



Title	Report describing data curation service
Work package n°	5
Deliverable n°	3
Lead beneficiary	NILU
Author(s)	Lise Eder Murberg (NILU), Lynn Hazan (ICOS), Damien Boulanger (IAGOS), Guillaume Brissebrat (CNRS)
Deliverable Type	R
Dissemination Level	PU
Estimated delivery date	31 st of March 2023 (M24)
Actual delivery date	23 rd of May 2023
Version	Final
Reviewed by	Cathrine Lund Myhre (NILU)
Accepted by	
Comments	





Table of Content

Summary	3
1. Virtual Access Portal	3
2. ATMO-ACCESS Homeless Data Portal	4
2.1 Starting page	5
2.2 Data Curation request page	6
2.3 Data Access page	7
2.4 Support Request page	8
3. Redmine Issue Tracker	8
4. Data curation	9





Summary

This report describes the data curation service offered within ATMO-ACCESS. The ATMO-ACCESS Homeless Data Portal (<https://homeless-data-portal.nilu.no>) is the core and entrance to this, and the service is available for all users through this portal. The portal is set up to serve scientists producing atmospheric measurements and time series resulting from research campaigns and TNA activities that are normally not included into any data management and data curation system and activity. Depending on the needs, the portal will route the user requests seamless to the corresponding RI experts. This report describes the data curation service set up in more detail.

1. Virtual Access Portal

The ATMO-ACCESS Homeless Data Portal is available through the ATMO-ACCESS Virtual Access Portal, which provides access to several new online services developed in the ATMO-ACCESS project.

ATMO ACCESS

Contacts Documentation ATMO-CONNECT

PROJECT ACCESS FACILITIES EVENTS & OUTREACH

ATMO-ACCESS Virtual Access Portal

This portal provides you with access to the new online services developed in the ATMO-ACCESS project. This include access to:

- Homeless data portal:**
A portal for submission of measurement data from research projects, not associated to any long-term projects/networks nor sustainable data centres. - **New and online**
- Footprint analysis tool for greenhouse gases, aerosols, reactive trace gases:**
Model tools for interpretation of measurement data both measured at the ground and from aircrafts. You can request model runs to produce data products (e.g footprint, source contributions) for your decided locations, or search and use the already produced products. - **Will be ready and launched May 2023**
- Time series analysis:**
Identify, utilize and combine data more effectively across RIs and data repositories, including data coverage, collocation of data and visualisation of data. - **Will be ready and launched May 2023**
- Massive open online course (MOOC)**
Online course on atmospheric observations in European infrastructure scheduled autumn 2023. - **Scheduled autumn 2023**

[LOGIN](#)

i You need to register to get access to the services. These services offered by ATMO-ACCESS are pilots for improvement over the project period. Accordingly, we need registration of users to get feedback and access statistic of the various services to meet the project requirements and increase the possibility of more funding for further development. When we collect personal data (e-mail) it is only for making it possible to contact you by e-mail. The information will be stored maximum 6 months after project completion, and you will only be contacted to provide feedback.

Figure 1: ATMO-ACCESS Virtual Access Portal (link: <https://www.atmo-access.eu/virtual-access/#/>)

When using the ATMO-ACCESS Virtual Access Portal the user is required to log in, as seen in Figure 1. This is to gain access statistics for the project, and to collect feedback on the services developed. Figure 2 shows the ATMO-ACCESS Virtual Access Portal after login.





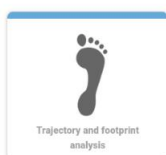
ATMO-ACCESS Virtual Access Portal

This portal provides you with access to the new online services developed in the ATMO-ACCESS project. This include access to:

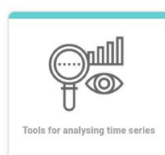
- **Homeless data portal:**
A portal for submission of measurement data from research projects, not associated to any long-term projects/networks nor sustainable data centres. - **New and online**
- **Footprint analysis tool for greenhouse gases, aerosols, reactive trace gases:**
Model tools for interpretation of measurement data both measured at the ground and from aircrafts. You can request model runs to produce data products (e.g footprint, source contributions) for your decided locations, or search and use the already produced products. - **Will be ready and launched May 2023**
- **Time series analysis:**
Identify, utilize and combine data more effectively across RIs and data repositories, including data coverage, collocation of data and visualisation of data. - **Will be ready and launched May 2023**
- **Massive open online course (MOOC)**
Online course on atmospheric observations in European infrastructure scheduled autumn 2023. - **Scheduled autumn 2023**



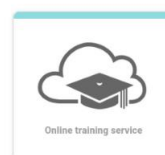
Data curation of homeless data



Trajectory and footprint analysis



Tools for analysing time series



Online training service

Figure 2: ATMO-ACCESS Virtual Access Portal page after login

By clicking the 'Data curation of homeless data' button, as seen in Figure 2, the user will be directed to the homeless data portal web page (<https://homeless-data-portal.nilu.no>).

