

ERA-PLANET Transnational Project Acronym:

iCUPE

DATA MANAGEMENT PLAN

Version 3

Alexander Mahura, Tuukka Petäjä, Hanna K. Lappalainen (University of Helsinki); Steffen M. Noe (Estonian University of Life Sciences); Stefano Nativi, Paolo Mazzetti (Consiglio Nazionale delle Ricerche), and DS Leaders¹

Helsinki, Finland, 15 November 2019

WP5: Data provision, interoperability and facilitation of data and services

Task 5.1: iCUPE data management plan / Deliverable 5.1.3: iCUPE data management plan (ver 3)

Version3

¹ see responsible persons DS Leaders in Appendix “iCUPE deliverables as datasets (DS)” of this Data Management Plan; appendix has contact information of persons responsible for delivering the datasets, and this information is available only at the internal iCUPE project website.

Table of Content

<i>Timetable of updates</i>	3
<i>Timetable of acronyms of organizations (partners and collaborators)</i>	4
<i>Current Data Management Plan</i>	4
List of iCUPE planned data collections / datasets	4
List of iCUPE teasers for data collections / datasets	5
List of iCUPE delivered data collections / datasets	6
<i>Data Collection Management Plans</i>	6
<i>Introduction</i>	7
1. Data Summary	7
1.1 Purpose of data collection/ generation & relation to project objectives	8
1.2 Types and formats of data generated/ collected	8
1.3 Existing data reuse (if any)	9
1.4 Origin of data	9
1.5 Expected data size	9
1.6 Data utility	9
2. FAIR (Findable, Accessible, Interoperable and Re-usable) Data	10
2.1. Making data findable, including provisions for metadata	10
2.2. Making data openly accessible	11
2.3. Making data interoperable	12
2.4. Increase data re-use (through clarifying licences)	13
3. Allocation of Resources	14
3.1. Data security	14
3.2. Ethical aspects	15
3.3. Other issues	16
Appendix: "iCUPE deliverables as datasets (DS)"	16
4. Current Status of Compliance with ERA-PLANET Data Management Principles and Constraints (status on 15 November 2019)	17
4.1. Compliance with H2020 Open Research Data Pilot	17
4.2. Contribution to GEOSS	18
Future work	27

Timetable of updates

DMP version	Author	Updates
version 1	Mahura et al.	November 2017 (based on original iCUPE project template)
version 2	Mahura et al.	November 2018 (based on DMP template designed for 4 projects of ERA-PLANET)
version 3	Mahura et al.	November 2019 (updated DMP with information on delivered datasets, their teasers)
version 3b	Mahura et al.	Final month of iCUPE project (final DMP – version beta– is expected by the end of the project)

Timetable of acronyms of organizations (partners and collaborators)

Acronym	Organization	Status in iCUPE
CNR-IIA	Institute for Atmospheric Pollution Research, National Research Council of Italy	Partner (Italy)
UHEL	University of Helsinki	Partner (Finland)
HZG	Helmholtz-Zentrum Geesthacht	Partner (Germany)
TROPOS	Leibniz Institute for Tropospheric Research	Partner (Germany)
NSCR	N.C.S.R. Demokritos, Institute of Nuclear Technology and Radiation Protection	Partner (Greece)
CNR-IDPA	Institute for Dynamics of Environmental Processes, National Research Council of Italy	Partner (Italy)
CNRS	Centre National de la Recherche Scientifique, Laboratoire Geosciences Environnement Toulouse	Partner (France)
CNR	National Research Council of Italy (<i>Consiglio Nazionale delle Ricerche</i>)	Partner (Italy)
AWI	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research	Partner (Germany)
PSI	Laboratory of Atmospheric Chemistry, Paul Scherrer Institute	Partner (Switzerland)
GFZ	Helmholtz Zentrum Potsdam	Partner (Germany)
INEP KSC RAS	Institute of Northern Environmental Problems, Kola Science Centre, Russian Academy of Sciences	Collaborator (Russia)
SIO RAS	P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences	Collaborator (Russia)
IAO SB RAS	V.E. Zuev Institute of Atmospheric Optics, Siberian Branch, Russian Academy of Sciences	Collaborator (Russia)
MSU	Lomonosov Moscow State University	Collaborator (Russia)
WMO	World Meteorological Organization	Collaborator (Switzerland)
NERSC	Nansen Environmental and Remote Sensing Centre	Collaborator (Norway)

Current Data Management Plan

List of iCUPE planned data collections / datasets

Data collection	Responsible organization
<i>List of the iCUPE deliverables as the datasets, DS (in a chronological delivering since beginning of the project; M01 = September 2017):</i>	
M16 – DS anthropogenic contaminants in snow from polar regions	(R - CNR-IDPA; D 2.1.1; T2.1)
M16 – DS emerging organic contaminants in air from the Arctic	(R – HZG; D 2.4.1; T 2.4)

