

DATA MANAGEMENT PLAN DU PROJET "OBS4CLIM - SYSTEME D'OBSERVATION INTEGRE POUR L'ATMOSPHERE"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - DMP template (english)" fourni par Agence nationale de la recherche (ANR).

INFORMATION ON THE PLAN

Plan title DMP of the project "Obs4Clim - Système d'Observation Intégré pour l'Atmosphère"
Version First version
Plan purpose/scope

The goal of the OBS4CLIM Data Management Plan (DMP) is to i) document the key elements of the data management life cycle, ii) describe the data collected, processed and generated in the course of the OBS4CLIM project, iii) detail the technical solutions to manage, archive, curate and access these data, and vi) outline the strategy and development needed for making OBS4CLIM data FAIR. OBS4CLIM data is embedded in an integrating observing system tying the French atmospheric research infrastructures (RI) ACTRIS-FR, ICOS-France and IAGOS-France for providing data integrated within the atmospheric data centre AERIS, in close alignment with the GAIA Data project and the digital RI DATA TERRA. On European level, the OBS4CLIM DMP builds on the interoperable framework for FAIR data which is jointly developed by the atmospheric landmark RIs ACTRIS, IAGOS, and ICOS on European level as part of the ENVRI-FAIR project and which is currently under implementation. Through its investment and service developments, OBS4CLIM will enhance the capacity of the collaborating atmospheric RIs to provide state-of-the-art data and innovative data services addressing the needs of the national and international user communities.

The DMP is a living document that will be updated regularly. The goal is to make the DMP accessible for all stakeholders (repository operators, funders, researchers, publishers, infrastructure providers etc.).

Fields of science and technology (from OECD classification)

Earth and related environmental sciences

Language

eng

Creation date

2022-03-04

Last modification date

2022-03-04

Licence

sd

Associated documents

(publications, reports, patents, experimental plan...), website

- Project website <https://www.obs4clim.fr/>

Management plans related to the project

- IAGOS Data Management Plan <https://github.com/iagos-dc/iagos-dmp>
- ACTRIS Data Management Plan <https://github.com/actris/data-management-plan/blob/master/DMP/ACTRIS-DMP.md>

ICOS improved Data Lifecycle <https://www.icos-cp.eu/sites/default/files/2021-02/D5.5.%20ICOS%20improved%20data%20lifecycle.pdf>

INFORMATION ON THE PROJECT

Project title Obs4Clim - Système d'Observation Intégré pour l'Atmosphère
Acronym Obs4Clim
Abstract OBS4CLIM is the joint effort for innovation of the three French components of the European Research Infrastructures (RI) in the atmospheric domain: the ESFRI Landmark ACTRIS, the ESFRI Landmark IAGOS, and the atmospheric component of ESFRI Landmark ICOS. OBS4CLIM develops a joint strategy for investment and service developments to reinforce the integration of "atmosphere RIs" and their synergies via e.g., innovation, joint utilisation of equipment and harmonisation of access conditions. OBS4CLIM is, thus, the EquipEx "atmosphere" project of the entire community concerned at a national level, and will enable the three French mirror RIs (ACTRIS-FRANCE, IAGOS-FRANCE, ICOS-FRANCE) to respond to the new challenges posed for Earth observation and provide their users with qualified and relevant data sets, as well as innovative services.

OBS4CLIM provides atmospheric RIs with adequate investment to keep serving the users at the highest level of quality over the next 15 years and to engage in developments to further respond to emerging needs, e.g. enhancing the networks in their four dimensions (longer and uninterrupted time-series, synergies with space-based observations, expanding global, denser network in specific areas, smart specialisations,...). The 8-year investment plan in OBS4CLIM has the following main objectives: 1) Fostering attractiveness of atmospheric facilities (i.e. central laboratories, reference instruments, fixed and mobile observation, and chamber facilities) to meet the evolving needs of research communities and the private sector, 2) Enhancing the capacity of atmospheric RIs to provide state of the art data services addressing the needs of stakeholders at national, European, and international levels, and 3) Expanding spatial and temporal coverage provided by atmospheric RIs to respond to future needs of a global atmospheric observing system, including the use of non-conventional, lower-cost and autonomous instruments, to support wider coverage, faster data and efficient service delivery to users.

OBS4CLIM will offer unique services in order to remain a place for innovation in research and technology. It is integrated into a mature framework for access, recognized at both national and international levels, that includes physical and remote access to atmospheric facilities as an integral part of the RI service portfolios. OBS4CLIM will further strengthen the capacity to translate the wealth of climate and atmospheric data into data services that will drive innovation and help decision-makers find ways to achieve a clean-air, climate-resilient, and low-carbon society.

