

GHG-Europe Data policy

1. Scope

The GHG-Europe Data Policy addresses issues related to the provision, exchange, availability, maintenance, use and publication of data and products produced by the GHG-Europe research activities or acquired from third parties as a support to this research.

The GHG-Europe Data Policy is a voluntary commitment negotiated with all GHG-Europe Parties supporting the Open Access philosophy.

2. Purpose

The purpose of the GHG-Europe Data Policy is to set up fundamental principles in view of

- easing collaboration among the participants of the GHG-Europe project;
- ensuring timely submission of quality-checked, consistent data and products for the use within the consortium;
- protecting the researchers' IPR and rights to publish their results;
- providing rules for the use of the data within the consortium and by third parties;
- building a homogeneous body of valuable scientific information on GHG fluxes over European ecosystems;
- providing the broader scientific community with an easy access to the above integrated knowledge during and beyond the lifetime of GHG-Europe.

3. Principles

- (1) All scientific data, tools and products used in publications of GHG-Europe are unambiguously identified and made publically accessible via a central database ("GHG-Europe database") so that the research results and the way towards the results are fully traceable by independent experts.
- (2) The GHG-Europe data policy is in line with the recommendations of the Global Climate Observing System (GCOS) and the Group on Earth Observation's (GEO) guidelines and data principles. The GHG-Europe Consortium endorses the GEOSS Data Sharing Principles¹:
 - a. There will be full and open exchange of data, metadata, and products inside and outside GHG-Europe, recognizing relevant international instruments and national policies and legislation.
 - b. All shared data, metadata, and products will be made available with minimum time delay.
 - c. All shared data, metadata, and products will be free of charge or no more than cost of reproduction for research and education.
- (3) An open data policy on its own is not sufficient to ensure scientific quality and traceability. Data quality control and assessment and rigorous, consistent uncertainty analysis need to be an integral part of an open data policy.
- (4) The GHG-Europe data policy ensures fair data use that involves several stages of information and opportunity of data owners to actively participate in scientific work, rules for crediting authorship and consent to publication by data owners.

¹ GEO-VI – 17-18 November 2009 Document 7(Rev2), [http://www.earthobservations.org/documents/geo_vi/07_Implementation Guidelines for the GEOSS Data Sharing Principles Rev2.pdf](http://www.earthobservations.org/documents/geo_vi/07_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf)

- (5) Recognizing the strong synergy with running EU projects NitroEurope IP, CC-TAME and Carbo-Extreme, the Parties agree, where possible, and subject to existing or future third party rights, to grant full immediate access to their data, metadata and products needed for research in the respective projects via cooperation agreements. For the avoidance of doubt however, no Party is under any obligation to agree to deliver such data.
- (6) GHG-Europe engages to include its database into a web portal of carbon and GHG data.
- (7) GHG-Europe encourages the Partners to endorse the Open Database License (PDbL; Open Data Commons 2010²) with its social and professional norms for acknowledgement and putting derived works in the public domain.

4. Data management infrastructure and operations

GHG-Europe is building the GHG-Europe database. The Università della Tuscia (UNITUS) is developing and operating this database in coordination with the Johann Heinrich von Thünen-Institut (vTI).

UNITUS handles all data, metadata and products except for the atmospheric concentration data and the results of the atmospheric model inversions from Task 5.3, which are handled by Commissariat à l'Énergie Atomique (CEA-LSCE).

5. Contents of the GHG-Europe database

The GHG-Europe database will host consolidated data, metadata and products from CarboEurope-IP and data collected and produced in the GHG-Europe project. The GHG-Europe database will not archive raw data and interim steps of data processing and modeling. Any information hosted in the database is a “dataset”, or “data”

- a. Documented, quality-checked observations and secondary data products with uncertainties + author/owner
- b. Documented, quality-checked modeling protocols with uncertainties + author/owner
- c. Documented, quality-checked model parameters with uncertainties + author/owner
- d. Documented, quality-checked model results with uncertainties + author/owner
- e. Executable files of the models used to generate the results
- f. Documented, quality-checked drivers with uncertainties + author/owner
- g. Meta-analysis data with uncertainties + references

In recognition that raw data represent a potentially valuable source of future science developments, including possible revision of the processed data, the Principal Investigators (PIs) agree to ensure that they are stored safely with the relevant processing software or, alternatively, with documentation on retrieval algorithms.

6. Data set identification

Data and products submitted to the database, which are not yet consolidated for publication or yet restricted for internal use within the GHG-Europe project are identified by file name and versioning. For all data and products in the public domain of the GHG-Europe database, uniform resource identifiers and universal numerical fingerprints are applied to the data sets. Recognizing recommendations of the World Data Centres individual data sets and products are identified by

² <http://www.opendatacommons.org/licenses/odbl/>

Scientific and Technical Data Digital Object Identifiers (STD-DOI)³, which are globally administered by the German National Library of Science and Technology.

7. Metadata standards

GHG-Europe complies to metadata standards for inclusion of the database in other networks. Metadata standards will be defined in the Protocols developed during the project. The Protocols will ensure that the metadata standards comply with the requirements for unambiguous identification, authorship and ownership of the data, transparency, versioning and publication with STD-DOI.