2. ATMO-ACCESS Homeless Data Portal

This report describes the data curation service offered within ATMO-ACCESS. The ATMO-ACCESS Homeless Data Portal (<https://homeless-data-portal.nilu.no>) is a portal for data curation of atmospheric measurement data. The portal is set up to serve scientists producing atmospheric measurements and time series resulting from research campaigns and TNA activities that are normally not included into any data management and data curation system and activity. These data sets are "homeless data", not associated with any long-term projects nor sustainable data centres. The goal of the portal is to make more data available to the end-user, as well as providing benefits for data collectors in terms of usage tracking and access to RI (Research Infrastructure) specific tools for curation and data access.

Making the data accessible through a FAIR data centre is the most efficient way to ensure long term curation and promote visibility and access to a large research community. This homeless data portal makes these data well documented for use and easily available.

All requests for data curation and archiving services will be handled by the relevant research infrastructures, seamless for the users. The user needs to log in at the Redmine support system for ATMO-ACCESS at <https://redmine.aeris-data.fr> to see the status of their request.



The architecture of the portal is described in section 2 of deliverable [D5.1 – Detailed requirements and \(non-\) technical specifications for all services based on the user consultations](#).

2.1 Starting page

Homeless Data Portal Home Data Curation Data Access Support

ATMO-ACCESS Homeless Data Portal

The Homeless data portal is a portal for data curation of atmospheric measurement data. The portal is set up to serve scientists producing atmospheric measurements and time series resulting from research campaigns and TNA activities that are normally not included into any data management and data curation system and activity. These data sets are "homeless data", not associated with any long-term projects nor sustainable data centers. The goal of the portal is to make more data available to the end-user, as well as providing benefits for data collectors in terms of usage tracking and access to RI (Research Infrastructure) specific tools for curation and data access.

All request for data curation and archiving services will be handled by a relevant research infrastructure. You must log in at the Redmine support system for ATMO-ACCESS at <https://redmine.aeris-data.fr> to see the status of your request.

Workflow:

- Researchers request access to data storage and curation services offered by the Research Infrastructures (RI)
- Request is forwarded to the relevant RI in the Homeless Data Portal management system
- Data Curator at the relevant RI trigger RI specific workflows for curating and storage of the data
- Data will be handled according to FAIR principles and assigned DOIs for datasets and collection of datasets
- After the data curation process is finished. The is available for download and plotting

Services:

- Data Archiving and Curation:** Apply for data curation and archiving services through the atmospheric RIs. [Access Services](#)
- Data Access:** Search ATMO-ACCESS related data from TNA and campaign activities. [Search Data](#)
- General Support:** Send a request: if you have questions regarding the data in ATMO-ACCESS. [Get Support](#)

© 2022–2023 ATMO-ACCESS

Figure 3: ATMO-ACCESS Homeless Data Portal Starting page (link: <https://homeless-data-portal.nilu.no>).

The starting page of the portal, seen in Figure 3, consists of a description of the service, a graphic of the workflow and buttons to the Data Archiving and Curation page, Data Access page and General Support page.



2.2 Data Curation request page

Request data curation services

Fill in the information below and you will be contacted by one of the ATMO-ACCESS partners.

All request for data curation and archiving services will be handled by a relevant research Infrastructure. You must log in at the Redmine support system for ATMO-ACCESS at <https://redmine.aeris-data.fr> to see the status of your request.

Contact Information

First Name

Last Name

Position

Organization Name

Country

-- select a country --

Email

Measurement Information

Atmospheric component(s) you are working with

Aerosol
Greenhouse gases
Reactive trace gases
Clouds

Observation Type

-- select a observation type --

What specific kind of measurement data do you want to submit and archive for long term access?

-- select a measurement type --

Location of measurements site(s)?

Europe

Data specific information

What is the nature of the data you would like to store?

-- select a data type --

Is your data currently formatted and undergoing certain quality assurance criteria, or unformatted/not quality assured?

-- select an option --

Specific Constraints

-- select an option --

Is this a collection of datasets

Yes

Expected volume of the data (Estimated size of file or size of collection)

Data has been produced as part of one or more of the following projects

AMAP
ATMO-ACCESS TNA
EMEP
GAW

Other relevant information (e.g. other associated projects, station information etc.)