Funding ● Agence Nationale de la Recherche : ANR-21-ESRE-0013

Start date 2021-09-15

End date 2029-06-14

Partners

- Centre National de Recherche Scientifique (CNRS)
- Aix-Marseille University (AMU)
- Sorbonne University (SU)
- Clermont Auvergne University (UCA)
- University of La Réunion (UR)
- University of Lille (LILLE)

- The Research Institute for Development (IRD)
- The Institut Mines-Télécom Northern Europe (IMT NE)
- Commissariat à l'énergie atomique et aux énergies alternatives (CEA)
- Toulouse III University - Paul-Sabatier (UT3)
- National Research Institute for Agriculture, Food and Environment (INRAE)
- Météo-France (MF)
- French National Institute for Industrial Environment and Risks (INERIS)
- University of Reims Champagne-Ardenne (URCA)
- Versailles Saint-Quentin-en-Yvelines University (UVSQ)
- Grenoble Alpes University (UGA)
- Paris-Est Créteil University (UPEC)
- École Polytechnique (EP)

Research products:

1. Default research output (Dataset)

Contributors

Contributors	Affiliation	Roles
Guillaume Brissebrat	AERIS/CNRS	DMP manager
Damien Boulanger	AERIS/CNRS	IAGOS data contact person
Léo Rivier	ICOS/CEA	ICOS data contact person
Jérôme Tarniewicz	ICOS/UVSQ	ICOS data contact person
Christophe Boitel	AERIS/CNRS-IPSL	ACTRIS data contact person
Paolo Laj	UGA/CNRS-IGE	Project coordinator
Valérie Thouret	UT3/CNRS-LAERO	Project co-coordinator, IAGOS-France director
Michel Ramonet	CNRS-LSCE	Project co-coordinator
Stéphane Sauvage	IMT NE	ACTRIS-FR director
Denis Loustau	INRAE	ICOS-France director
Sabine Philippin	CNRS-LAMP	Project manager

DATA MANAGEMENT PLAN OF THE PROJECT "OBS4CLIM - SYSTÈME D'OBSERVATION INTÉGRÉ POUR L'ATMOSPHÈRE"

1. DATA DESCRIPTION AND COLLECTION OR RE-USE OF EXISTING DATA

1A. HOW WILL NEW DATA BE COLLECTED OR PRODUCED AND/OR HOW WILL EXISTING DATA BE RE-USED?

OBS4CLIM is a joint initiative of the national research infrastructures (RI) IAGOS, ACTRIS and ICOS for which the operational data collection and processing is already in place within a close collaborative framework with the French atmospheric data centre AERIS. Part of the OBS4CLIM investments are allocated for upgrading of the existing instrumentation or their deployment on new instrumented ACTRIS, IAGOS or ICOS sites for which the data procedures and standards are already in place.

OBS4CLIM furthermore develops a number of services for new variables that are intended to be integrated with these 3 RIs. The development of these services will consider the entire value chain starting from the initial phase of preparations, the operations for the provision of data, as well as the provision of access to the data services to the user communities. The new variables/services will be included in the revised version of Data Management Plan as part of its next update.

Data levels

Data levels definition vary slightly across RIs but common patterns apply:

- L0: raw data
- L1: partly qualified data (incl. NRT data)
- L2: fully qualified data
- >L2: elaborated data product

Only fully qualified data and elaborated data products are made openly available to the public. For a more precise description see ACTRIS, ICOS and IAGOS DMPs.

1B. WHAT DATA (FOR EXAMPLE THE KIND, FORMATS, AND VOLUMES), WILL BE COLLECTED OR PRODUCED?

The primary goal of Obs4Clim is to produce high quality integrated datasets in the area of atmospheric sciences and provide services on those data. The purpose of the data collection and generation of data products in Obs4Clim is to provide open access to reactive and greenhouse gases, cloud and aerosols in measurements of high quality, benefiting a large community of scientists involved in atmospheric science and related areas as well as policy makers, private sector, educators and the general public.

ACTRIS data

ACTRIS data are data from observational or exploratory National Facilities complying with the procedures established within ACTRIS. ACTRIS data comprises ACTRIS variables resulting from measurements at National Facilities that fully comply with the standard operating procedures (SOP), measurement recommendations, and quality guidelines established within ACTRIS.

