbon Observation System, TCCON (Total Carbon Column Observing Network): A network of ground-based Fourier Transform Spectrometers recording direct solar spectra in the near-infrared spectral region; Copernicus Global Land Service, ESA GlobSnow and EUMETSAT H-SAF: Climate databases and near-real-time services for hemispheric snow mapping; Development of satellite systems for the monitoring of cryospheric processes and arctic atmosphere: Reference systems and measurements, e.g. for ESA SMOS, NASA SMAP, NASA OCO-2, NASA AURA OMI, NASA TERRA/AQUA MODIS, NASA/NOOA NPP VIIRS and OMPS, NASA/JAXA AMSR2, JAXA/MOE/NIES GOSAT, ESA/Copernicus Sentinel 1, 2 and 3.
Princess Elisabeth Antarctica	Each project that carries out monitoring reports to their own networks, e.g. for ozone, weather data, climate modelling. The Station does not control the data management of individual projects. When the Station staff carry out activities such as weather balloon launches, the data is sent to the WMO networks using the dedicated Station ID.
Research Station "Samoylov Island"	GTN-P
Sonnblick Observatory	GAW, GCW, GTS, NDACC, BSRN
Sverdrup Station	NPI report MET data to WMO, GNIP (Global Network of Isotopes in Precipitation). Our station is also responsible for maintaining hundreds of instruments from several other institutions. Many of these report to monitoring frameworks like GAW, etc., but they are not listed here.

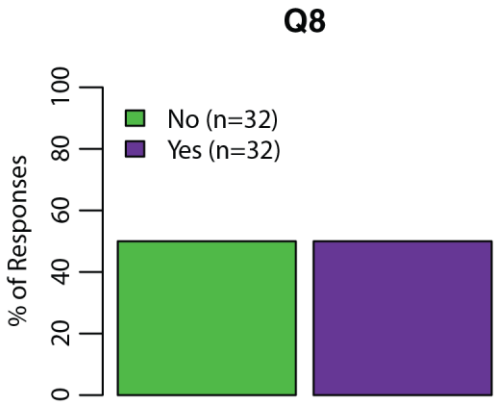
Tarfala Research Station	WGMS
Värriö	ICOS
Villum Research Station	GAW, AMAP, EMEP
Zackenberg, Greenland Ecosystem Monitoring	GEM is cross-disciplinary, and each of the sub-programs / disciplines reports individually to relevant networks such as: Program for Monitoring the Greenland Ice Sheet; World Hydrological Cycle Observing System; Global Runoff Data Centre; World Glacier Monitoring Service; Global Terrestrial Network – Permafrost; Arctic Climate Impact Assessment; Snow, Water Ice and Permafrost Assessment; Arctic Monitoring and Assessment Program; Circumarctic Active Layer Monitoring; Nordic Centre of Excellence; eScience Tools for Investigating Climate Change at High Northern Latitudes; Center for Permafrost, University of Copenhagen; Arctic Research Centre, Aarhus University; Integrated Carbon Observation System; Arctic Coastal Dynamics; Permafrost Carbon Network; Fluxnet; European Fluxes Database Cluster; Circumpolar Biodiversity Monitoring Program; Global Biodiversity Information Facility; International Tundra Experiment; Global Observation Research Initiative in Alpine Environments; Arctic Birds; Network for Arthropods of the Tundra; Herbivory network; ArcticWEB; Danish Meteorological Institute; US National Space Agency; European Space Agency; Arctic Science Partnership; DTU Space; Intergovernmental Panel on Climate Change; World Meteorological Organization - Global Cryosphere Watch; Global Terrestrial Network for Glaciers; International Hydrological Program – UNESCO; INTAROS – Integrated Arctic Observing System; Greenland Institute of Natural Resources

Q7. Does your station have a clearly defined data policy?



A data policy describes the principles for sharing data. It is complemented by licenses on the data like Creative Commons-BY or similar, but with statements indicating which types of data are restricted or not available freely. The results of the study indicate that roughly half of the stations do not have a data policy in place.