M20 – DS ground based measurements for particle number, black carbon mass and ozone concentration	(R – CNR; D 1.1.2; T 1.1b)
M21 – DS Arctic parameters exactly based on ground-based remote sensing and airborne platforms	(R – TROPOS; D 1.3.2; T 1.3)
M22 – Pilot DS – NRT parameters of Arctic Research Infrastructures	(R – NSCR; D 1.2.1; T 1.2)
M22 – DS anthropogenic contaminants in ice cores	(R – CNR-IDPA; D 2.1.2; T 2.1)
M22 – DS Arctic atmospheric Hg isotope observations	(R – CNRS; D 2.2.3; T 2.2)
M22 – DS emerging organic contaminants in snow from the Arctic	(R – HZG; D 2.4.1; T 2.4)
M22 – DS emerging organic contaminants in water from the Arctic	(R – HZG; D 2.4.1; T 2.4)
M24 – DS snow spectral reflectance	(R – CNR; D 1.1.3; T 1.1c)
M24 – DS aerosol vertical profiles from ground-based and satellite observations in Finland and Russia	(R – CNRS & UHEL; D 4.3.1; T 4.3)
M28 – DS Arctic atmospheric Hg(II) observations: updated GMOS database	(R – CNRS; D 2.2.1; T 2.2)
M29 – DS blueprint for novel proxy variables integrating in-situ and satellite RS data with a exemplary dataset	(R – UHEL; D 4.1.1; T 4.1)
M30 – DS novel optical remote sensing products on snow & on vegetation and gas flaring mapping in selected sites	(R – CNR-IIA & GFZ; D 3.2.2; T 3.2)
M29 – DS precipitation in the high-latitudes	(R – UHEL; D 4.2.1; T 4.2)
M32 – DS time series of lakes’ size changes in Northeast Greenland	(R – AWI; D 3.3.2; T 3.3)
M32 – Pilot DS aerosol reanalysis for SMEAR-II	(R – UHEL; D 4.1.2; T 4.1)
M33 – DS organic aerosols in the Arctic (based on source apportionment)	(R – PSI; D 1.4.2; T 1.4)

See details <https://www.atm.helsinki.fi/icupe/index.php/datasets/list-of-datasets-as-deliverables>

List of iCUPE teasers for data collections / datasets

Data collection teaser	Responsible organization
List of teasers of the iCUPE deliverables as the datasets, DS	
• Fractional snow cover area in selected sites of Svalbard islands, Norway (D3.2.2 - teaser)	CNR-IIA
• Proxies for mixing layer height, condensation sink and gross primary production (D4.1.1 - teaser)	UHEL
• Dataset for ground-validation of precipitation measurements in high-latitudes and Arctic region, (D4.2.1 - teaser)	UHEL
• Multi-year dataset on mercury measurements at the Amderma station, Russian Arctic (teaser)	INEP KSC RAS
• Elemental and organic carbon over the northwestern coast of the Kandalaksha Bay of the White Sea (teaser)	SIO RAS
• DS on atmospheric composition at Fonovaya Observatory, West Siberia (teaser)	IAO SB RAS
	MSU, WMO, NERSC

• DS on micro-climatic features and Urban Heat Island Intensity in cities of Arctic region (teaser)	CNRS
• Arctic atmospheric Hg speciation and isotope observations (D2.2.1 & 2.2.3 - teaser)	AWI
• Time series of lake size changes in Northeast Greenland (D3.3.2 - teaser)	CNR-IDPA
• Concentration of organic contaminants, mercury and other heavy metals in annual snow and shallow core records (D2.1.1 & 2.1.2 - teaser)	PSI
• Source apportionment of organic aerosols in the Arctic including the source regions (D1.4.2 - teaser)	
• Occurrence, transport and exchange fluxes of emerging organic contaminants in the Arctic (D2.4.1 - teaser)	HZG
• Investigating the small-scale vertical and horizontal variability of the atmospheric boundary layer aerosol using unmanned aerial systems (D1.3.2 - teaser)	TROPOS
• Absorption Coefficient / Equivalent Black Carbon standardized dataset for long term impacts in the Arctic (D1.2.1 - teaser)	NSCR
• Continuous vertical observation of aerosol and cloud properties during the Polarstern cruises PS 106.1 and PS106.2 using a CLOUDNET station (D1.3.2-n2 - teaser)	TROPOS

See details: <https://www.atm.helsinki.fi/icupe/index.php/datasets/submitted-datasets>

List of iCUPE delivered data collections / datasets

Title of delivered data collection / dataset	Delivery month-year, (R - responsible organization; from deliverable D#; from task T#)
<i>List of iCUPE delivered data collections/ datasets, DS</i>	
• Dataset on emerging organic contaminants in air from the Arctic	Dec 2018 (R – HZG; D 2.4.1; T 2.4)
• Dataset on emerging organic contaminants in snow in the Arctic	May 2019 (R – HZG; D 2.4.1; T 2.4)
• Dataset of anthropogenic contaminants in snow from polar regions	Jun 2019 (R - CNR-IDPA; D 2.1.1; T2.1)
• Dataset of anthropogenic contaminants in ice cores from polar regions	Jun 2019 (R - CNR-IDPA; D 2.1.2; T2.1)
• Dataset on emerging organic contaminants in water from the Arctic	Sep 2019 (R – HZG; D 2.4.1; T 2.4)
• Pilot dataset on Near-Real Time aerosol absorption measurements from Zeppelin Station, Ny Ålesund, Svalbard	Sep 2019 (R – NSCR; D 1.2.1; T 1.2)

See details: <https://www.atm.helsinki.fi/icupe/index.php/datasets/delivered-datasets>

Data Collection Management Plans

Introduction

The iCUPE WP5 (“*Data provision, interoperability and facilitation of data and services*”) facilitates the data provision and services from the iCUPE project (www.atm.helsinki.fi/icupe) to the end-users, decision-makers, and stakeholders. Data obtained, integrated and accessed in WPs 1-4 are distributed in WP5 as data products and assessments. Data (ground-based, time-series, column, etc. observations) will be harmonized and assessed through developed novel methods, proxies and observables. These will be delivered through interoperability tools and services and will be available to iCUPE partners and any other end-users, decision-makers, and stakeholders (in agreement with existing data policies and open source principles).