ACTRIS observational platforms are fixed ground-based stations that produce long-term data based on a regular measurement schedule and common operation standards. These platforms perform measurements of aerosol, clouds,

and reactive trace gases from the Earth surface throughout the troposphere up to the stratosphere by applying state-of-the-art remote-sensing and in situ measurement techniques under consideration of harmonised, standardised, and quality controlled instrumentation, operation procedures and data retrieval schemes.

ACTRIS exploratory platforms are atmospheric simulation chambers or mobile platforms. The chambers are among the most advanced tools for studying and quantifying atmospheric processes in controlled conditions and are used to provide many of the parameters incorporated in air quality and climate models. Atmospheric simulation chamber data contribute to better predicting the behaviour of the atmosphere and its processes over all time scales through a detailed understanding of the physical and chemical processes which affect air quality and climate change. Mobile platforms are set up to perform dedicated experiments and provide data on atmospheric constituents, processes and events.

ACTRIS-FR deploys in its perimeter about 200 different types of instruments that generate an equally high number of geophysical parameters. Standard data processing is applied to most of the data. ACTRIS-FR data and data products are provided by the atmospheric data centre AERIS in a harmonised format using about 15 *ACTRIS-compliant* processing procedures including data quality control, analysis and visualisation. It generates more than 500 data products, for a total of 15 TB and about 60 million files. The principle ACTRIS-FR instruments and parameters are categorised according to the scientific topics. A simplified list is provided in table 1 below.

Table 1. List of principle ACTRIS-FR instruments and parameters. The complete list is available online: <https://docs.google.com/spreadsheets/d/1wkP3sjqTCO1mWRxt58ZdZUS1gww13mKcHISViJDDu3w/edit?usp=sharing>

Theme	Parameters / Instruments	Processing by AERIS	Provision to external database
Aerosol remote sensing (GT1)	Aerosols Lidar	BASIC STRAT / STRAT + GARRLiC BASIC-Evolution	EARLINET NDACC E-PROFILE
	Photometer	GRASP-AOD	
Aerosol <i>in situ</i> (GT2)	ACSM, Aethalometer, SMPS, TEOM, Nephelometer, CPC, OPC/OPS	Calibration, Format, QL	
Cloud and precipitation (GT3)	Cloud radar		
	Precipitation radar	Format	
	OPC cloud		
Vertical profiles of thermodynamic and dynamic parameters and Ozone (GT4)	Imager	ELIFAN, mosaic, animation	
	GPS	Algo IGN	
	Lidar H2O	Filiere H2O OPAR	
	Lidar temperature and stratospheric Ozone	QL	NDACC NORS
	Lidar tropospheric Ozone	Format / QL	NDACC
	Radar UHF / VHF	DESMAN	E-PROFILE
	ozone sonde, Dobson Ozone, Sazoz Ozone/NO2	Format / QL	NDACC
	Micro-onde radiometer		MWRNET
	Radiosoundings	Format	GRUAN
Surface flux (GT5)	Turbulent fluxes	Eddypro	
	Radiation	BSRN	BSRN
	UV spectrometer	Format / QL	NDACC
	Surface meteorological stations	Format / QL	
Reactive trace gas <i>in situ</i> (GT6)	NOx, O3, COV, SO2	Format	
Atmospheric Deposition (GT7)	Deposition	Format	

IAGOS data

IAGOS (In-service Aircraft for a Global Observing System) aims to provide long-term, regular and spatially resolved in situ observations of the atmospheric composition. IAGOS observational platforms are commercial aircraft hosting IAGOS sensors that produce data based on a regular measurement during all the flights performed by the aircraft. These platforms perform measurements of reactive and greenhouse gases, cloud particles, aerosols and also meteorological variables by applying state-of-the-art in situ measurement techniques under consideration of harmonised, standardised, and quality controlled instrumentation, operation procedures and data retrieval schemes. The fleet is currently composed of 8 commercial aircraft.

The data measured by IAGOS sensors are labelled as IAGOS-CORE. The data centre also manages the data acquired during the former project MOZAIC with similar sensors (labelled IAGOS-MOZAIC) and by the current project CARIBIC with similar sensors (labelled IAGOS-CARIBIC). The observed variables by IAGOS instruments are listed in the table 2 below.

Table 2: List of variables covered by IAGOS.