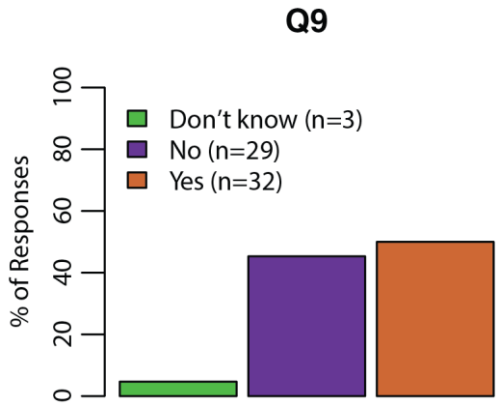
Q8. Does your station have a clearly defined data management plan?



A data management plan outlines the procedures of data curation, e.g. how data is described (metadata), what file format is archived and at which repository. Such plans aid

standardized data collection. Half of the stations in the network do not have a data management plan.

Q9. Are your data archived in a data center?



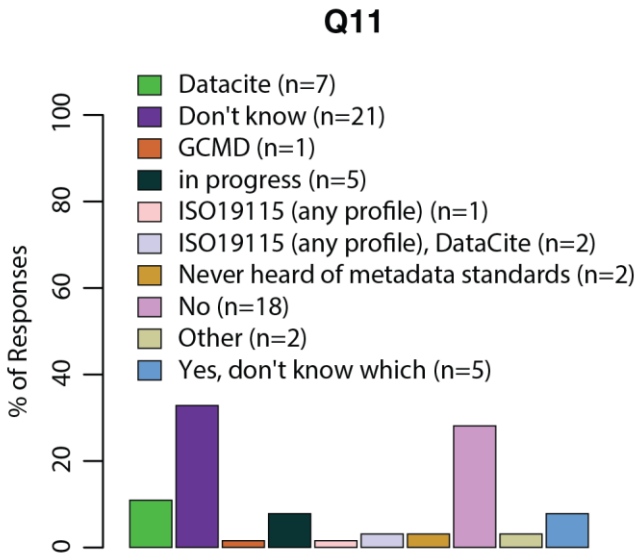
A data center should provide long term preservation the data. Most often data archived at data centers is accompanied by metadata enabling searchable data catalogs, ideally provide a Digital Object Identifier (DOI), and potentially direct access to data and metadata. Half of INTERACT Stations do not archive their data in such centers.

Q10. If you answered yes on the previous question, which data center(s) do you archive your data in?

The detailed responses provided an elaboration on Q9. This question also aimed at evaluating the long term sustainability of INTERACT data archived in data centers. Respondents could provide multiple answers. All instances of a data center being listed were counted in the table below. Although approximately half of the stations claim to utilize data centers, in some cases the data are stored locally, i.e. not findable and not accessible externally.

Responses	Data center
8	Canadian Geological Survey of Canada/Weather Network
8	Nordicana D
2	AVAA (Finland)
2	Dynamic Ecological Information Management System (DEIMS, EU)
2	Integrated Carbon Observation System (ICOS) Carbon Portal (EU)
2	Others
2	PANGAEA (Germany)
1	Arctic Portal (Iceland)
1	Arctic Science and Technology Information System (ASTIS, Canada)
1	Bolin Center Database, Stockholm university (Sweden)
1	Centre for Ecology and Hydrology (UK)
1	ECN Data Centre (UK)
1	Finnish Meteorological Institute
1	Finnish Meteorological Institute's satellite data center
1	Greenland Ecosystem Monitoring
1	International Arctic Research Center (IARC)
1	local servers
1	Netherlands Institute for Sea Research (NIOZ)
1	NILU ebas
1	Norwegian Bird Ringing Centre - Museum Stavanger
1	Norwegian Polar Data Centre
1	NSF Arctic Data Center
1	Open Data and Source Code of the Finnish Meteorological Institute
1	own database at Aarhus University
1	project databases
1	Research data descriptions discovery service of Natural Resources Institute Finland (RADAR)
1	Tilda - a university wide system
1	University of Oulu data services
1	World Glacier Monitoring Service

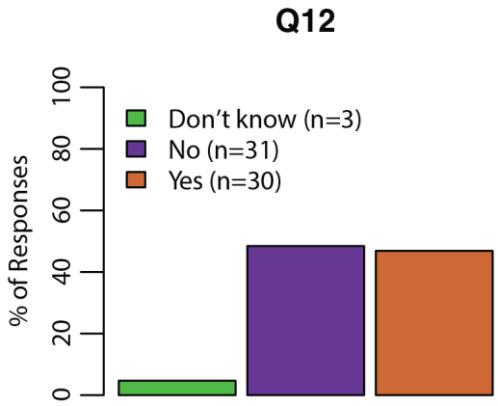
Q11. Are your data described using standardized discovery metadata?