To achieve the ERA-PLANET (European Network for Observing our Changing Planet; www.era-planet.eu) overall objectives, and to pursue cross-thematic interoperability and contribute effectively to GEOSS (Global Earth Observation System of Systems; www.earthobservations.org/geoss.php), the iCUPE project will implement the best practices and recommended approaches of ERA-PLANET. This will allow project to contribute to GEOSS via the GCI (GEOSS Common Infrastructure) and to utilize relevant Copernicus data and Core Services and EU capabilities in the EO domain. iCUPE will promote and implement the use of open specifications (i.e. international standards, community specifications) for data sharing and will foster technological development to deliver more timely and high quality data and information, in compliance with the GEOSS Data Management Principles.

The **iCUPE Data Management Plan (DMP)** (Task 5.1) will assure interoperability between the ERA-PLANET projects/ strands and with other activities carried out as part of GEO Strategic Plan (2017-2019) (i.e., GEO Initiatives, Flagships, Foundational tasks). As part of making iCUPE research data findable, accessible, interoperable and re-usable (FAIR), iCUPE DMP will include information: on handling of obtained research data (during and after the end of the project); on types of data to be collected, processed, analyzed, etc.; on applied methodological approaches and standards; on whether data will be shared and/or made open access; and on how data will be curated and preserved (including after the end of the project). The most important aspects for data management include: discoverability (data and metadata should be discoverable); accessibility (data should be accessible in online services); usability (encoding, traceability, documentation, quality); preservation (preservation, verification); and curation (review and reprocessing, persistent and resolvable identifiers).

Following the signed iCUPE Consortium Agreement /Section 11 on Data Management/, the appropriate and secure use of material and data of the project will be enabled according to the application of common standards. In the iCUPE project, the DMP will follow the "*Guidelines on FAIR Data management in Horizon 2020*" (version July 2016; ²) and according to the ERA-PLANET Data Management Plan (Deliverable 4.5). The iCUPE DMP will be updated during the lifetime of the project.

1. Data Summary

- *State the purpose of the data collection/generation & explain relation to 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)*

² http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

- *Outline the data utility: to whom will it be useful*

1.1 Purpose of data collection/ generation & relation to project objectives

The iCUPE addresses the overarching objective of ERA-PLANET to simplify access to information required by decision-makers and brings together and strengthen the European national and regional research and innovation programs in the EO domain. In particular, the iCUPE connects to the thematic strand 4 on polar areas and natural resources by integrating national and international monitoring and assessment activities in relation to ecosystem and environment quality in Arctic and Antarctic regions.

The work in iCUPE answers to the ERA-PLANET horizontal objectives and overarching goals of GEOSS and Copernicus integration by facilitating integration of the environmental pollution and sources and their transformation and impacts. The iCUPE platform takes on-board national agendas and improves co-alignment of national activities, and in particular, in the Arctic environmental observations. The iCUPE project endorses open data policies and Key Enabling Technologies (KET) via horizontal work within the three other ERA-PLANET strands. The iCUPE work provides ground-truthing for satellites in relation to snow, precipitation and aerosol observations.

The iCUPE will: 1) synthesize data from comprehensive long-term measurements, intensive campaigns and satellites, collected during the project or provided by on-going international initiatives; 2) relate the observed parameters to impacts; and 3) deliver novel data products, metrics and indicators to the stakeholders concerning the environmental status, availability and extraction of natural resources in the polar areas. The obtained results/ data will be useful for policy development and for improving and clearly communicating multidisciplinary understanding of status of the polar environment and pollution dynamics in the future.

1.2 Types and formats of data generated/ collected

The iCUPE will apply an integrated approach through combining and providing several types of in-situ, ground-based, airborne observations, satellite remote sensing and multi-scale modeling data. The produced datasets (DSs) for Arctic/Polar regions and for selected locations (to be specified for each dataset separately) will include information/ data on: in-situ, ground-based, remote sensing and airborne measurements; anthropogenic contaminants in snow and in ice cores; atmospheric Hg(II) and Hg isotope observations; organic contaminants in air; ground-based measurements for particle number, black carbon mass and ozone concentration; organic contaminants in snow and water; satellite derived cryospheric measurement data; snow spectral reflectance; aerosol vertical profiles from ground-based and satellite observations; blueprint for novel proxy variables integrating in-situ and satellite remote sensing data; novel optical remote sensing products on snow and on vegetation and gas flaring mapping; precipitation in the high-latitudes.

The methods and approaches for collecting mentioned above data are described in the iCUPE project Description of Work (DoW) as part of the project research plan. These standard collection methods and validated protocols are commonly used in the research field.

All DSs will be immediately available for the iCUPE consortium partners following a principle of "Internal use until publication", and then these DSs will be publicly available for all other potential users (e.g. decision-makers, stakeholders, end-users and other researchers, whom are not directly involved into iCUPE project activities).

A series of brief descriptions (1 page summary including data examples, contact information, etc. and extracts from the data available online together with the metadata) of planned DSs will appear (e.g. in advance before each planned deliverable expected as the dataset) at the iCUPE Datasets web-page (<https://www.atm.helsinki.fi/icupe/index.php/datasets/submitted-datasets>). This will allow potential users to learn and test applicability of DSs and to consider opportunities for establishing collaboration with the DSs owners (responsible for delivering DSs) as well as other iCUPE partners.

The formats and specifications for each individual dataset will be provided in the individual surveys on datasets /or data collections/ designed by ERA-PLANET for all 4 projects (*based on information provided by responsible persons for delivering the specific DSs*). Where it is applicable, the data formats may be migrated in case of new technologies will become available and are proved to be robust enough in order to ensure digital continuity and continued availability of data.

