Data
Models
Inventories

PARIS

Process Attribution of Regional Emissions

GA 101081430, RIA

Data Management Plan

Deliverable D1.2			
Delivery due date Annex I	PM 6 | 30 June 2023		
Actual date of submission	12 July 2023		
Lead beneficiary: UU	Work package: 1	Nature: Report	Dissemination level: PU
Responsible scientist	Remco de Kok(WU)		
Contributors	Wouter Peters (WU)		
Internal reviewers	Sylvia Walter (UU)		
Version: 1			

Horizon Europe Cluster 5: Climate, energy and mobility

"This project has received funding from the European Union's Horizon Europe Research and Innovation programme under HORIZON-CL5-2022-D1-02 Grant Agreement No 101081430 - PARIS".

Table of content

Abbreviations

AGAGE	Advanced Global Atmospheric Gases Experiment
CF	Climate and Forecast
CMIP	Coupled Model Intercomparison Project
DARE-UK	Detection and Attribution of Regional Emissions in the UK
ESFRI	European Strategy Forum on Research Infrastructures
GHG	Greenhouse gas
ICOS	Integrated Carbon Observation System
NCAS	National Centre for Atmospheric Science (UK)
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration

1. Changes with respect to the DoA (Description of the Action)

The deliverable D1.2 belongs to task T1.4 of work package WP1 and has been executed as planned in the DoA.

2. Dissemination and uptake

The deliverable is a public document and will be made available on the project website and the ICOS repository.

3. Short Summary of results

This deliverable is the initial draft of the Data Management Plan (DM Plan) of the HorizonEurope project *PARIS – Process Attribution of Regional emISsions*.

PARIS as a HorizonEurope research project aims to significantly increase our knowledge about greenhouse gas emissions, the evaluation & combination of scientific approaches and a progressive use of collaborative data. 17 European partners focus on emissions of common greenhouse gases such as carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O), and new emissions estimates for fluorinated gases (F-gases). With expected large data sets from measurements and modelling it is essential to implement a tailored data management strategy right from the beginning of the project.

The DM Plan implements the procedures for this and ensures FAIR access to and re-use of research data generated by PARIS as a HorizonEurope project. The DM Plan will follow the **FAIR** principles to ensure data are **F**indable, **A**ccessible, **I**nteroperable and **R**e-usable.

As the DM Plan is a key element of project management, it describes preparation, handling, preservation, and storage of data collected and generated by PARIS. The DM Plan is a living document and will be updated or adjusted over the course of the project whenever significant changes arise. The DM Plan is a public document and an approved copy – by the EU and the consortium - will be available on the project website to ensure FAIR data.

4. Evidence of accomplishment

4.1 Introduction | Background of the deliverable

To ensure knowledge integration and reuse of data, four principles - referring to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure - are necessary:

1. Findability
2. Accessibility
3. Interoperability
4. Reusability

Findability is the first step in (re)using data. It includes both humans and machines, whereas particularly the latter is essential for an automatic discovery of (large) data sets. Key for findability are adequate identifiers of data, detailed descriptions of data by clear and explicit metadata and searchable resources/repositories of data.

In a next step, users must be able to access the data, e.g., by using a standardised communications protocol including authentication and authorisation. The protocol must be universally implementable and ensuring free and open access to the (meta)data. These (meta)data must be interoperable, which means that they can be integrated with other data, applications and workflows for analysis, storage, and processing. In particular, (meta)data should use broadly applicable languages and vocabularies and should also include qualified references. In combination with detailed prevenance, domain-relevant community standards and clear data usage licences, well described (meta)date are optimised to be reusable.

4.2 Scope of the deliverable

The purpose of this document is to draft the criteria to data management within the PARIS project. The DM Plan is a living document and will be updated or adjusted over the course of the project whenever significant changes arise. The DM Plan is a public document and an approved copy – by the EU and the consortium - will be available on the project website to ensure FAIR data.

4.3 Content of the deliverable

4.3.1 Data Summary

The project will use and generate data of different levels, from raw data that is the direct output of measurement devices, to calibrated, quality-controlled data and model-derived data products. The latter two are used for the deliverables in this project. This document deals with the data products that are directly related to the deliverables. Management of the raw data is assumed to be the responsibility of the relevant PIs and such data are available upon request only.