8. Authorship of data

Every dataset has an author and an owner identified in the metadata of the dataset. An author is a natural person. An owner is either an institution, a natural person or a group of persons or a combination of persons and institutions according to the legal situation of the data contributors. Every dataset has one contact person identified in the metadata of the dataset. The contact person receives all requests and information about downloads of data. The contact person is responsible for quality control, updating the dataset or changes in the metadata of the dataset.

9. Submission of data to the GHG-Europe database

In order to warrant timeliness and continuity to the research activities within GHG-Europe, it is essential that data produced by one Partner is delivered as soon as possible to the other Partners via the GHG-Europe database. Datasets collected over some defined period of time (of the order of three to six months) and products have been allocated submission deadlines related to deliverables in the Description of Work. Data submitted to the database cover the reporting periods and previous periods in case of updates. Data submission has to follow the Protocols established during the project.

10. Use of data

The GHG-Europe data policy ensures fair data use that involves several stages of information and opportunity of data owners to actively participate in scientific work, rules for crediting authorship and consent to publication by data owners.

The database is separated into a *public* and an *internal project domain*.

Data within the database will be immediately public unless individual property rights restrict or delay publication. In the latter case the data are stored in the internal project domain of the database. PIs and their PhD students are obliged to grant immediate access to the data for project partners for project purposes according to the rules below. Agreement to publication of the data shall not unnecessarily be delayed. The period between producing the data and their publication, to validate and analyse the data, should not exceed two years.

Conditions of use of data in the internal project domain

These data were generated or collected within the framework of GHG-Europe project.

The user agrees

- to contact the Data Originator prior to any use of the data,
- not to disclose the data to other parties,
- to restrict the use of the data to the context of the research topic as agreed with the Data Originator,
- that the Intellectual Property Right remains with the Data Originator,

³ STD-DOI. 2010. Publication and Citation of Scientific Primary Data. <http://www.std-doi.de/>, viewed 2010-04-22.

- not to use the data for commercial purposes. Access Rights to Foreground data, generated within the GHG Europe Project, if Needed for Use of a Party's own Foreground shall be governed by the terms of Article 9.4. of the GHG-Europe Consortium Agreement.
- to offer the Data Originator(s) an intellectual participation in the planned research,
- to offer co-authorship of any publication or communication based on these data, if the Data Originator(s) participated intellectually in the research;
- in the event that the Data Originator(s) only contributed the data, Data Originators must be duly acknowledged,
- to send the publication to all Data Originator(s) 45 days before submission. Any objection to the planned publication by the Data Originator(s) shall be made in writing to the Coordinator and to any Party concerned within 30 days after receipt of the notice. If no objection is made within the time limit stated above, the publication is permitted. The reception of such a request for publication has to be acknowledged/confirmed. The 30 days period starts with confirmation of reception.

Conditions of use of data in the public domain

Data sets in the public domain of the GHG-Europe database shall be regarded as scientific publications and shall be quoted with their STD-DOI. The data is provided without cost for online access.

11. Publication of data

To make primary scientific data citable as publications via STD-DOI³, data must meet the criteria stated below:

- Quality control of the primary data set by the author and by the data publishing agency,
- Quality control of the descriptive metadata set by the author and by the data publishing agency
- Long-term availability of the published data in online repositories (World Data Centres and GFZ Potsdam)

Published data must be openly accessible and cannot be modified, like articles in scientific literature. Changes in the data require a new STD-DOI.

To allow the use of STD-DOI in the GHG-Europe database access to the primary data with assignment of a persistent identifier and resolver system (DOI resolver) will be implemented. The author of the data is ultimately responsible for quality control. The database manager coordinates the quality control and manages the organizational and technical issues of publication via STD-DOI. The database manager establishes a practicable mechanism to define consolidated data sets and to organize the publication with STD-DOI. The assignment of a data DOI is free of charge.

12. Long-term accessibility of GHG-Europe data

To guarantee the long-term accessibility of GHG-Europe data, metadata and products, publication via STD-DOI and cooperation with emerging research and IT infrastructures (e.g. ENCORE, GENESI-DR⁴, INSPIRE⁵, ICOS⁶) is sought.

Amalgamation of the project data with other data requires the authorization of the project coordinator, the General Assembly, or where applicable, the original providers (in case of external data sets) used in GHG-Europe. Amalgamated data shall recognize the contribution of GHG-Europe via acknowledgement: *The data were acquired in the EU project GHG-Europe (EC FP7, GA N° 244122).*

⁴ <http://www.genesi-dr.eu/>

⁵ <http://inspire.jrc.ec.europa.eu/>

GHG-Europe encourages the Partners to endorse the Open Database License (PDbL; Open Data Commons 2010⁶) with its social and professional norms for acknowledgement and putting derived works in the public domain. The PDbL governs the rights over the data base, and not the contents of the database individually. Endorsing the PDbL means that the database is made freely available to any kind of user in the belief that its wide dissemination will lead to greater understanding and new scientific insights and that global scientific problems require international cooperation. The endorsement of the PDbL requires the authorization by the General Assembly.

13. Data quality and completeness

It is the PIs' responsibility to perform required calibration and validation prior to data submission to the GHG-Europe database, and ensure that the data are of the best possible quality and completeness. It will be one of the roles of Work package 7 to support and supervise quality checks.

⁶ <http://www.icos-infrastructure.eu/>