1.3 Existing data reuse (if any)

The iCUPE activities to a high degree is based on data currently residing in existing data and information repositories and these will be used to produce new compilations and other derived datasets as products. The possibility of data reusing is also considered, where it is applicable; and it will allow avoiding duplication of work already done. This is clarified during realisation of the iCUPE project and based on information provided for each DS by responsible persons for delivering such DS.

1.4 Origin of data

Data originate from various national data and information repositories as well as some are taken from on-going and planned observations and measurements during the project lifetime. Compilations of data is produced by iCUPE partners responsible for delivering and at some degree it is also linked with other ERA-PLANET strands/ projects. This is clarified during the realisation of the iCUPE project and based on information provided for each DS by responsible persons for delivering such DS.

1.5 Expected data size

The expected sizes of datasets (DSs) to be delivered as products became known after the delivery date and these are clarified during lifetime of the iCUPE project and based on information provided for each DS by responsible persons for delivering such DS. The expected size will depend on extend and nature of the data that is made available.

1.6 Data utility

The data generated by the iCUPE project will be useful for decision-makers, policy-makers, stakeholders, end-users as well as for other researchers (whom are not involved in the iCUPE project). In particular, iCUPE data can be widely utilized by decision-makers, governmental organizations, regional and local authorities, national environmental agencies and ministries, national weather and air pollution control services at national and European levels, etc. Among these are the research communities working on tasks linked with the air quality and climate modelling, and numerical weather prediction. Moreover, international in-situ observational networks such as PEEX, WMO-GAW, AMAP, INTAROS, and others (monitoring tasks and filling gaps in the polar regions) are potential users. And of course, concerned citizens (whom are interested in environmental problems) are also potential users, including both school and universities educational systems promoting various programmes for better understanding of environment and sustainable development of society in a changing climate

(and especially in most vulnerable polar regions). In addition, a detailed information is also available in the iCUPE Stakeholder Engagement Plan (Deliverable 6.1.1).

2. FAIR (Findable, Accessible, Interoperable and Re-usable) 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).*

Q1: Are the data produced and/or used in the programme discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

All data will be discoverable with the metadata. Many will also be identifiable and locatable by means of standard identification mechanisms. The iCUPE will be encouraged to consider Digital Object Identifiers (DOIs) for produced datasets.

Q2: What naming conventions do you follow?

The proposed for dataset, metadata, and template names, the iCUPE will define naming convention consisting the following key parts: i) prefix indicating if it is dataset, metadata or template; ii) root composed by: short and meaningful name of dataset/template & acronym or short name of the data provider (i.e. organization; for example: iCUPE - by default for templates); and iii) suffix indicating date of the last upload into the repository in YYYYMMDD (year-month-day format).

Q3: Will search keywords be provided that optimize possibilities for re-use?

The metadata system will provide opportunities for tagging the datasets and their content with keywords. In particular, the dataset information reported into the metadata will be published, where specific filters, based on the metadata elements, will allow to refine the search across datasets (e.g. search dataset by key words, by temporal or spatial coverage/location of data, by selected parameters or group of parameters, etc.).

Q4: Do you provide clear version numbers?

It is expected that many of delivered datasets will have version numbers. E.g. after the project ended (or during lifetime of the project), for example, for dynamic in nature datasets: new quality controlled measurement data can be added to already existing dataset represented as a time-series of observations. Although, for selected datasets some providers may not define the versions. This will be clarified during the project realisation. In general, the versioning management of data and files stored into the repository, will be realised via the naming convention and use of the date as a suffix (indicating the latest version of the file uploaded into the repository).

Q5: What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

The data will be documented following metadata standards. Each dataset documentation will include measurement stations/site basic information, explains terms, variable names, codes (or abbreviations used), etc. (questionnaire to be distributed among the persons/ partners responsible for the datasets). Considering possibility of DS usage in the future, a set of information needed to find, use and interpret the data and description of documentation's types (that will accompany DS) will be prepared. Metadata will provide standardized structured information with explanation of purpose, origin, time references, spatial (geographic) location, creator, access conditions and terms of dataset usage.

Consistency and quality of iCUPE data will be controlled and documented (on how data were collected). Data quality control will ensure that no data will be lost or accidentally changed. Such procedure is an integral part of DS creation and it takes place during data collection, entering or digitization, and checking. Quality control measures include e.g. standardized methods and protocols for making observations, alongside recording forms with clear instructions, calibration of instruments, etc. Missing data codes will be defined that actual data values fall within the range of expected values. Processed datasets will be reviewed by the iCUPE DMP responsible persons.

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*

Q1: Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

All produced datasets will be public of nature and shared. No restrictions are applied. Except, the raw data (used to produce datasets) will be available through request to the responsible person/ partners.

Q2: How will the data be made accessible (e.g. by deposition in a repository)?

Basic user requirements document will be posted on the iCUPE project website with instructions on how to access and use datasets.

Q3: What methods or software tools are needed to access the data?

Datasets will be accessible via HTTP and FTP by downloading/ ftp-ing files containing data and files with descriptions of datasets.

Q4: Is documentation about the software needed to access the data included?

Information about possible (preferably, publicly available) software/ tools to access datasets will be also included in description (as basic instructions on how-to-do and with corresponding web-links to technical aspects) of datasets produced.

Q5: Is it possible to include the relevant software (e.g. in open source code)?

A pointer to documentation on relevant standards can be included in open source code.

*Q6: Where will the data and associated metadata, documentation and code be deposited?
Preference should be given to certified repositories which support open access where possible.*

Datasets will be available through the iCUPE project website.

Q7: Have you explored appropriate arrangements with the identified repository?

Several options to be explored.

Q8: If there are restrictions on use, how will access be provided?