Distributed data management relies on exchange of standardized discovery metadata (like index cards in libraries). We asked the respondents to select among some common standards or provide a free text answer. A majority of respondents could not name which, if any, metadata standards were used at the station. The responses indicate that 1) station managers must be educated on metadata standards; 2) a divergence of profiles in use exists, although most utilize either Datacite or ISO19115.

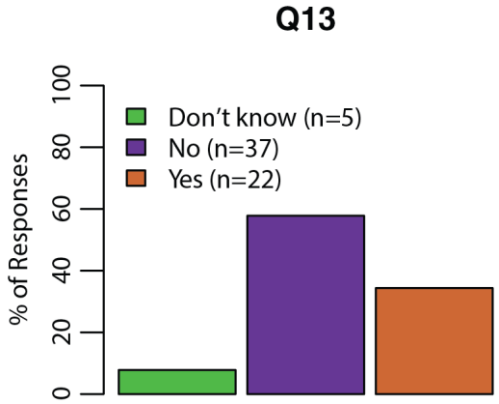
Q12. Does your discovery metadata contain information on how to cite the data?

Data citations are a way to credit the originator of the data and should be contained within the metadata. The results indicate most stations do not have such information provided in the metadata.



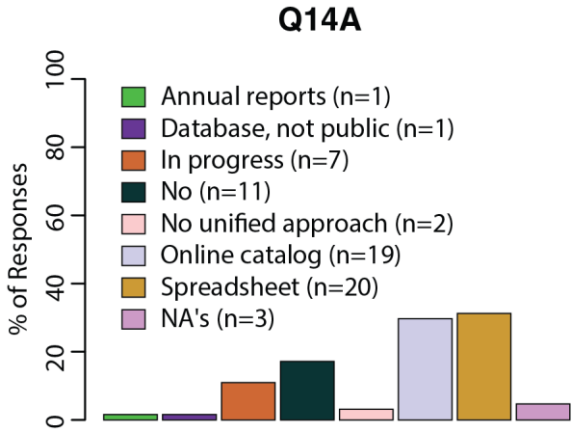
Q13. Are your discovery metadata searchable online?

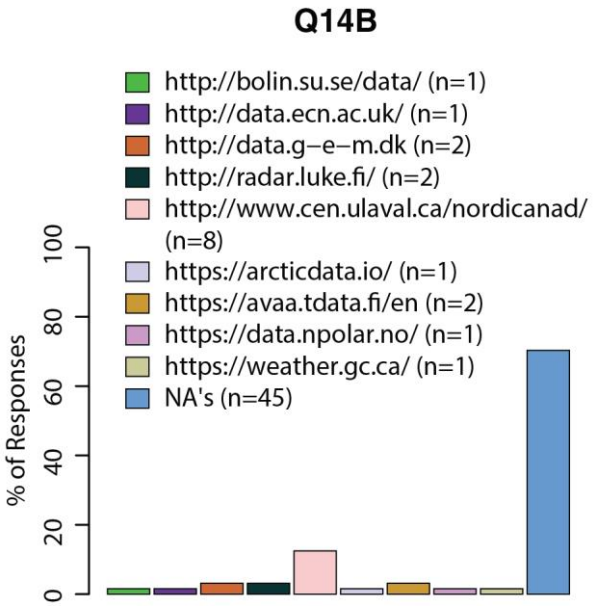
The question aimed to inquire about publicly accessible data catalogs. Indexing discovery metadata in searchable catalogs serving both human and machine users promotes station visibility in the global context. More than 65% of respondents either do not have such catalogs or do not know if discovery metadata is searchable online.