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

The main aim of the PARIS project is to bring the estimates of climate forcing greenhouse gases (GHG) and aerosols from bottom-up approaches (based on inventories) and top-down (based on atmospheric measurements) closer together. This involves performing improved top-down estimates by existing and new atmospheric measurements, as well as comparing these with bottom-up inventories (national reporting and other regional/global products) and process-models.

Will you re-use any existing data and what will you re-use it for?

Work packages WP3-7 all include top-down estimates of emissions. Measurement networks are already in place for several GHGs and particulates. We will use existing networks where possible, and our project added new measurements where significant improvements over existing infrastructure can be expected. A summary of all data used, existing and new, is given in Table 1. Atmospheric transport models used in work packages WP3-7 rely on operational meteorological input data (ECMWF, UKMO) that is available to the project partners.

Table 1: Data summary

Data description	Data type	Data volume	Data origin
WP1: Project Management and coordination			
Lead Beneficiary: UU | Contact: s.walter@uu.nl			
General documentation and reporting	.docx, .pptx, .xlsx, .pdf	< 1 GB	Created by WP1
WP2: National top-down, bottom-up reconciliation			
Lead Beneficiary: MO | Contact: alistair.manning@metoffice.gov.uk			
Annex reports to the National Inventories	Docx, PDF	< 1 GB	Created by WP2
WP3: European F-gas emissions evaluation			
Lead Beneficiary: GUF | Contact: an.engel@iau.uni-frankfurt.de			
F-gas flask and continuous measurements	csv, netCDF	< 1 GB	New measurements at individual sites.
F-gas ObsPack for inversions	netCDF	1 GB	Created from new and existing (AGAGE database) measurements
F-gas inversion prior and posterior flux maps and/or country aggregates	netCDF	< 1 GB	Created by WP3
Bottom-up F-gas flux estimates of point sources	netCDF	< 1 GB	Created by ORB
WP4: Sector-level CH_4 emission quantification			
Lead Beneficiary: UU | Contact: Tim.Arnold@ed.ac.uk			
Per-station contributions from OSSE	netCDF		Created by WP4
CH_4 isotope ratio measurements of sampling campaigns	csv	< 1 GB	New measurements
CH_4 isotope ratio measurements of 6 month continuous series	csv	< 1 GB	New measurements
CH_4 ObsPack for inversions	netCDF	1 GB	Created from new and existing (ICOS, other networks, other stations) measurements.
CH_4 inversion prior and posterior flux maps and/or country aggregates	netCDF	< 1 GB	Created by WP4
Per-sector inversion results for local campaign	netCDF	< 1GB	Created by WP4
WP5: Advancing N_2O process models and their reconciliation with top-down estimates			
Lead Beneficiary: EMPA | Contact: stephan.henne@empa.ch			
N_2O atmospheric measurements	csv, netCDF	< 1 GB	New measurements
N_2O ObsPack for inversions	netCDF	1 GB	Created from new and existing (ICOS, other networks) measurements.
N_2O isotope flux measurements	csv	<1 GB	New measurements
N_2O inversion prior (process models) and posterior flux maps and/or country aggregates	netCDF	< 1 GB	Model products created by WP5
WP6: Separation of biosphere and fossil fuel CO_2 flux			
Lead Beneficiary: UNIVBRIS | Contact: a.ganesan@bristol.ac.uk			
O_2 and CO_2 measurements	csv, netCDF	< 1 GB	New measurements
CO_2 and O_2 ObsPack for inversions	netCDF	2 GB	Created from new and existing (ICOS) measurements.

Fossil fuel emission inversion prior and posterior flux maps and/or country aggregates	netCDF	< 1 GB	Created by WP6
Verification game protocol	pdf	< 1 MB	Created by WP6
Simulated observations for inversions verification game, different scenarios	netCDF	6 GB	Created by WP6
Fossil fuel emission prior and posterior flux maps and/or country aggregates for inversion game	netCDF	< 1 GB	Created by WP6

WP7: Source attribution for European aerosols

Lead Beneficiary: NUIG | Contact: jurgita.ovadnevaite@nuigalway.ie

Aerosol measurements for source appointment	csv, netCDF	< 1 GB	From existing EBAS database and new data created by WP7
WRF model results	netCDF	100 GB	Created by WP7
Black Carbon ObsPack for inversions	netCDF	1 GB	Created from existing measurements
Black Carbon emission inversion prior and posterior flux maps and/or country aggregates	netCDF	< 1 GB	Created by WP7
FLEXPART model results	netCDF	50 GB	Created by WP7