As the iCUPE produced datasets are all public in nature, there are no restriction on their usage. The iCUPE data to be stored are encouraging an unlimited and open data policy for non-commercial use. Data will be available and cited in publications. Interested persons will be able to contact with partners-owners of DSs (with CC to PI) for accessing data. For research and educational purposes, access to these data is unlimited and provided without a charge. By using such data, the person should accept that an offer of co-authorship will be made through personal contact with the owners whenever substantial use of such data is made. In all cases, an acknowledgement should be made to the owners and to the project name when these data are used within a publication.

Q9: Is there a need for a data access committee?

The iCUPE Project Office can take a role of the Data Access Committee, when it might be necessary. For example, for providing access to collaborators (i.e. researchers, whom are not involved into the iCUPE project) to specific dataset after discussion and agreement with the owner/ partner, whom produced DS.

Q10: Are there well described conditions for access (i.e. a machine readable license)?

To be specified and described at the later stage of the project.

Q11: How will the identity of the person accessing the data be ascertained?

Issues regarding authentication, authorization and accounting will be dealt with on general terms. The quality checked newly generated datasets (DS) including the metadata descriptions together with the harmonized database of long-term time-series produced in the iCUPE will be curated by the data producing partner, and saved and archived by national activities. The data will be accessible via iCUPE virtual platform (WP5).

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?*

Q1: Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Produced datasets will be interoperable allowing data exchange and reuse. Documentation for each dataset will be produced.

Q2: What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

iCUPE will use well-established European and international standards for storage, exchange and dissemination of produced data. Every dataset will have metadata.

Q3: Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

iCUPE will use standard vocabularies to an extent that they exist (or might be also novel ones developed in the project).

Q4: In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Although this is not anticipated to be an issue, but if it is found unavoidable, iCUPE will be required to document a usage of uncommon or to generate specific ontologies or vocabularies. In a case, if mapping to more commonly used ontologies will be possible, then such mapping will be required to establish.

2.4. Increase data re-use (through clarifying licences)

- *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.*

Q1: How will the data be licensed to permit the widest re-use possible?

Datasets will be publicly available. Information to be available at the later stage of the project. To be decided by owners/ partners of the datasets.

Q2: When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

It is not envisaged that iCUPE will seek patents. The data collected, processed and analyzed during the project will be made openly available following deadlines (for deliverables as the datasets) and once the corresponding peer-reviewed papers will be published. All datasets are expected to be publicly available by the end of the project.

Q3: Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

The iCUPE general rule will be that data produced after lifetime of the project will be useable by third parties. For shared information, standard format, open source software, and proper documentation will guarantee re-usability by third parties. Regarding the iCUPE digital research data, the ownership of raw data is on responsible partner for delivering dataset.

Q4: How long is it intended that the data remains re-usable?

The data are expected to remain re-usable (and maintained by the partner/ owner) as long as possible after the project ended, at least, 5 year period.

Q5: Are data quality assurance processes described?

Data quality assurance processes will be described before the public release of the dataset.

3. Allocation of Resources

Allocation of resources includes costs for making data FAIR (Findable, Accessible, Interoperable and Re-usable), responsibilities for data management, and resources for long-term data preservation. The iCUPE data management will be realized by Steffen Noe (WP5 leader, EULS), Alexander Mahura (UHEL), Stefano Nativi (CNR), Paolo Mazzetti (CNR) in close collaboration with the iCUPE Project Office as well as with active involvement of WP leaders and key researchers, whom are responsible for delivering data. Costs of making data FAIR are already included in this and other ongoing relevant projects as part data production and analysis. Time has been already allocated in WP5 for reach team by involvement of all partners, and it covers costs of preparing datasets and documentation for archiving.

Data will be stored at the coordinator's (UHEL) repository, and will be kept maintained, at least, for 5 years after the end of the project (with a possibility of further prolongation for extra years). UHEL repository will be managed and supported by a team of experts, and it is free of charge. UHEL will be setting up and upgrading, when it is needed, the hardware and software components of the storage repository; creation, maintenance and upgrading database for users accessing datasets; collecting users requests for access to and download of data; co-creation of the data repository's folders/sub-folders for each dataset and keeping document type (e.g. data, metadata, templates); capacity management of hardware and software components.

3.1. Data security

The iCUPE consortium partners endorse the open data policy as stated in the project DoW – all deliverables as datasets are "P" public of nature. The long-term storage of iCUPE project data (in form of datasets, DS) and legacy will be planned in more details during the project. Several options are considered, such as topical data storages, EBAS database (ebas.nilu.no), smartSMEAR

(www.atm.helsinki.fi/smartsMEAR), GEOSS portal (www.geoportal.org), World Data Centre of Aerosols (www.gaw-wdca.org), for mercury observations we could connect to GEO flagship Global Observation System for Mercury (GOS4M). The aerosol and trace gas observations will be continued in collaboration of ACTRIS Research infrastructure and its data services. In the future, PEEX offers continuation in the Arctic observations and integration of in-situ and satellite data in the Arctic context. For applicable data, another option could also be Copernicus Climate Change Service (C3S) that offers a possibility for operational and quality controlled information sharing. These aspects will be developed further in the evolving iCUPE data management plan).

For duration of the iCUPE project, data will be also stored at the project website (<https://www.atm.helsinki.fi/icupe/index.php/datasets>; initially - at internal, and then - at public domain) with corresponding links to access and description information. It will be managed and supported by a team of experts at the University of Helsinki (UHEL) and subject to the university's policies ensuring long-term security of the datasets, including version control and secure backups (i.e. there is no issue with the DSs recoverability). By depositing datasets, the iCUPE will ensure that the research data will be migrated to new formats, platforms, and storage media as required by good practice. The individual DOI's [=persistent identifier] will be generated enabling access to the iCUPE datasets via persistent links. The transfer of data is via a zip-process of distribution (at internal website – with access password separately distributed). Note that there are no iCUPE data of sensitive nature.

