GEOSS Data Management Principles
Prepared by the GEO Data Management Principles Task Force
Approved by Data Management Principles Task Force on April 28, 2015

Background and Context

The background discussion for these initially proposed principles can be found in Document 7[1] from GEO XI.

The GEOSS Data Management Principles build on the GEOSS Data Sharing Principles in the sense that they adumbrate what is required in terms of data management to allow data to be shared as Open Data, promptly and at minimum cost. Good data management implies a number of activities which ensure that data are discoverable and accessible, that they may be understood and used, and that they are looked after in the long term.[2]

A priority mission for GEO is to encourage the implementation of the Principles by organizations contributing to GEOSS. Guidelines for implementation will be provided elsewhere.

GEOSS Data Management Principles

To further maximize the value and benefit from data sharing, GEO will continue to work with partners to promote the use of key Data Management Principles, including the need for common standards and interoperability arrangements. This will ensure that data and information of different origin and type are comparable and compatible, facilitating their integration into models and the development of applications to derive decision support tools. GEO therefore strives to promote and encourage the implementation of Data Management Principles laid out below under five headings: discoverability, accessibility, usability, preservation, and curation.

Discoverability
DMP-1. Data and all associated metadata will be discoverable through catalogues and search engines, and data access and use conditions, including licenses, will be clearly indicated.

Accessibility
DMP-2. Data will be accessible via online services, including, at minimum, direct download but preferably user-customizable services for visualization and computation.

Usability

---

[1] GEO XI Document 7: *Towards Data Management Principles.*
[2] G8 Science Ministers' Statement (on Open Scientific Research Data), 13 June 2013 https://www.gov.uk/government/news/g8-science-ministers-statement : 'Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards.'

DMP-3. Data will be structured using encodings that are widely accepted in the target user community and aligned with organizational needs and observing methods, with preference given to non-proprietary international standards.

DMP-4. Data will be comprehensively documented, including all elements necessary to access, use, understand, and process, preferably via formal structured metadata based on international or community-approved standards. To the extent possible, data will also be described in peer-reviewed publications referenced in the metadata record.

DMP-5. Data will include provenance metadata indicating the origin and processing history of raw observations and derived products, to ensure full traceability of the product chain.

DMP-6. Data will be quality-controlled and the results of quality control shall be indicated in metadata; data made available in advance of quality control will be flagged in metadata as unchecked.

Preservation

DMP-7. Data will be protected from loss and preserved for future use; preservation planning will be for the long term and include guidelines for loss prevention, retention schedules, and disposal or transfer procedures.

DMP-8. Data and associated metadata held in data management systems will be periodically verified to ensure integrity, authenticity and readability.

Curation

DMP-9. Data will be managed to perform corrections and updates in accordance with reviews, and to enable reprocessing as appropriate; where applicable this shall follow established and agreed procedures.

DMP-10. Data will be assigned appropriate persistent, resolvable identifiers to enable documents to cite the data on which they are based and to enable data providers to receive acknowledgement of use of their data.