WP2-7

Refereed papers, conference presentations, any other output	.docx, .pptx, .pdf	< 1 GB	Created by work packages

Table 2: Measurement summary

Data description	Measured parameter	Instrument used	Method used	Measured by (responsible person / institute)
WP3: European F-gas emissions evaluation				
Hegyhatsal	F-gases	Laboratory GC/MS	GC/MS	A. Engel, GUF
Birkenes	F-gases	Laboratory GC/MS	GC/MS	C. Lunder, NILU
Southern Italy	F-gases	ADS GC/MS	GC/MS	J. Arduini, UNIURB
Cabauw	F-gases	Medusa/samples	GC/MS	S. O'Doherty, UNIVBRIS
Taunus	F-gases	Medusa	GC/MS	A. Engel, GUF
Monte Cimone	F-gases	Medusa	GC/MS	J. Arduini, UNIURB
Zeppelin	F-gases	Medusa	GC/MS	C. Lunder, NILU
Mace Head	F-gases	Medusa	GC/MS	S. O'Doherty, UNIVBRIS
Tacolneston	F-gases	Medusa	GC/MS	K. Stanley, UNIVBRIS
WP4: Sector-level methane emission quantification				
SOAR or LUT	$\delta^{13}C$-CH_4 and ethane	Aerodyne laser spectrometer	Laser spectrometry	A. Ganesan, UNIVBRIS
Heathfield	$\delta^{13}C$-CH_4, $\delta^{2}H$-CH_4	Aerodyne laser spectrometer	Laser spectrometry	T. Arnold, NPL
Zurich	$\delta^{13}C$-CH_4, $\delta^{2}H$-CH_4	Aerodyne laser spectrometer	Laser spectrometry	J. Mohn, Empa
Mobile platform	$\delta^{13}C$-CH_4, $\delta^{2}H$-CH_4	IRMS	Mass spectrometry	T. Röckmann, UU
Mace Head	$\delta^{2}H$-CH_4	IRMS	Mass spectrometry	T. Röckmann, UU
WP5: Advancing N2O process models and their reconciliation with top-down estimates				
Hegyhatsal	N_2O	Los Gatos Research	CRDS	L. Haszpra, ATOMKI
Reckenholz	N_2O, $\delta^{15}N_a$, $\delta^{15}N_b$, $\delta^{18}O$	Picarro Laser spectrometer	CRDS	J. Mohn, Empa
WP6: Separation of biosphere and fossil fuel carbon dioxide flux				
Rotterdam, APO	CO_2, O_2	CO_2 (Uras), O_2 (Oxzilla)	Continuous in situ measurements	RUG, WU
Heathfield	CO_2, O_2	CO_2 (Ultramat), O_2 (Oxzilla)	Continuous in situ measurements	UEA
Weybourne	CO_2, O_2	CO_2 (Ultramat), O_2 (Oxzilla)	Continuous in situ measurements	UEA
WP7: Source attribution for European aerosols				
Dublin	Aerosol mass concentration and chemical speciation	Aerosol Chemical Speciation Monitor (Aerodyne Inc.)	Aerosol Mass Spectrometry	L. Lei / J. Ovadnevaite, NUIG