	IAGOS-CORE & IAGOS-MOZAIC	IAGOS-CARIBIC
Greenhouse Gases		
CO ₂	Package2d (since 2021)	since 2005
CH ₄	Package2d (since 2021)	since 2005
H ₂ O	ICH (since 1994)	since 1997
Reactive Gases		
O ₃	Package1 (since 1994)	since 1997
CO	Package1 (since 2001)	since 1997
NO _x	Package2b (since 2015)	X
Aerosols		
Number Density	Package2c (Certification in progress)	
Size Distribution	Package2c (Certification in progress)	
5-75 μ	BCP (since 2011)	
Cloud Particles (5-75 μm)		
Number Density	BCP (since 2011)	
Size Distribution	BCP (since 2011)	

In addition to the observed variables are provided systematic L4 ancillary parameters based on ECMWF data and on FLEXPART back-trajectories along each flight measurements. These ancillary parameters are available from 1994 to present day. The following ECMWF data are interpolated along flight trajectories using 6 hours ECMWF analysis and 3 hours ECMWF forecasts. Some parameters are derived from systematic FLEXPART simulations along each flight measurements:

- Air potential temperature

- Geopotential height at 500 hPa
- Orography
- Pressure at four different PV values (1.5, 2, 3 and 4)
- Surface Pressure
- Vertical wind velocity
- Boundary layer height
- Potential vorticity

Observational data L2 provides all observations merged in a single file for each IAGOS flights. L4 ancillary parameters are provided in another file. The database includes more than 63000 flights since 1994. Table 3 presents the current volume of the data and the estimation for the future years. The volume depends on the number of aircraft instrumented and the number of flights operated.

Table 3. Overview of past and future predicted data volumes from IAGOS flights.

	aircraft	flights	Total flights	Volume GB
2020	7	3235	62193	133.20
2021	7	3235	65428	140.13
2022	8	3697	69125	148.04
2023	9	4159	73284	156.95
2024	10	4621	77905	166.85
2025	11	5083	82988	177.74
2026	12	5545	88534	189.61
2027	13	6007	94541	202.48
2028	14	6470	101011	216.33
2029	15	6932	107942	231.18
2030	16	7394	115336	247.01

ICOS data

The Integrated Carbon Observation System (ICOS) provides long term, high quality observations that follow (and cooperatively set) the global standards for the best possible quality data on the atmospheric composition for greenhouse gases (GHG), greenhouse gas exchange fluxes measured by eddy covariance and CO₂ partial pressure at water surfaces. Within each domain (Atmosphere, Ecosystem, Ocean), a Thematic Centre coordinates the observations and supports the stations. In addition to the Thematic Centres, there are Central Analytical Laboratories (CALs) that provide gas analyses and calibration gases. ICOS has established a network of tall towers, coastal and mountain stations where data on greenhouse gas concentrations in the atmosphere are collected. The ICOS Atmosphere network covers a large part of the European continent.

Table 4. List of parameters measured at ICOS Atmosphere stations.

CATEGORY	GASES, CONTINUOUS SAMPLING	GASES, PERIODICAL SAMPLING	METEOROLOGY, CONTINUOUS	EDDY FLUXES
Class 1 Mandatory parameters	CO ₂ , CH ₄ , CO: at each sampling height	CO ₂ , CH ₄ , N ₂ O, SF ₆ , CO, H ₂ , ¹³ C and ¹⁸ O in CO ₂ ; sampled every three days at highest sampling height ¹⁴ C (radiocarbon integrated samples): at highest sampling height	Air temperature, relative humidity, wind direction, wind speed: at highest and lowest sampling height* Atmospheric Pressure Planetary Boundary Layer Height**	
Class 2 Mandatory parameters	CO ₂ , CH ₄ : at each sampling height		Air temperature, relative humidity, wind direction, wind speed: at highest and lowest sampling height* Atmospheric Pressure	
Recommended parameters***	²²² Rn, N ₂ O, O ₂ /N ₂ ratio CO for Class 2 stations	CH ₄ stable isotopes, O ₂ /N ₂ ratio for class 1 stations: weekly sampled at highest sampling height		CO ₂ : at one sampling height

Each ICOS Atmosphere station is an observatory established to continuously measure the dynamics in the concentration of greenhouse gases (CO₂, CH₄) and other trace gases (for example, CO), which are the result of regional and global surface fluxes as well as of complex atmospheric transport mechanisms. A site chosen for installing an Atmosphere station will typically be representative of a footprint area of more than 10 000 km². The ICOS Atmosphere stations are equipped with commercially available instruments integrated into a digital control system run by ICOS ATC custom-made software.