3.2. Ethical aspects

The iCUPE scientific focus is to provide novel insights and observational data on global grand challenges with a polar focus. The project is examining the environmental questions also having high ethical impact for the sustainable development of the polar regions. The project partners are in bringing forward knowledge about climate variability, climate change and adaptation to climate change. The iCUPE research approach will follow EU standards of ethical principles in its tasks and outcomes ensuring equality, quality and integrity in all conductions.

As an important part of the work also includes observations in Russia and use of such data available via collaboration between the Institute of Atmospheric Optics (IAO, Siberian Branch, Russian Academy of Sciences, SB RAS), CNRS, CNR and UHEL partners. For such iCUPE datasets, the data acquirement, storage, openness and sharing will be also based on requirements and guidelines of the participating Russian institutions. These partners are addressing the importance to adhere the ethical norms in research related to aspects of knowledge, truth, avoidance of error, falsifying, misrepresenting research data, promoting truth and minimizing error. The existing collaboration with Russian partners is based on well-established contacts through bi-lateral projects and contacts.

Another aspect is personal data (i.e. name-surname, e-mail address, telephone, and skype) of researchers responsible for producing and maintaining (after the end of the project) the iCUPE datasets (DSs). This information will be internally available for the project partners during lifetime of the project for more efficient work and discussions on creation, evaluation, testing, quality control, integration, etc. of DSs into the web-based system. For others (e.g. potential decision-makers, stakeholders, end-users as well as other researchers), only e-mail contact information will be available at the iCUPE Datasets web-page (<https://www.atm.helsinki.fi/icupe/index.php/datasets>).

Storage and access to iCUPE project DSs will be allowed. Dataset owner will have access to data and consortium partners will have access as far as it is needed to follow the project research plan. Intellectual Property Rights, in case when needed, will be agreed following the iCUPE Consortium Agreement.

3.3. Other issues

Concerning the iCUPE data management, the iCUPE project does not involve any issues raising security issues: i.e. there are no results obtained (including datasets to be delivered) raising such issues, and there are no EU-classified information as background or expected project results.

Information on suitable repositories for datasets produced by iCUPE can be also located using the Registry of Research Data Repositories (www.re3data.org) and Zenodo (zenodo.org) with providing tools to link publications and data. Information on research data management is also available from the Digital Curation Centre (www.dcc.ac.uk/dmponline) and ScienceMatters (www.sciencematters.io). In addition, the Research Data Alliance provides the Metadata Standards Directory (rd-alliance.github.io/metadata-directory) that can be searched for discipline-specific standards and associated tools, and the EUDAT B2SHARE (b2share.eudat.eu) tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.

Appendix: “iCUPE deliverables as datasets (DS)”

Appendix has contact information of persons responsible for delivering the datasets, and this information is available only at the internal iCUPE project website.

4. Current Status of Compliance with ERA-PLANET Data Management Principles and Constraints (status on 15 November 2019)

This section summarizes the current status of compliance with the ERA-PLANET Data Management Principles and constraints. Plans for future compliance can be provided in the comment sections.

At the final release of the DMP, all answers should be YES, or NO with clear justification.

Data generated in

4.1. Compliance with H2020 Open Research Data Pilot

Concerning "**Compliance with H2020 Open Research Data Pilot**", the first datasets are expected to be available in December of 2018, and hence, there are no data collection or dataset yet uploaded in repositories. At current stage, only teasers of selected future datasets are available (see <https://www.atm.helsinki.fi/icupe/index.php/datasets/submitted-datasets>) from partners and collaborators of the iCUPE project. Concerning scientific results and publications, the produced datasets will be publicly available free of charge with information on required tools and instruments.

Data generated in the project are available in a **repository with proper metadata**:

Yes

Partially: **_33_** % of datasets are available in a repository with proper metadata

No

Comment/Justification: (Not Available Yet = N, Available = Y)

	Data collection / dataset (DS)	Repository	Metadata	Responsible
1.	Anthrop. contaminants in snow ...	Y	Y	1. CNR-IDPA
2.	Emerg. organ. contaminants in air ...	Y	Y	2. HZG
3.	Meas. for particle number, BC, O3 ...	N	N	3. CNR
4.	Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5.	NRT parameters of Arctic Research Infrastructures	y	y	5. NSCR
6.	Anthropogenic contaminants in ice cores	Y	Y	6. CNR-IDPA
7.	Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8.	Emerg. organic contaminants in snow from Arctic	Y	Y	8. HZG
9.	Emerg. organic contaminants in water from Arctic	Y	Y	9. HZG
10.	Snow spectral reflectance	N	N	10. CNR
11.	Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12.	Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13.	Blueprint for novel proxy variables ...	N	N	13. UHEL
14.	Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15.	Pecipitation in the high-latitudes	N	N	15. UHEL
16.	Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17.	Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18.	Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data generated in the project are available **free-of-charge**:

Yes (100%) – all iCUPE dataset are expected to be free-of-charge available

Partially: % of datasets are available free-of-charge

No

Comment/Justification:

*Scientific results (e.g. publications) of the project are reproducible (data are available free-of-charge, information on **required tools and instruments** are provided):*

Yes

Partially: **_33_ % of scientific results are reproducible**

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Info on Tools	Info on Instrum	Responsible
1. Anthrop. contaminants in snow ...	Y	Y	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	Y	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	y	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	Y	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	Y	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	Y	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

4.2. Contribution to GEOSS

Concerning "**Contribution to GEOSS**", the first datasets are expected to be available in December of 2018, and hence, there are no data collection or dataset yet uploaded in repositories. At current stage, only teasers of selected future datasets are available (see <https://www.atm.helsinki.fi/icupe/index.php/datasets/submitted-datasets>) from partners and collaborators of the iCUPE project.