Upload file example (Maximum size: 2 MB)

Browse... No file selected.

Send Request





Figure 4: ATMO-ACCESS Homeless Data Portal – Data Curation Request page (link: <https://homeless-data-portal.nilu.no/request>).

The Data Curation request page of the Homeless Data Portal consists of a form where the user can request data curation services from the ATMO-ACCESS partners. This form is seen in Figure 4 and includes **Contact Information** (name, position, organization, country and e-mail), **Measurement Information** (atmospheric component(s), observation type, measurement type and location of measurement site) and **Data specific information** (data type, formatted or not, specific constraints, collection of datasets, expected volume of data, associated projects, other relevant information). Lastly, there is the possibility of uploading an example file.

2.3 Data Access page

Figure 5: ATMO-ACCESS Homeless Data Portal – Data Access page (link: <https://homeless-data-portal.nilu.no/data-access/>).

The Data Access page of the Homeless Data portal, seen in Figure 5, gives access to data from the different Research Infrastructures that are part of the ATMO-ACCESS project by describing and linking to their respective data portals. The data portals are: ACTRIS Data Portal (<https://actris.nilu.no/>), IAGOS Data Portal (<https://www.iagos.org/>) and ICOS Data Portal (<https://data.icos-cp.eu/portal/>). Homeless data managed by ACTRIS will appear in the ACTRIS Data Portal with an 'ATMO-ACCESS' tag, while homeless data managed by IAGOS and ICOS will be accessible through a common catalogue managed by AERIS (<https://www.atmo->



access.eu/homeless-data-aeris/). More details on homeless data curation and data access are available in section 4.

2.4 Support Request page

Homeless Data Portal

Home Data Curation Data Access Support

General Support

Let us know if you have questions or requests related to homeless data in ATMO-ACCESS

Contact Information

First Name

Last Name

Position

Organization Name

Country
- select a country -

Email

Request

Send Request

Figure 6: ATMO-ACCESS Homeless Data Portal – General Support Request page (link: <https://homeless-data-portal.nilu.no/support>).

The General Support request page of the portal, seen in Figure 6, consists of a simple support form. This form includes name, position, organization name, country, email address and a free text request field.

3. Redmine Issue Tracker

Redmine is an open-source project management web application. In the framework of ATMO-ACCESS we use its issue tracking feature to record, follow the progress and keep the history of all data curation requests. This service is provided by the French cluster for Atmospheric data AERIS. Issues are stored in a MySQL database with a daily backup.





Homeless Data Portal

Overview Activity Issues New issue Spent time Gantt Calendar News Documents Wiki Files Settings

Issues

Filters

Status

any

Add filter

Options

Apply Clear Save custom query

#	Tracker	Status	Priority	Subject	Assignee	Updated	Category
99	Bug	New	Normal	[Homeless Data Support] [ICOS, IAGOS, ACTRIS] Request data curation services for the following components: [reactive_trace_gases]	Homeless Data Technical Account	05/15/2023 10:04 AM	...
98	Bug	New	Normal	[ACTRIS-INSITU] Request data curation services for the following components: [aerosol] (Issue from November 2022 moved from https://atmo-access-support.nilu.no/ to Redmine)	Yong Lin	05/08/2023 04:11 PM	...
97	Support	In progress	Urgent	RI-URBANS data submission-List of sites	Lise Eder Murberg	04/27/2023 10:11 AM	...
95	Support	In progress	Normal	RI-URBANS Black Carbon data submission-filter absorption photometers	Marjan Savadkoochi	04/21/2023 10:30 AM	...
94	Bug	New	Normal	[Homeless Data Support] [ACTRIS In-Situ] Request data curation services for the following components: [reactive_trace_gases]	Yong Lin	04/28/2023 02:04 PM	...
90	Bug	In progress	Normal	[Homeless Data Support] [ACTRIS-INSITU] Request data curation services for the following components: [aerosol]	Noemi Perez	03/27/2023 12:44 PM	...
86	Support	Closed	Normal	[Homeless Data Support] [ICOS] Request data curation services for the following components: [ghg]	Lynn Hazan	03/31/2023 10:33 AM	...
85	Bug	Closed	Normal	[Homeless Data Support] [ACTRIS-INSITU] Request data curation services for the following components: [aerosol]	Cathrine Lund Myhre	03/27/2023 12:45 PM	...
82	Bug	In progress	Normal	[Homeless Data Support] [ACTRIS-INSITU] General Support request from Marvin Dufresne, marvin.dufresne@imt-nord-europe.fr	Yong Lin	04/17/2023 02:47 PM	...