2. DOCUMENTATION AND DATA QUALITY

2A. WHAT METADATA AND DOCUMENTATION (FOR EXAMPLE THE METHODOLOGY OF DATA COLLECTION AND WAY OF ORGANISING DATA) WILL ACCOMPANY THE DATA?

In order to increase their reusability, data are completed with rich metadata which are in open access from the website. Data will be organised by AERIS according to the desires of the OBS4CLIM science team. Each dataset (collection of data stored together) will be connected to a metadata document according to AERIS requirements. The connection will be accomplished through standardised naming conventions (i.e. metadata file names and data file names will be similar). Data will be clearly described, including processing procedures and methods for dealing with instrument problems and calibrations, in the metadata file linked to the instrument.

The specified AERIS format for metadata is compatible with ISO 19115, INSPIRE and other standards and extensible to adapt to the specific features of certain types of data (e.g. aircraft and ground platforms, remote sensing, and output of 0D and 3D models). They can be exported to standard formats or other portals on request. Each dataset will also be assigned a Digital Object Identifier (DOI) by AERIS.

Two naming conventions are used in the metadata profiles and the data files, to represent the variables provided in the datasets, the instruments and the platforms:

- GCMD (Global Change Master Directory)
- CF convention (Climate and Forecast)

A system is currently implemented to provide the full provenance metadata for each dataset. This information will be available on the Data Portal in a human-readable way and in the data files in a machine-readable way. The system is based on the implementation W3C standard PROV-O and will allow to provide full history for all the datasets including:

link to the previous versions of the datasets, identifier of the input datasets used to create this dataset, identification of the software, data providers and instruments, etc. A SPARQL endpoint will provide access to the provenance metadata.

2B. WHAT DATA QUALITY CONTROL MEASURES WILL BE USED?

ACTRIS

Quality control procedures have been defined and are applied at each step of the data processing chain, from the measurement to the provision of the data.. The standard operation procedures (SOPs) have been developed by ACTRIS & SNO working groups involving the instrument PIs. These SOPs are applied at the measurement platforms on-site to provide documented and traceable data and data products

Beside the data validation by the instrument PIs, specific automatic processing tools are applied to the different ACTRIS-FR data by the AERIS DC in order to ensure compliance between recorded information in the database, the standardised format or higher geophysical data products (>L0).

A detailed description of these procedures is available in the [ACTRIS DMP](#).

IAGOS

The quality assurance (QA) of the measurements acquired by the instruments and provided to the user is performed by following the Standard Operating Procedures (SOPs). The SOPs documentation are available on the IAGOS website for each instrument (<https://www.iagos.org/iagos-core-instruments/>). The quality control (QC) of the data is performed by each instrument PI using automatic or semi-automatic data processing protocols. All observations are provided with uncertainties and quality flags. Data processing operations are traceable within the database. The internal consistency of the data is checked for every new flight operation. The external consistency is also checked by comparing the IAGOS data with other measurement platforms such as ground stations, balloons, satellite and research aircraft field campaigns.

The IAGOS - Quality Assessment Group (I-QAG) composed of the IAGOS instrument PIs or designated QA/QC managers, meets twice a year to ensure the application of the IAGOS-QA/QC concept into routine operation (see Figure below). The chair of the I-QAG will be nominated by the IAGOS-AISBL. The I-QAG prepares yearly the internal QA/QC evaluation reports and reviews them. The I-QAG reports regularly (every year) or on demand to the IAGOS Executive Board. On invitation, an external panel of experts review regularly (every 3-5 years) the QA/QC documentation prepared by the I-QAG. The IAGOS QA/QC process is illustrated in figure 1 below.