Are **data** of interest for GEO stakeholders (e.g. scientists, policy-makers) discoverable and accessible through GEOSS?

Yes, they are **accessible through GEOSS** or through a **repository registered in the GEO Yellow Pages**

Partially: **_IN PROGRESS_ % of required datasets are accessible through GEOSS or through a repository registered in the GEO Yellow Pages**

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Access through GEOSS	Registered in GEO Yellow Pages	Responsible
1. Anthrop. contaminants in snow ...	N	N	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	N	N	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	N	N	5. NSCR
6. Anthropogenic contaminants in ice cores	N	N	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	N	N	8. HZG
9. Emerg. organic contaminants in water from Arctic	N	N	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data accessible through GEOSS are available as GEOSS DATA-Core?

Yes

Partially: **_IN PROGRESS_** % of datasets available through GEOSS are GEOSS DATA-Core

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Access through GEOSS	Available as GEOSS DATA-Core	Responsible
1. Anthrop. contaminants in snow ...	N	N	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	N	N	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	N	N	5. NSCR
6. Anthropogenic contaminants in ice cores	N	N	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	N	N	8. HZG
9. Emerg. organic contaminants in water from Arctic	N	N	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL

14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data and all associated metadata are discoverable through catalogues and search engines, and data access and use conditions, including licenses, are clearly indicated

Yes

Partially: **33%** of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Discov. through catalogues & search engines	Access, use, license – clearly indicated	Responsible
1. Anthrop. contaminants in snow ...	Y	Y	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	Y	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	y	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	Y	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	Y	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	Y	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data are accessible via online services, including, at minimum, direct download but preferably user-customizable services for visualization and computation:

Yes

Partially: **22%** of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Direct download	User-custom service	Responsible
1. Anthrop. contaminants in snow ...	IntWebSite	N	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	N	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	N	5. NSCR
6. Anthropogenic contaminants in ice cores	IntWebSite	N	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	N	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	N	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data are structured using encodings that are widely accepted in the target user community and aligned with organizational needs and observing methods, with preference given to *non-proprietary international standards*.

Yes

Partially: **33 %** of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Structured using accepted encodings	Non-proprietary internat. standards	Responsible
1. Anthrop. contaminants in snow ...	Y	Y	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	Y	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	y	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	Y	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	Y	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	Y	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI

17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data are comprehensively documented, including all elements necessary to access, use, understand, and process, preferably via formal structured metadata based on international or community-approved standards. To the extent possible, data are also **described in peer-reviewed publications** referenced in the metadata record.

Yes

Partially: **_33_ % of datasets**

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Compreh. docum.	Descr. peer-reviewed paper	Responsible
1. Anthrop. contaminants in snow ...	Y	in progress	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	in progress	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	in progress	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	in progress	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	in progress	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	in progress	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data include provenance metadata indicating the **origin and processing history of raw observations and derived products**, to ensure full traceability of the product chain.

Yes

Partially: **_IN PROGRESS_ % of datasets**

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Origin & history of process. raw observ.	Origin & history of process. derived products	Responsible
1. Anthrop. contaminants in snow ...	in progress	in progress	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	in progress	in progress	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	in progress	in progress	5. NSCR
6. Anthropogenic contaminants in ice cores	in progress	in progress	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	in progress	in progress	8. HZG
9. Emerg. organic contaminants in water from Arctic	in progress	in progress	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	N	18. PSI

Data are quality-controlled and the results of quality control shall be indicated in metadata; data made available in advance of quality control will be flagged in metadata as unchecked.

Yes

Partially: **33 % of datasets**

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Data are quality controlled	Results of quality control	Responsible
1. Anthrop. contaminants in snow ...	Y	N	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	N	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	Y	N	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	N	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	N	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	N	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI

17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source appart.)	N	N	18. PSI

Data are protected from loss and preserved for future use; preservation planning will be for the long term and include guidelines for loss prevention, retention schedules, and disposal or transfer procedures:

Yes

Partially: **33%** of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Protected from loss	Preserved for future	Responsible
1. Anthrop. contaminants in snow ...	Y	Y	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	Y	2. HZG
3. Meas. for particle number, BC, O3 ...	N	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	y	y	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	Y	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	Y	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	Y	9. HZG
10. Snow spectral reflectance	N	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	N	14. CNR-IIA, GFZ
15. Pecipitation in the high-latitudes	N	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	N	17. UHEL
18. Org. aerosols in Arctic (based on source appart.)	N	N	18. PSI

Data and associated metadata held in data management systems are periodically verified to ensure integrity, authenticity and readability:

Yes

Partially: **33%** of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Data period. verified	Metadata period. verified	Responsible