Table 3: Model data summary

Data description	Modeled parameter	Model used	Input data used	Modelled by (responsible person / institute)
WP3: European F-gas emissions evaluation				
F-gas inversion prior and posterior flux maps and/or country aggregates	F-gases	NAME, FLEXPART, ICON	ObsPacks	A. Manning, MO | S. Henne, Empa | A. Kaiser-Weiss, DWD
Bottom-up F-gas flux estimates of point sources	Emission of F-gases	AnaFgas (https://web.jrc.ec.europa.eu/policy-model-inventory/explore/models/model-anaf-gas), New modelling	Use of F-gases from EU F-gas reporting, Expert information	ORB
WP4: Sector-level methane emission quantification				
CH_4 inversion prior and posterior flux maps and/or country aggregates	CH_4 emissions	InTEM, Bristol-MCMC, Empa-Bayesian	Footprints from transport models (NAME, FLEXPART), concentration observations, isotope ratio observations, prior emission estimates by sector, isotope ratio source signatures	A. Ganesan, UNIVBRIS | A. Manning, MO | S. Henne, Empa
WP5: Advancing N2O process models and their reconciliation with top-down estimates				
Process model derived N_2O flux, also input to inverse model	N_2O flux	DayCent, LandscapeDNDC	Agricultural management, soll data, land use, climate data	S. Kehl, AG | B. Wolf, KIT
N_2O posterior flux maps from inverse model and country aggregates	N_2O flux	InTEM, Bristol-MCMC, Empa-Bayesian	Footprints from transport models (NAME, FLEXPART), concentration observations	A. Ganeson, UNIVBRIS | A. Manning, MO | S. Henne, Empa
WP6: Separation of biosphere and fossil fuel carbon dioxide flux				
Fossil fuel emission inversion prior and posterior flux maps and/or country aggregates	CO_2 flux	NAME, CarbonTracker	EDGAR, UK National Atmospheric Emissions Inventory, national inventories, oceanic and terrestrial CO2 models (TBD)	UNIVBRIS, WU
Simulated observations for inversions verification game, different scenarios	CO_2 flux	NAME, CarbonTracker	EDGAR, UK National Atmospheric Emissions Inventory, national inventories, oceanic and terrestrial CO2 models (TBD)	UNIVBRIS, WU
Fossil fuel emission prior and posterior flux maps and/or country aggregates for inversion game	CO_2 flux	NAME, CarbonTracker	EDGAR, UK National Atmospheric Emissions Inventory, national inventories, oceanic and terrestrial CO2 models (TBD)	UNIVBRIS, WU
WP7: Source attribution for European aerosols				
Black Carbon emission inversion prior and posterior flux maps and/or country aggregates	Black Carbon Flux	FLEXPART + BRISTOL-MCMC/FLEXINVERT	Footprints from transport model (FLEXPART), concentration observations (ACTRIS)	S. Annadate, UNIURB | J. Arduini, UNIURB | M. Maione, UNIURB | A. Ganesan, UNIVBRIS

State the reasons if re-use of any existing data has been considered but discarded.

N/A

What types and formats of data will the project generate or re-use?

See Table 1.

What is the expected size of the data that you intend to generate or re-use?

See Table 1.

What is the origin/provenance of the data, either generated or re-used?

See Table 1.

To whom might your data be useful ('data utility'), outside your project?

The new measurements and top-down estimates will be useful for all parties that are interested in monitoring greenhouse gases and aerosols concentrations and fluxes in Europe. These include governments and the public. Measurements can be used by other scientific groups to perform their own top-down estimates or perform other analysis to better understand Europe's emissions. Specifically, our new observational data can be useful to our two "sister" project EYE-CLIMA, and AVENGERS. Inferred flux data will be compared against officially reported values on the national scale and can be compared to results from other model studies.

4.3.2 FAIR data

4.3.2.1 Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

Yes, all publicly available data will have a unique persistent identifier, assigned by the ICOS Carbon Portal. ICOS has chosen to primarily work with PIDs built on the Handle system, including PIDs from the European Persistent Identifier Consortium (ePIC) and DOIs (Digital Object Identifiers) from DataCite.

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Metadata on the ICOS Carbon Portal will comply to the relevant community standards, netcdf data provided by the project partners will have to comply with the CF conventions (Climate and Forecast) see http://cfconventions.org) as much as possible, which provides metadata within the dataset. Each dataset also has a landing page on the ICOS Carbon Portal that has detailed information and search capabilities.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Keywords are given for all datasets and will be searchable from the ICOS Carbon Portal.

Will metadata be offered in such a way that it can be harvested and indexed?

Metadata can be accessed in human and machine-readable way through content negotiation (xml, json-ld, rdf/turtle, rdf/xml) in the ICOS Carbon Portal. Data will be indexed by gas, measurement instrument and location, amongst other parameters depending on the data type making it easily indexed and harvested.

4.3.2.2 Making data accessible - Repository

Will the data be deposited in a trusted repository?

Data will be deposited and shared on the ICOS Carbon Portal (see https://www.icos-cp.eu/observations/carbon-portal), which is currently applying for a CoreTrustSeal (see https://www.coretrustseal.org/), which is a certification for data repositories that ensures a reliable operation of the repository. All data is securely copied using the EUDAT CDI B2SAFE service, which stores two copies of each digital artefact at two locations (one at CSC, Finland, and one at FZJ, Germany). The ICOS Carbon Portal also stores a local copy of the data, with separated backups at two different buildings at Lund University, Sweden. All relevant project partners will be assisted by CP in upload of the data and accompanying metadata.