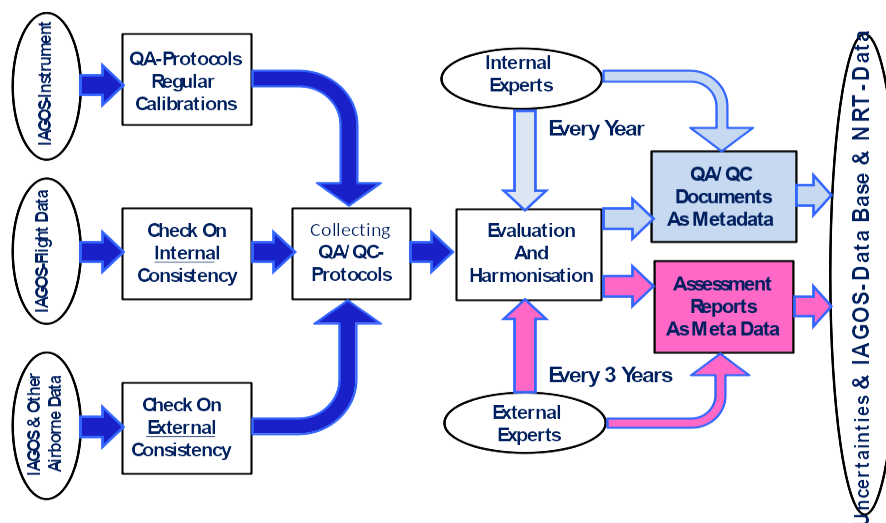


Figure 1: Illustration of IAGOS data QA/QC process.

ICOS

Quality assurance is at the heart of ICOS and the reason for the existence of the research infrastructure. The quality assurance procedures are described in the relevant papers and reports that are published by ICOS, the Thematic Centers and their contributors. The Central Facilities in ICOS ensure that all data are treated and quality controlled with the same algorithms and are properly archived. Thematic centres receive online data from the ICOS stations, typically on a daily basis and in near-real-time, and perform automatic quality checks. They interact with the Monitoring Station Assemblies, which consist of the stations' Principal Investigators, researchers and technicians, on matters related to the coordination and improvement of the ICOS National Networks, to produce ICOS Final Fully Quality Controlled Observational Data (Level 2). Harmonisation in procedures and equipment with the appropriate Quality Assurance/Quality Control plan ensures the robustness and reliability of the collected atmospheric dataset. An in-house software was developed and is maintained to centrally process and quality control the data from ICOS Atmosphere stations.

The Atmospheric Thematic Centre (ATC) is composed of a data centre and a metrology lab and is complemented by an in situ station quality-control mobile lab. The ATC is coordinated and operated by the French Laboratoire des Sciences du Climat et de l'Environnement, supported by the Finnish Meteorological Institute. The ATC has the following long-term objectives:

- Develop and operate the atmospheric data processing chains, from the data transmission from stations to the routine delivery of the quality-checked data-stream used by modellers.
- Carry out a regular measurement-technology
- survey and analysis and enable the development of new sensors and their testing, for instance, through research and development programmes.
- Service the network with spare instruments, training and high-level technical assistance.

Link the ICOS atmospheric data-collection programme with other ICOS Central Facilities within the framework of European and international monitoring programmes.

For ATC ICOS station specifications see: https://meta.icos-cp.eu/objects/JUOeklwSb3fh8hdK9eL3_V9V

ATC Data processing: Hazan, L., Tarniewicz, J., Ramonet, M., Laurent, O., and Abbaris, A.: Automatic processing of atmospheric CO₂ and CH₄ mole fractions at the ICOS Atmosphere Thematic Centre, Atmos. Meas. Tech., 9, 4719-4736, doi:10.5194/amt-9-4719-2016, 2016

3. STORAGE AND BACKUP DURING THE RESEARCH PROCESS

3A. HOW WILL DATA AND METADATA BE STORED AND BACKED UP DURING THE RESEARCH?

All RI data centres have a long-term commitment to dedicate the required resources and procedures to ensure sustainable databases, incl. OBS4CLIM data.

ACTRIS

The archiving of the database and data files is assured through redundancy of the archive in a dedicated storage space, and is backed-up on a daily basis on external support devices. All platforms have furthermore set up equivalent solutions based on their own resources or informatic services with which their facility is linked.

IAGOS

In the case that the data would be lost, a sustainable archiving is provided by a dedicated service at OMP. The data is duplicated on different geographical sites in Toulouse, France.

ICOS

A task of the ICOS Carbon Portal (<https://www.icos-cp.eu/>) is to organise the long-term archiving of ICOS data products, with the aim of guaranteeing both safe storage and future access long after the cessation of the RI itself. This activity complements the data storage and backups routinely performed by the Thematic Centres.

3B. HOW WILL DATA SECURITY AND PROTECTION OF SENSITIVE DATA BE TAKEN CARE DURING THE RESEARCH

No sensitive data will be collected during OBS4CLIM. The data are from scientific measurements of atmospheric composition and other related quantities.

In the event of data loss, the intent is to make use of the redundant copies of the data at various locations.

4. LEGAL AND ETHICAL REQUIREMENTS, CODE OF CONDUCT

4A. IF PERSONAL DATA ARE PROCESSED, HOW WILL COMPLIANCE WITH LEGISLATION ON PERSONAL DATA AND ON SECURITY BE ENSURED?