(1-9/9)

Also available in: Atom | CSV | PDF

Custom queries

Demandes mises à jour
Demandes qui me sont assignées
Demandes soumises
Demandes surveillées

Powered by Redmine © 2006-2022 Jean-Philippe Lang

Figure 7: ATMO-ACCESS Redmine issue tracker (link: <https://redmine.aeris-data.fr/projects/homeless-data-portal/issues>).

The Homeless data portal is connected to the issue tracking system through the Redmine REST API. When a user sends a request for data curation or general support, a new issue appears at <https://redmine.aeris-data.fr>. The issues will be assigned to data curators from the relevant research infrastructure(s), and this/these data curator(s) will then contact and assist the user with its request. The data curator is responsible for updating the Redmine issue. The user needs to log in at to see the status of their request. The virtual access portal and the redmine rely on the AERIS authentication system (based on keycloak) to identify users and manage authorizations. It offers a Single Sign On (SSO) feature. Users can login once, using their institutional account (EduGain federation) or their ORCID, and access both systems.

4. Data curation

The users will submit their requests and the data curation will be treated seamlessly by the RIs in accordance with the workflow description in *deliverable D5.1 - section 2*. The Homeless data will be handled according to the FAIR principles as far as possible; the service also includes DOI minting for the data sets if requested by the user, either for individual data sets, or a collection DOI for a large set of data, e.g. full campaigns provide easy access to all data from a specific campaign or research project. The various RIs has different responsibilities distributed in variable and observational platform. Depending on the needs, the portal will route your request seamless to the corresponding RI experts:





- IAGOS for data collected on moving platforms (aircraft, drones, balloons);
- ICOS for surface data on greenhouse gasses including halocarbons;
- ACTRIS for data on reactive trace gasses, cloud and aerosol particle properties.

For all data compiled through the ATMO-ACCESS homeless data portal will have “ATMO-ACCESS” tag in the data centre in the respective RIs, to be able to identify these data later. Furthermore, the data can have additional tags, associating the data to a corresponding research project.

ACTRIS is offering services for data curation of aerosol, trace gas and cloud variables both in-situ and remote sensing techniques for ground-based stations. These data will undergo the ACTRIS standard data curation chain offered by the RI with traceable data production and rich meta data to document the quality of data. All data will have “ATMO-ACCESS” label in the data centre. However, if, **and only if**, the data are fully compliant with ACTRIS SOPs and recommendations (e.g. calibration at TCs) the data set will also have the tag “ACTRIS_compliant”. Definition of “ACTRIS compliant” is in the ACTRIS vocabulary; https://vocabulary.actris.nilu.no/skosmos/actris_vocab/en/page/ACTRIScompliant. The data can also be tagged with the associated research project name, upon request from the user. The ACTRIS data curation system is advanced and complex and set up to handle data from about 30 methodologies both in-situ and remote observations (50-80 instrument types), online and offline data and with time resolution of *s* to *weeks*. More details on the data curation are available in the [ACTRIS Data Management plan](#) and by contacting the data centre.

IAGOS is offering service for data curation of data collected on moving platforms (aircraft, balloons and drones). Rich metadata describing the datasets, their provenance and quality will be asked to the data providers. A specific Catalogue (<https://www.atmo-access.eu/homeless-data-aeris/>) managed by AERIS, the French Data and Services Cluster for Atmosphere provides access to metadata and data of all the curated datasets. All datasets will be labelled as “ATMO-ACCESS”.

Recommendations will be provided to the users for the data (recommended format for the files, required metadata, etc.) and simple quality control will be applied to the data files. Data providers can provide metadata and data through the edition features provided by the Catalogue or through machine actionable RESTful services. Vocabularies provided by AERIS will have to be used for the platforms (https://skosmos.aeris-data.fr/aeris_pf/en/), instruments (https://skosmos.aeris-data.fr/aeris_inst/en/) and variables (https://skosmos.aeris-data.fr/aeris_param/en/).

For more information see the [ATMO-ACCESS Data Management Plan](#).

ICOS is offering service for data curation of greenhouse gasses on fixed platforms and ships. In order to correctly trace the data, an extensive as possible set of metadata will be asked to the data providers (about the owner, measurement site and instrument as well as the measurement procedures). Curated dataset and most of the metadata will be accessible





ATMO ACCESS
Access to Atmospheric Research Facilities

through a specific catalogue managed by AERIS (<https://www.atmo-access.eu/homeless-data-aeris>).

Requirements concerning the metadata and data (naming and format of the files) will be provided to the users. A simple quality control will be applied to the data and the original quality flag if any will be traced. If requested averages can be computed, a more elaborated quality control can be applied and an access to the ICOS quality control tool can be granted.

