ECO2 – Data Policy

In accordance with the FP7 Rules governing the Intellectual Property Rights (IPR) of the funded project clearly stated in the ECO2 documents [ECO2 grant agreement (Article 7 – Special clause 39), Part B of Annex I DoW to ECGA (B2.1 – *Management structure and procedures – data management – communication*, B3.1 – *Strategic impact*, B3.2 – *Plan for the use and dissemination of foreground – Dissemination of project information and results – IPR and their management*, B4 – *Ethical Issues*) and the ECO2 Consortium Agreement (Section 8 – *Foreground*, Section 9 – *Access Rights* (Attachment 1 – *Background included*, Attachment 2 – *Background excluded*), Section 10 – *Non-disclosure of information*) ECO2 is thereby identifying the rules of its data policy.

The aim of this data policy is to facilitate operation and use of the ECO2 data base by the research consortium of ECO2. The ECO2 database that gathers all field, experimental and modelling data, will be archived in and published by the PANGAEA data library, operated at the Alfred-Wegener-Institute in Bremerhaven. Data generated within WP5 – *Risk Assessment, Economics, Legal Studies and Policy Stakeholder Dialogue* and WP6 – *Public Perception Assessment* will be stored at the secure part of the ECO2 website as text files. This policy recognises the benefits of providing free and open access to good quality data from environmental and socioeconomic sciences in order to achieve ECO2 goals as specified in the DoW as well as for future use in global CCS studies, ocean acidification, and natural seepage studies, research projects and operational services such as portals and search engines. The steering committee and the scientific board of ECO2 encourages the widest possible use of the data base, in order to best realise the values of the ECO2 data collection during the project as well as after its completion.

General Principles

- The ECO2 project data base on PANGAEA is the central archive to all ECO2 field, experimental, and modelling data. The system supports free and Open Access to data in a partnership approach within the ECO2 consortium. The scientist or team generating the data should have the opportunity to first publish the data, and analysis based on them.

- For any data, provided to the ECO2 data base, the format and description of a data set must ensure its most widespread and easiest use by the scientific community following the principles of PANGAEA.

- ECO2 scientists using data from the ECO2 data base are urged to properly use the data set citation or quote the related reference of supplements.

Data provision

- Metadata of expeditions/stations/samples (events) are submitted to the project management office (email: areitz@ifm-geomar.de OR datenmanagement@ifm-geomar.de) within 14 days after the end of an expedition or field campaign.

- The data manager maintains a dictionary of parameters definitions with units, to be used as the agreed standard for all project data. Parameters are grouped into categories according to their related scientific field. Data submissions are required to use parameters and units as defined in the dictionary. New parameters are defined by the data librarian on request.

- Data are archived in a relational database, geo-referenced in space and time; if a data set is very large or must have a proprietary format, it is archived as an object in a file system with a meta-description only, linked to the file.

- As soon as data has been measured/produced and validated, the PI (principle investigator) is urged to submit the data in agreement with the import format. Data has to be provided latest 3 month subsequent to request by the project management office. The work package leader is responsible of data delivery. Any type of data must always be accompanied by a description (metadata) allowing future users to understand and process the data.

Quality assurance

- Data submitted have to be documented properly; documentation is stored together with each dataset.

- The scientific quality is always in the responsibility of the PI and all authors. Fields for its documentation like quality flags for single values, adjustable precision or documentation of methods are available.

- Technical quality control, i.e. completeness of metadata, consistence of formats, and correctness of download is in the responsibility of the data manager.

- After import, the PI/authors will receive the data DOI and a request for proof-read. Corrections has to be send to the data manager until the data set is finally approved by the PI.

Access to ECO2 data

- Unpublished data are password protected by default. Providers may decide to withdraw data from the archive as long as it is not published. Metadata are always freely accessible.

- Project members commit themselves neither to provide data nor access passwords to third parties. Excluded are the data that is owned by a scientist or team. In case one group or work package needs the data of another scientist or group the respective group has to be informed.

- Data are made available under a Creative Commons Attribution (CC-by) license if not otherwise requested and outlined in the metadata.

Publication of ECO2 data

- According to EU data policy, all data collected during the lifetime of the project are finally made public two years after the termination of the project but can be published any time before this deadline (e.g. supplement to a publication or as Earth System Science Data (ESSD) publication).

- Partner institutes and data providers agree, that data archived in Pangaea are made public available through appropriate technical setups on the Internet (e.g. portals, search engines, library catalogs, GIS) without further notification.

- Data sets can be made citable on its own. The citation, accompanied by an extended abstract, is added to a public library catalog and will receive a DOI. Those data publications may also be added to personal or project publication lists.

- The publication of data by others than the data owner during the running time of ECO2 requires the written agreement of the data owner. An email should be send to all partners listing the owners from which an answer is requested. However, others that work on conflicting aspects should have the opportunity to make an objection.

Operation

- Long-term availability of data and operation of the system is ensured by the institutions AWI and MARUM, also responsible for the consistency of the content.

- The Backup of the data inventory is in the responsibility of the computer centre at AWI with daily incremental backup and weekly full backup in two mirrored tape drive archives, located in different buildings.

- Data flow is organised from the work pack ages via the d ata manager to the a rchiving facility, monitored by the project management and supervised by the data manager.