The Data Centre stores personal data about the users as they need to register. Personal data management follows GDPR. It will be the only personal data collected during OBS4CLIM. The data are from scientific measurements of atmospheric composition and other related quantities.

4B. HOW WILL OTHER LEGAL ISSUES, SUCH AS INTELLECTUAL PROPERTY RIGHTS AND OWNERSHIP, BE MANAGED? WHAT LEGISLATION IS APPLICABLE?

All data collected during OBS4CLIM are the property of the measurement teams and/or their institutions. These teams agree to share their data with the rest of the science team, and when the embargo period is over, with the wider scientific community and the public. Since these data were collected during a project supported by public funds, the data are ultimately owned by the public.

All data from the project will be distributed under a CC-BY 4.0 licence. This licence allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. It allows commercial use.

4C. WHAT ETHICAL ISSUES AND CODES OF CONDUCT ARE THERE, AND HOW WILL THEY BE TAKEN INTO ACCOUNT?

The Data Centre does not manage any data with disclosure risk.

The data providers are well aware of their responsibility for the correctness of data and metadata. In general, everyone in OBS4CLIM should work in a socially ethical way keeping the integrity and FAIRness, and maintaining a high level of trust and respect among the people working in the project and with the users and other stakeholders. One should always take into account that the mission of OBS4CLIM is to provide effective access for a wide user community to its resources and services, in order to facilitate high-quality Earth system research, to increase the excellence in Earth system research, and to provide information and knowledge on developing sustainable solutions to societal challenges.

5. DATA SHARING AND LONG-TERM PRESERVATION

5A. HOW AND WHEN WILL DATA BE SHARED? ARE THERE POSSIBLE RESTRICTIONS TO DATA SHARING OR EMBARGO REASONS?

The OBS4CLIM website: <https://www.obs4clim.fr/> will include information about the data produced during the project. The data will be accessible through the RIs data centres and the AERIS catalogue: <https://www.aeris-data.fr/catalogue>.

OBS4CLIM data is licensed under the Creative Commons Attribution 4.0 International licence (CC BY 4.0). OBS4CLIM data is public data open to all data users. Users are free to copy and redistribute the material in any medium or format, remix, transform, and build upon the material for any purpose, even commercially. Users must give appropriate credit, provide a link to the licence, and indicate if changes were made. Users may do so in any reasonable manner, but not in any way that suggests the licensor endorses himself or its use. Users may not apply legal terms or technological measures that legally restrict others from doing anything the licence permits. Use of the data requires proper reference and citation of the OBS4CLIM data, using the exact citation (including the provided DOI or PID) as provided in the metadata.

Only fully qualified data and elaborated data products are made openly available to the public. Raw data are only available for the members of the project as their quality is limited. Some NRT datasets are only provided to Copernicus on the basis of a contract.

IAGOS data is open but the users need to be registered to get access. Registration allows the IAGOS DC to provide rich data usage information and help to provide added-value services to the users such as subscription to data sets, queries history, etc. The authentication and authorization system is based on the Single Sign-On solution hosted and maintained by AERIS. It allows the users to authenticate using their ORCID or EDUGAIN accounts.

5B. HOW WILL DATA FOR PRESERVATION BE SELECTED, AND WHERE DATA WILL BE PRESERVED LONG-TERM (FOR EXAMPLE A DATA REPOSITORY OR ARCHIVE)?

The OBS4CLIM final data will be the most important data for long-term archival. The archive will be distributed over three of the four Data and Services Centre (DSC) of AERIS: ESPRI, ICARE and SEDOO. These centres are already hosting data from ACTRIS, IAGOS and ICOS. They have experts in computer hardware and software. Their expertise will be employed to ensure data is backed up regularly and safely (see section 3a).

The mid-term preservation (10 years) is guaranteed by AERIS (see section 6a). The question of long-term preservation will be addressed soon in the framework of Data Terra RI. In case any of the DSC finds it cannot maintain this database, the responsibility could be transferred to another data centre within AERIS. AERIS is strongly supported at national level (CNRS, CNES, Météo-France...). In the highly unlikely event that AERIS will have to close operations, we guarantee that we will migrate all content to other suitable repositories, and since all datasets will have DOIs (see 5d), all citations and links to datasets will not be affected.

Other data will be primarily archived by the individual measurement groups. AERIS can provide a backup if necessary.

5C. WHAT METHODS OR SOFTWARE TOOLS ARE NEEDED TO ACCESS AND USE DATA?