Q14A. Do you maintain a catalog of your data? (e.g. an online catalog, an excel sheet, a database etc.), and if so, Q14B which online catalog(s)?

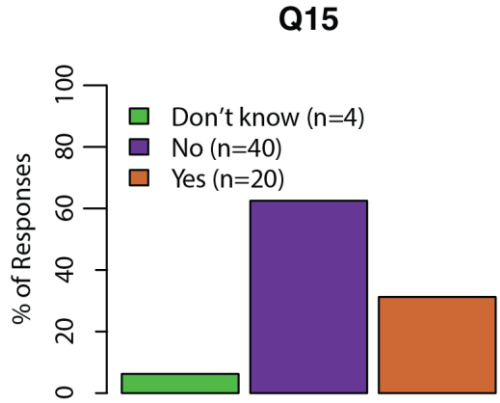
This question requested free text replies which were then sorted and edited to identify the nature of the catalog, and identify the online catalog. Results indicate that at least 19 stations maintain an online catalog. Many stations maintain spreadsheets with such information (local or cloud), but many do not maintain a catalog at all.





Q15. Are your data accessible through the discovery metadata? (i.e. does your metadata contain a URL to where data can be obtained)

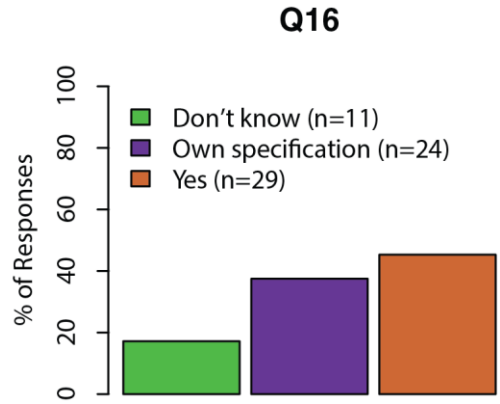
Discovery metadata of INTERACT station data typically does not contain a reference to where data could be obtained (URL).



D

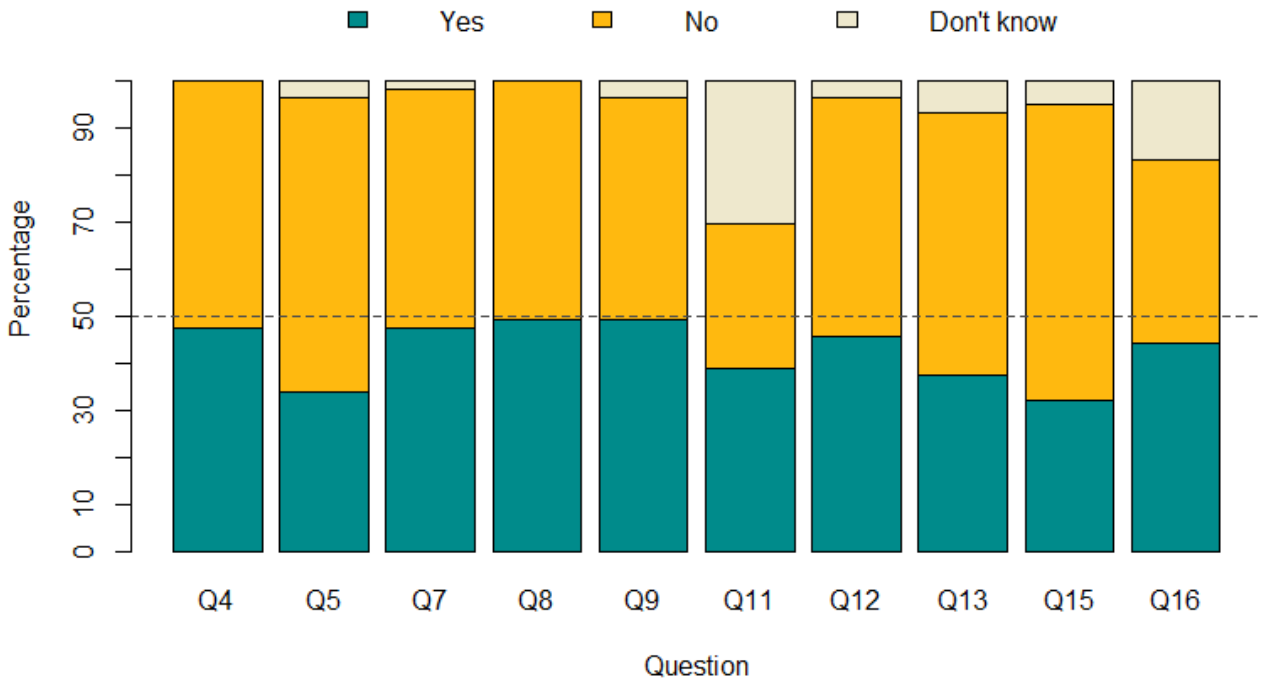
Q16. Are your data available in a standardized format?