A more detailed procedure for how to upload data to the ICOS Carbon Portal will be given in the updated Data Management Plan in 2024.

The aerosol data in WP7 will be stored in the EBAS database and linked via the Carbon Portal.

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Arrangements with the ICOS Carbon Portal have been made to deposit data, archive it, and make it available according to the FAIR principles. All data will be distributed under a CCBY4 data license (see https://www.icos-cp.eu/data-services/about-data-portal/data-license). Metadata is licensed as public domain (CC0).

For WP7, the Dublin site is registered on EBAS and new aerosol data will be deposited there in a similar manner to previous campaigns/project data for this site.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

Any dataset will have assigned a Handle PID that resolves into a landing page with a qualified link to the dataset. DOIs can be minted also for example for collections of data through the ICOS Carbon Portal (see https://doi.icos-cp.eu/) and are linked to the datasets through the Handle PID. Before uploading to the ICOS Carbon Portal, the data will be curated by experts from the ICOS Carbon Portal to ensure suitable data format and the availability of rich metadata. Data directly related to the deliverables (see Table 1) and that requires easy citation in publications will be assigned a DOI.

4.3.2.3 Making data accessible - Data

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

All data relevant to the deliverables (see Table 1) will be made publicly available. Other data will in principle also be publicly available, with the possibility of defining a moratorium date for release. And exception can be made to provide data only as available upon request on explicit request by the data provider. Publications will be made Open Access.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g., patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

In general, we do not intent to apply embargos to the data products made available through deliverables. However, in the case individual publications we would apply an embargo (moratorium) on a given data set until a publication has reached a first citable stage. An embargo can be applied on request the data provider for a maximum duration at 6 months after the end of the project.

Will the data be accessible through a free and standardized access protocol?

Data will be freely available from the ICOS Carbon Portal, which has access protocols in place (see https://www.icos-cp.eu/data-services/about-data-portal/how-to-use).

Downloads use standard http protocol and do not require special software or registration. Aerosol data will be findable via the ICOS Carbon Portal and freely available via the EBAS database (ebas-data.nilu.no).

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

There are no restrictions on access to the data.

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

Data can be downloaded anonymously through the ICOS Carbon Portal. User can register and then enable bypass of the license check per individual download after acknowledgement of the data license agreement in their profile. Data downloads are logged for statistical analysis by IP, and in the personal profile. The personal profile data can be deleted on request by the user at any time or when the account is removed by the user.

Only the user has access to the personal download history. The ICOS data license requires users to contact the data provider if the data is used in a publication.

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

There is no need for an access committee. The only personal data that is gathered are the e-mails of the participants of the projects for e-mail lists. They will be treated confidentially and according to EU privacy regulations.

4.3.2.4 Making data accessible - Metadata

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Yes, data and metadata will be available on the Carbon Portal under a CC BY 4.0 licence and contains information on how to access the data.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Data and metadata will be available over timespans longer than 5 years after the end of the project through the ICOS Carbon Portal. ICOS Carbon Portal has a long term (>20 years) commitment on preserving the data and metadata and has a contingency plan in case the ICOS Carbon Portal should cease to exist by providing a safe harbour for the data.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

Data are output in formats that follow the community standards.ICOS also hosts Jupyter services where Notebooks to access the data can be created, shared, and archived (see https://www.icos-cp.eu/data-services/tools/jupyter-notebook). These are easily accessible and useable computer scripts that allow the user to read and manipulate data. These will be linked from the dataset landing pages to assist users in accessing data, where needed.

4.3.2.5 Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

Data are output in formats that are commonly use in the GHG and aerosol field. This includes NetCDF data formats (supported by NCAR), and ObsPacks (supported by NOAA and ICOS). These files built on the CF-conventions also adhered to in CMIP intercomparisons and other international geophysical projects. These formats can be read and manipulated by open-source software packages and on multiple computer platforms. They have been used for more than a decade and are expected to remain stable for the near future. The metadata is provided by ICOS Carbon Portal following schema.org (indexed by Google Data Search), rdf, xml (iso19115) and turtle. All metadata can be queried through the open Sparql endpoint of the repository.

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? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining, or extending them?

We will use the ATMODAT standard (https://www.atmodat.de/atmodat-standard) and CF conventions (http://cfconventions.org) where possible. The ATMODAT standard defines requirements for rich metadata with controlled vocabularies, structured landing pages, and for the format and internal structure of the data files.