Data will be accessible and downloadable from a dedicated website created and managed by AERIS. Access to the data will not require intervention or handling of requests. Once the website is accessed, the data can be immediately downloaded. The data will be stored in formats that are widely used by the user's community.

Except data from simulation chambers, all data will be converted to the NetCDF format (using CF and ACDD conventions). Data from simulation chamber data will be distributed using the EUROCHAMP Data Format (EDF).

As it is sometimes more practical for users, an ascii format is also provided (CSV or NASA Ames).

5D. HOW WILL THE APPLICATION OF A UNIQUE AND PERSISTENT IDENTIFIER (SUCH AS A DIGITAL OBJECT IDENTIFIER (DOI)) TO EACH DATA SET BE ENSURED?

One of the services of AERIS (a trustworthy, long-term data repository and manager; see description in section 6a) is to provide a DOI for each dataset. These will be requested for each component of the OBS4CLIM database and will be stored with the metadata viewable on the website.

AERIS has an agreement to assign DOIs to datasets with the prefix 10.25326. A DOI will be requested for fully qualified data and elaborated products. It will be registered on Datacite (<https://datacite.org/>) using its metadata scheme. The landing page, provided by AERIS, will contain a sample citation, full metadata and information on how to access the data.

For the internal datasets (raw data and non-fully qualified data), non-persistent homemade identifiers are generated. A new system based on ePIC handles is currently under implementation. AERIS already has an agreement to assign such PIDs with the prefix 21.11148. An ePIC PID will also be assigned to all instruments, using the PIDInst metadata schema designed by RDA (Research Data Alliance).

6. DATA MANAGEMENT RESPONSIBILITIES AND RESOURCES

6A. WHO (FOR EXAMPLE ROLE, POSITION, AND INSTITUTION) WILL BE RESPONSIBLE FOR DATA MANAGEMENT (I.E. THE DATA STEWARD)?

The cluster for Atmospheric data AERIS (<https://www.aeris-data.fr>), part of the French Data Terra Research Infrastructure (RI), has the objective to facilitate and enhance the use of atmospheric data, whether from satellite, aircraft, balloon, or ground observations, or from laboratory experiments. It generates advanced products and provides services to facilitate data use, to prepare campaigns, and to interface with modelling activities. It consists of four Data and Service Centres (DSC) with more than 20 years' experience and strong expertise in data curation, storage, preservation, and dissemination by these entities: ICARE, ESPRI, SATMOS and SEDOO. Most of these data centres are involved in European initiatives and projects promoting the FAIR data principles and participating in the European Open Science Cloud (EOSC). AERIS is hosting the IAGOS data centre and parts of the ACTRIS data centre (trace gas remote sensing, atmospheric simulation chamber studies, and other types of data). AERIS is also in charge of the distribution of the data produced by the french stations of ICOS.

The AERIS organisation will be ultimately responsible for the many data management/stewardship activities involved with the OBS4CLIM data.

This DMP will be regularly reviewed and updated to ensure its content reflects the intent and needs of the OBS4CLIM science team and the public interested in OBS4CLIM scientific issues.

6B. WHAT RESOURCES (FOR EXAMPLE FINANCIAL AND TIME) WILL BE DEDICATED TO DATA MANAGEMENT AND ENSURING THAT DATA WILL BE FAIR (FINDABLE, ACCESSIBLE, INTEROPERABLE, RE-USABLE)?

AERIS will ensure that the OBS4CLIM data will be FAIR. This is a component of their agreement to manage the OBS4CLIM data. The data are distributed by data centres, and so, by concept, the infrastructure and maintenance costs are mutualized.

AERIS is still in charge of the data management and "FAIRisation") for the 3 research infrastructure involved in OBS4CLIM. A part of the resources needed for OBS4CLIM is then still planned in each infrastructure. In addition, OBS4CLIM will benefit from GAIA DATA EQUIPEX+ project activities especially for new observations based on fast-developing innovative and non-conventional technologies, including low-cost sensors as well as embedded and automatic sensors. GAIA DATA plan to implement an integrated distributed data and services platform, develop services accessible via portals allowing research and processing between and transdisciplinary from multi-source data acquired by satellites, ships, planes, drones, submersibles, balloons, in situ devices, inventories, observatories and

experimentation, as well as, on data from reference simulations. OBS4CLIM has been identified as a use case to test this implementation.

Another source of funding is the involvement of AERIS (and the OBS4CLIM RI) in European projects such as ENVRI-FAIR and EOSC related projects.