Standardized formats ensure a common interpretation of the data. Based on the replies, most stations utilize some form of standardized formats.



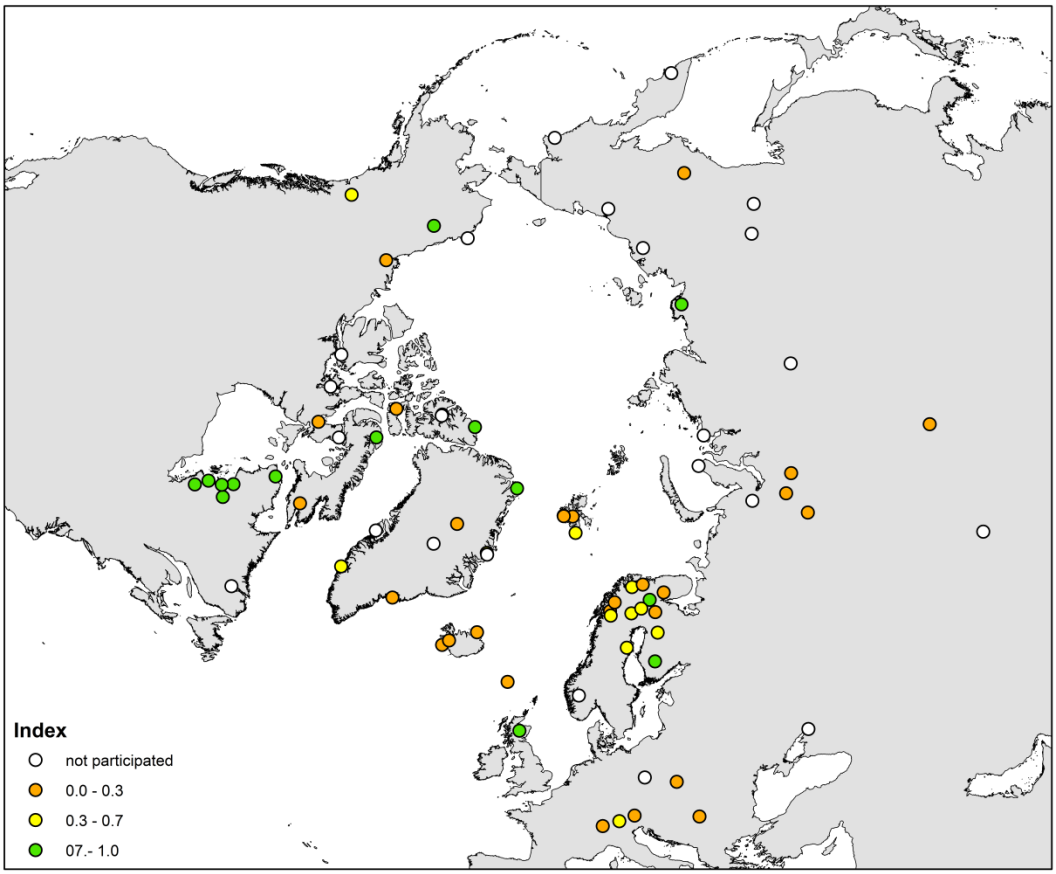
1.2 Summary of poll

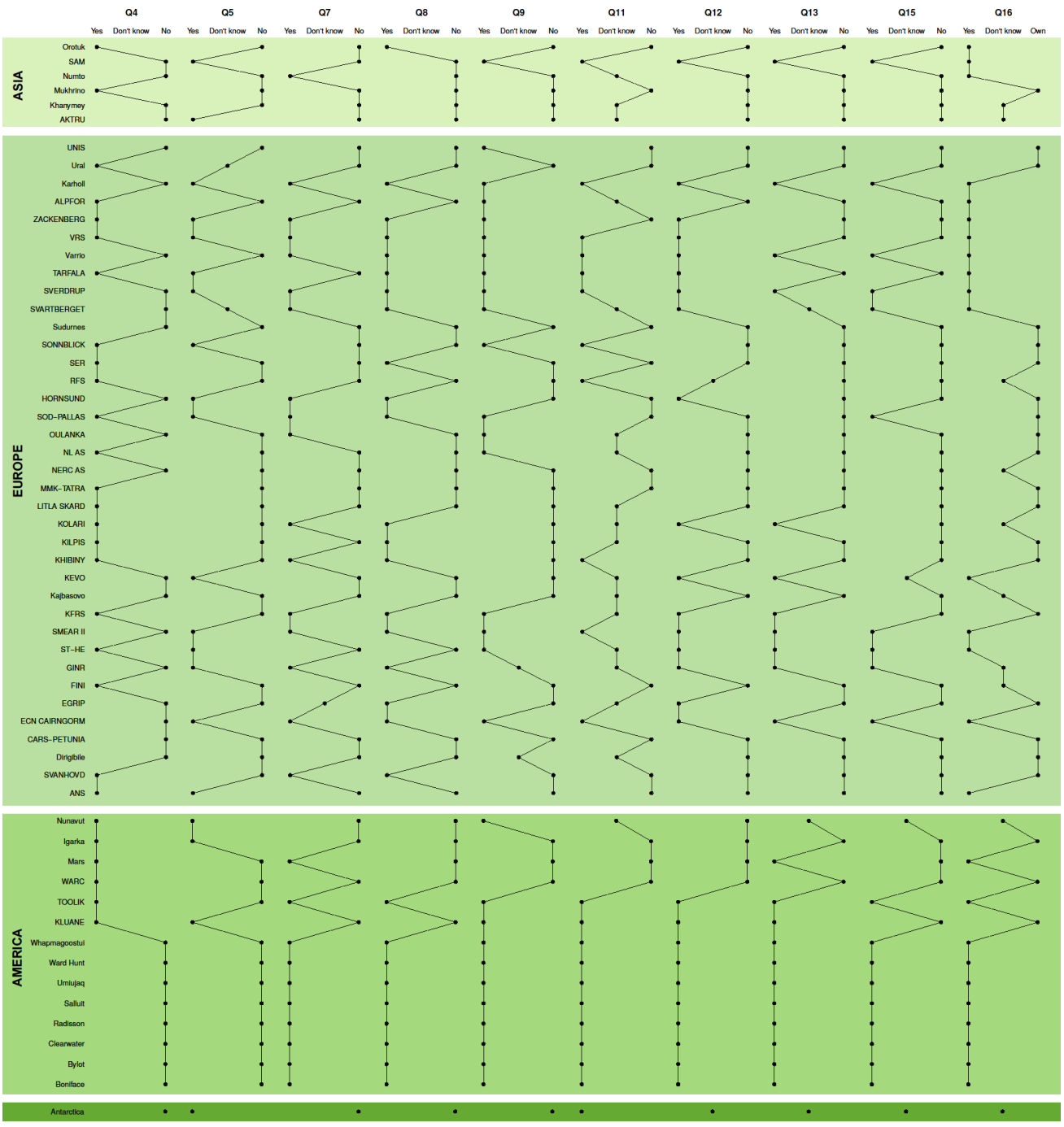
The bar chart below summarizes all answers which could be simplified to “yes”, ”no”, ”don’t know”. It shows that there is still considerable potential for the improvement of data management strategies. None of the questions could be answered by more than 50 % of the stations with “Yes”, meaning that less than 50 % of all stations which participated in the survey are reporting to a monitoring framework, have a clearly defined data policy, archive their data in a data center, or use standardized discovery metadata. The direct comparison of all answers reveals that especially questions about metadata (Q 11, Q 12, Q 13, Q 15) were answered with “No”, meaning that most of the stations do not have a good control over their metadata, yet. More in depth questions, like Q 11 (Are your data described using standardized discovery metadata?) were answered by over 20 % of the participants with “don’t know”, which might be an indicator for missing capacity for station-wide data management and/or missing awareness for the importance of station-wide data management.