Will your data include qualified references[1] to other data (e.g., other data from your project, or datasets from previous research)?

Yes, qualified references to other dataset will be included in the metadata where applicable.

[1] A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)

4.3.2.6 Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Landing pages of the datasets on the ICOS Carbon Portal contain documentation, e.g. variable definitions, units, and a description of the dataset. These are defined in a standard way by the ICOS Carbon Portal.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Data will be available under the CC BY 4.0 licence, also CC0 is available. Other data licenses can be adopted on request.

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Data will be publicly and freely available.

Will the provenance of the data be thoroughly documented using the appropriate standards?

Metadata and documentation will include a description of the provenance, and will include proper citation to other datasets where applicable. Metadata in the ICOS Carbon Portal will include a citation string for proper citation of the data.

Describe all relevant data quality assurance processes.

QA of the ICOS observational data is assured through procedures documented at the ICOS Atmospheric Thematic Center and the Ecosystem Thematic Center, or through the procedures applied by the AGAGE community. Observations that fall outside this scope will be calibrated to the international air gas standards of the international governing bodies (e.g., WMO).

Inverse modeling data are computer-generated and their QA is not based on international protocols, but on expert judgement from the Work Package members instead. Software packages used for inverse modelling (transport models, inversion codes) are developed through version control (git) and versions/tags will be included as part of the metadata. A more detailed QA procedure for data sets and deliverables will be included in the updated DMP in 2024.

Aerosol data will follow ACTRIS and Cost Action COLOSSAL standard measurement protocols:

https://www.actris.eu/sites/default/files/Documents/ACTRIS-2/Deliverables/WP3_D3.3_M16.pdf

https://cost-colossal.eu/wp-content/uploads/2021/05/COLOSSAL-WG1_D1.1_Q-ACSM-Standard-Operating-Procedure_May2021.pdf)

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

The main outcome of the project is digital data. In addition, we expect scientific publications, which will follow EU-guidelines for open access and archival. Code created in this project will be freely available, where possible distributed through GIT repositories. Notebooks for data/model analyses will be archived on the ICOS Carbon Portal.

4.3.3 Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

The management of other research output will be with the creators.

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

Besides the mentioned outputs in 2.4, we foresee project deliverables in the form of reports. These will be archived on the file sharing service based on Nextcloud called Fileshare of the project, on the ICOS Carbon Portal, as well as submitted to the EU for dissemination and archiving.

4.3.4 Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

We have budgeted 100 k€ for direct costs and 25 k€ for indirect costs. These are largely costs of personnel to handle all data streams and to interface with the ICOS Carbon Portal. Storage/archival/security is handled at the ICOS CP to no extra costs of this project, as part of their community data services as an ESFRI.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

The grant for this project covers the mentioned costs.

Who will be responsible for data management in your project?

Dr Remco de Kok, PostDoctoral researcher at Wageningen University tasked with ICOS Carbon Portal data services for the inverse modeling community.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

The data will be stored in the long term on the ICOS Carbon Portal, which is already set up for this. Long term for ICOS means well beyond the 5-year recommitment period of all member countries, that support this ESFRI for carbon cycle research data.

4.3.5 Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

Data on the Carbon Portal has an extensive back-up and recovery capability, as mentioned under 2.2.

No sensitive data is stored.

4.3.6 Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

There are no ethics or legal issues foreseen. The ICOS Carbon Portal data license (https://www.icos-cp.eu/data-services/about-data-portal/data-license) includes the following disclaimer: "The ICOS data products are provided "as is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose, title and non-infringement. In no event shall the copyright holders or anyone distributing the ICOS data products be liable for any damages or other liability, whether in contract, tort or otherwise, arising from, out of or in connection with the ICOS data products"

Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

No personal data will be gathered in this project.

4.3.7 Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

We will not use additional data management procedures.

5. History of the deliverable

Version	Lead Author(s), contributors	Date	Changes
1	W. Peters, S. Walter, A. Vermeulen	December 2022	First discussions of set-up data management
	W. Peters, R. de Kok, WP leads	January – February 2023	Individual meetings with WP leaders to discuss needs and requirements for data management per WP
	R. de Kok	March – April 2023	Preparation draft version DM Plan
	S. Walter, WP leaders, A. Vermeulen	May – June 2023	Feedback rounds
	R. de Kok, S. Walter	July 2023	Finalising DM plan & submission