1.	Anthrop. contaminants in snow ...	Y	Y	1.	CNR-IDPA
2.	Emerg. organ. contaminants in air ...	Y	Y	2.	HZG
3.	Meas. for particle number, BC, O3 ...	N	N	3.	CNR
4.	Arctic param. on rem.sens. & airborne platforms	N	N	4.	TROPOS
5.	NRT parameters of Arctic Research Infrastructures	Y	Y	5.	NSCR
6.	Anthropogenic contaminants in ice cores	Y	Y	6.	CNR-IDPA
7.	Arctic atmospheric Hg isotope observations	N	N	7.	CNRS
8.	Emerg. organic contaminants in snow from Arctic	Y	Y	8.	HZG
9.	Emerg. organic contaminants in water from Arctic	Y	Y	9.	HZG
10.	Snow spectral reflectance	N	N	10.	CNR
11.	Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11.	CNRS, UHEL
12.	Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12.	CNRS
13.	Blueprint for novel proxy variables ...	N	N	13.	UHEL
14.	Novel optical rem. sens. snow and vegetation ...	N	N	14.	CNR-IIA, GFZ
15.	Precipitation in the high-latitudes	N	N	15.	UHEL
16.	Time series of lakes' size changes in NE Greenland	N	N	16.	AWI
17.	Pilot DS aerosol reanalysis for SMEAR-II	N	N	17.	UHEL
18.	Org. aerosols in Arctic (based on source apport.)	N	N	18.	PSI

Data are managed to perform **corrections and updates** in accordance with reviews, and to enable **reprocessing** as appropriate; where applicable this follows established and agreed procedures:

Yes

Partially: **33** % of datasets

No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Correct., updates	Re-process.	Responsible		
1.	Anthrop. contaminants in snow ...	Y	Y	1.	CNR-IDPA
2.	Emerg. organ. contaminants in air ...	Y	Y	2.	HZG
3.	Meas. for particle number, BC, O3 ...	N	N	3.	CNR
4.	Arctic param. on rem.sens. & airborne platforms	N	N	4.	TROPOS
5.	NRT parameters of Arctic Research Infrastructures	Y	Y	5.	NSCR
6.	Anthropogenic contaminants in ice cores	Y	Y	6.	CNR-IDPA
7.	Arctic atmospheric Hg isotope observations	N	N	7.	CNRS
8.	Emerg. organic contaminants in snow from Arctic	Y	Y	8.	HZG
9.	Emerg. organic contaminants in water from Arctic	Y	Y	9.	HZG
10.	Snow spectral reflectance	N	N	10.	CNR
11.	Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	N	11.	CNRS, UHEL
12.	Arctic atmos. Hg(II) obs.: updated GMOS database	N	N	12.	CNRS
13.	Blueprint for novel proxy variables ...	N	N	13.	UHEL
14.	Novel optical rem. sens. snow and vegetation ...	N	N	14.	CNR-IIA, GFZ
15.	Precipitation in the high-latitudes	N	N	15.	UHEL
16.	Time series of lakes' size changes in NE Greenland	N	N	16.	AWI
17.	Pilot DS aerosol reanalysis for SMEAR-II	N	N	17.	UHEL
18.	Org. aerosols in Arctic (based on source apport.)	N	N	18.	PSI

Data are **assigned** appropriate persistent, resolvable **identifiers** to enable documents to cite the data on which they are based and to enable data providers to receive acknowledgement of use of their data:

Yes Partially: **33** % of datasets No

Comment/Justification: (Not Available Yet = N, Available = Y)

Data collection / dataset (DS)	Assign identifier	Responsible
1. Anthrop. contaminants in snow ...	Y	1. CNR-IDPA
2. Emerg. organ. contaminants in air ...	Y	2. HZG
3. Meas. for particle number, BC, O3 ...	N	3. CNR
4. Arctic param. on rem.sens. & airborne platforms	N	4. TROPOS
5. NRT parameters of Arctic Research Infrastructures	Y	5. NSCR
6. Anthropogenic contaminants in ice cores	Y	6. CNR-IDPA
7. Arctic atmospheric Hg isotope observations	N	7. CNRS
8. Emerg. organic contaminants in snow from Arctic	Y	8. HZG
9. Emerg. organic contaminants in water from Arctic	Y	9. HZG
10. Snow spectral reflectance	N	10. CNR
11. Aerosol vert. profiles gr.-based & sat.obs. in FI+RU	N	11. CNRS, UHEL
12. Arctic atmos. Hg(II) obs.: updated GMOS database	N	12. CNRS
13. Blueprint for novel proxy variables ...	N	13. UHEL
14. Novel optical rem. sens. snow and vegetation ...	N	14. CNR-IIA, GFZ
15. Precipitation in the high-latitudes	N	15. UHEL
16. Time series of lakes' size changes in NE Greenland	N	16. AWI
17. Pilot DS aerosol reanalysis for SMEAR-II	N	17. UHEL
18. Org. aerosols in Arctic (based on source apport.)	N	18. PSI

Future work

This section reports the plans of actions for improving compliance for the next version of the DMP.

The nearest future work will include:

- All remaining (6 delivered) iCUPE datasets to be delivered will be uploaded and linked at the iCUPE project website <https://www.atm.helsinki.fi/icupe/index.php/datasets/delivered-datasets>
- Final reportings on the ERA-PLANET surveys on each iCUPE dataset are expected by each dataset leader than dataset delivered and uploaded in the repository. Surveys are aimed at collecting information about data sharing and data management conditions and procedures for all ERA-PLANET products (and hence, for the iCUPE products as well).
- Before uploading iCUPE datasets into research data repository and providing datasets for public access, the corresponding teasers for datasets will be linked and displayed at <https://www.atm.helsinki.fi/icupe/index.php/datasets/submitted-datasets>
- Once each iCUPE dataset /or data collection/ will become uploaded in the repository and become publicly available, the final update of surveys on each dataset will be done.
- The iCUPE generated datasets will be checked and updated on compliance with H2020 Open Research Data Pilot as well as dataset status on contribution to GEOSS.
- By the end of the project, the status on compliance with ERA-PLANET Data Management Principles and constraints will be updated for each iCUPE dataset accordingly in this Data Management Plan. Target is that majority of iCUPE datasets will be at 100% level of compliance with proposed data management principles.