1.3 Spatial analysis of answers

The figure below shows all answers listed by station. The stations were sorted by continent, starting with Asia, then Europe and North-America. Especially in Europe, answers were very heterogeneous. However, there was a tendency towards negative answers at stations based in Asia and positive answers of stations based in North America. Over double as many stations participated from North America then did from Asia, even though there are only three more INTERACT Stations in North America than in Asia. Thus, according to the survey, the main accumulation of station data flow bottlenecks are localized in the Eastern Arctic.




Q4
Q5
Q7
Q8
Q9
Q11
Q12
Q13
Q15
Q16
Yes
Don't know
No
Own
ASIA
Orotuk
SAM
Numto
Mukhrino
Khanymey
AKTRU
EUROPE
UNIS
Ural
Karholl
ALPFOR
ZACKENBERG
VRS
Varrio
TARFALA
SVERDRUP
SVARTBERGET
Sudurnes
SONNBLICK
SER
RFS
HORNSUND
SOD-PALLAS
OULANKA
NL AS
NERC AS
MMK-TATRA
LITLA SKARD
KOLARI
KILPIS
KHIBINY
KEVO
Kajbasovo
KFRS
SMEAR II
ST-HE
GINR
FINI
EGRIP
ECN CAIRNGORM
CARS-PETUNIA
Dirigibile
SVANHOVD
ANS
AMERICA
Nunavut
Igarka
Mars
WARC
TOOLIK
KLUANE
Whapmagoostui
Ward Hunt
Umiujaq
Salluit
Radisson
Clearwater
Bylot
Boniface
Antarctica

To get a better control over which stations have the highest development potential in terms of data management practices, based on the selection of answers displayed in the figure above, a simple index was created. If all nine questions were answered with “yes”, the station was ranked 1, stations with no positive answers were ranked 0. The indices were classified in three categories, going from satisfactory (green) to sufficient (yellow) and improvable (orange). Also, the stations which were not participating in the survey were displayed (white) (see map above).

Discussion

In general, the survey reflected a bimodality of data management practices existing in the INTERACT network. While many of the stations are well equipped in terms of data management personnel, infrastructure, and standards, the survey revealed significant gaps in data flows and bottlenecks. Even if a station has data management mechanisms in place, they still may not be FAIR. The survey did not attempt to investigate all aspects of FAIR data categories (particularly those related to repositories, e.g. A1.2, A2), but the principles (Box 1) still represent the yard stick for the survey.

1.4 INTERACT Station Data Management in terms of FAIR

1.4.1 Is INTERACT data Findable?

Questions 9, 11, 12, 13, and 14 assessed if data produced by stations is archived in data centers, that it be searchable online in online databases, and described using standardized discovery metadata that includes citation information. Responses to these questions reveal that more than half of the stations either do not archive their data in data centers, or do not know if they are. The usage of standardized metadata schema is not common practice. Furthermore, half of the stations discovery metadata do not include citation information (Q12). Utilization of searchable online databases should be encouraged in the network (Q13, Q14B) in order to make INTERACT data accessible.

Questions 14 and 15 inquired if data are catalogued, if these catalogs were online, and if data can be obtained by an URL in the metadata. Accessibility to INTERACT data represents a major data flow gap. Only 19 stations maintain online catalogs and provide access to data via the discovery metadata.

1.4.2 Is INTERACT data Interoperable?

Questions 15 and 16 addressed the interoperability of INTERACT data. As interoperability is also related to data archives' configurations (i.e. interoperability interfaces), the question of interoperability can only partially be answered. However, the survey results indicate that interoperability at the data and metadata levels is not established in the international stations network. Metadata does not include a URL to the data, and there is roughly a 50% chance the data will not be in a standardized format. The online repositories listed by the survey respondents also do not have interoperability interfaces (i.e. APIs) universally implemented.

1.4.3 Is INTERACT data Reusable?

Questions 5, 7, 9, 10, 11, 13, and 15 provide insight if data produced by INTERACT stations is suitable for reuse. We find a correlation between reporting to monitoring networks, archival in data centers, data policies, format and metadata standards, and reusability. The usage of standardized metadata including controlled vocabularies, the implementation of data policies, and standardized formats represent a major bottleneck in data flow at INTERACT.

1.5 Survey design

We expected that many respondents to the survey would not be familiar with the terminology. To clarify the questions, examples and online materials were included. Some station managers contacted WP4 for clarification, but most respondents completed the survey without further interaction. Given the responses it can be assumed that the content of the survey was understandable, although not universally. This is evident in Q11 where many respondents did not know the type of metadata standard used, or entered GCMD, which is not a metadata standard.

1.6 Concluding comments on survey outcomes

It is clear that many research stations need support to improve their data management practices and knowledge thereof. This is reflected in the high number of stations that do not have a data policy or management plan (Q7, Q8); that do not archive data in data repositories (Q9); that don’t know of, or never heard of, or do not utilize standardized metadata (Q11); do not securely catalog station data (Q14A); that do not have mechanisms in place to get credit for data (Q12) or get access to data (Q15). The efforts of WP4 should address this by providing appropriate information materials and training, exchange and support within the Data Forum. Possible reasons for the low participation of stations from Asia are, that either not enough effort was put in the communication with these stations, or language barriers hampered the fill-out of the questionnaire. A further possible reason might be that the topic of data management and data interoperability does not have such a high significance, yet. For the success of the INTERACT program, future work of WP4 has to integrate the Asian stations stronger and understand their needs better, in order to provide more suitable help in the development of their data management. Further regions which seem to have a high development potential are the Canadian High Arctic, Iceland, Eastern Siberia as well as south-east Europe However, a different reason for the overall performance of the stations might have been, that at many stations someone else than the responsible data manager filled out the questionnaire (Q4).

Summary and Conclusions

A survey aiming to assess data management practices, identify gaps and bottlenecks in data flows at INTERACT. We received a representative sample of 64 stations, or 78% of stations in the network. Data management practices in the INTERACT network fall into two camps with roughly the same proportion. Significant gaps and bottlenecks in data flows were identified.

Considerable lack of information on data and metadata standards and management exists in the network. A high number of stations are unclear on responsibilities and data management practices, and in many cases, data are not archived properly, i.e. long-term data security is not warranted.

Therefore:

- Information material should be provided by WP4 Data Forum
  - To improve understanding of data flows and standardization
  - To help individual stations implement sustainable data management practices in accordance with the data management plan
- To generate comprehensive guidelines on data management within INTERACT
  - An inventory of data types should be carried out to facilitate standardization of metadata standards and controlled vocabularies
  - An inventory of suitable repositories should be made
  - The data management plan should be amended accordingly

Furthermore, to facilitate data management aligned with FAIR principles, INTERACT should

- encourage usage of searchable data archives
- encourage utilization of data archives that provide unique identifiers
- survey the data archive landscape to ensure interoperability for metadata
- encourage usage of self-explaining file formats

Further, WP4 will encourage all stations to provide the contact of a responsible data manager, or data management team, to facilitate communication of future developments within the workpackage.

References

Elger, K., Biskaborn, B.K., Pampel, H., Lantuit, H., 2016. Open research data, data portals and data publication–an introduction to the data curation landscape. Polarforschung 85, 119–133. doi:10.2312/polfor.2016.009

G8 Open Data Charter and Technical Annex [WWW Document], 2013. URL https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex (accessed 9.30.17).

European Commission, 2016. Guidelines on FAIR Data Management in Horizon 2020. Version 3.0 [WWW Document]. URL http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf (accessed 10.30.17).

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018. doi:10.1038/sdata.2016.18