Implementing GEOSS Data Sharing Principles with GEOSS Data Management Principles

Robert R. Downs
October 23, 2017

Best Practices in the GEOSS Common Infrastructure
GEO XIV Side Event, Washington, DC
Data Sharing and Data Management

Improving Community Capabilities

How can we improve practices for sharing and managing Earth science data?

Implement the GEOSS Data Sharing Principles and the GEOSS Data Management Principles!

These principles are complementary and offer opportunities for improving data sharing and data management.
Let’s implement them together!

Both sets of principles are under review, culminating in a review for consistency.

A Joint GEO Subgroup on Data Documentation has been proposed as a new subgroup of the Data Sharing Working Group, with goals to:

1. Identify GEOSS DMP Contributing to GEOSS DSP and Identify GEOSS DSP Contributing to GEOSS DMP
2. Describe Opportunities for Implementing GEOSS DSP and GEOSS DMP in Concert
3. Produce Examples of Mutual Attainment of Principles by Implementing GEOSS DSP or GEOSS DMP
4. Provide Recommendations for Improving the GEOSS DSP Implementation Guidelines (IG) and GEOSS DMP IG Based on Synergies
### GEOSS Data Sharing Principles

#### GEOSS Data Sharing Principle 1.

Data, metadata and products will be shared as Open Data by default, by making them available as part of the GEOSS Data Collection of Open Resources for Everyone (Data-CORE) without charge or restrictions on reuse, subject to the conditions of registration and attribution when the data are reused;

#### GEOSS Data Sharing Principle 2.

Where international instruments, national policies or legislation preclude the sharing of data as Open Data, data should be made available with minimal restrictions on use and at no more than the cost of reproduction and distribution;

#### GEOSS Data Sharing Principle 3.

All shared data, products and metadata will be made available with minimum time delay.
<table>
<thead>
<tr>
<th>GEOSS Data Management Principles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discoverability</strong></td>
</tr>
<tr>
<td>DMP-1: Metadata for Discovery</td>
</tr>
<tr>
<td><strong>Accessibility</strong></td>
</tr>
<tr>
<td>DMP-2: Online Access</td>
</tr>
<tr>
<td><strong>Usability</strong></td>
</tr>
<tr>
<td>DMP-3: Data Encoding</td>
</tr>
<tr>
<td>DMP-4: Data Documentation</td>
</tr>
<tr>
<td>DMP-5: Data Traceability</td>
</tr>
<tr>
<td>DMP-6: Data Quality-Control</td>
</tr>
<tr>
<td><strong>Preservation</strong></td>
</tr>
<tr>
<td>DMP-7: Data Preservation</td>
</tr>
<tr>
<td>DMP-8: Data and Metadata Verification</td>
</tr>
<tr>
<td><strong>Curation</strong></td>
</tr>
<tr>
<td>DMP-9: Data Review and Reprocessing</td>
</tr>
<tr>
<td>DMP-10: Persistent and Resolvable Identifiers</td>
</tr>
</tbody>
</table>


While endorsing all the GEOSS Data Sharing Principles and Data Management Principles, we can identify ways to prioritize aspects of implementation and amplify the benefits of data sharing and data management.

- Attain synergies by improving capabilities
- Prioritize implementation of infrastructure enhancements
- Balance utilization of resources with needs
- Schedule implementation activities to optimize improvements
- Facilitate data discovery and selection by users
- Increase ease of using data
GEOSS DSP & DMP: Some Common Concepts

• **Data Sharing Principle 1**
  - Open Data
  - Without charge or restrictions
  - GEOSS Data-CORE
  - Subject to conditions
  - Attribution

  - DMP-1: Metadata for Discovery
  - DMP-2: Online Access
  - DMP-4: Data Documentation
  - DMP-10: Persistent and Resolvable Identifiers

• **Data Sharing Principle 2**
  - Minimal restrictions on use
  - Cost of reproduction and distribution

  - DMP-1: Metadata for Discovery
  - DMP-4: Data Documentation

• **Data Sharing Principle 3**
  - Minimum time delay

  - DMP-6: Data Quality-Control

Implementing GEOSS DMP-1: Metadata for Discovery with the GEOSS Data Sharing Principles

NASA Socioeconomic Data and Applications Center (SEDAC) Examples

- Data are discoverable via
  - SEDAC website, NASA Common Metadata Repository, GEOSS Portal, DataCite and harvesting catalogs, and popular search engines, e.g., Google
- Data described in metadata
  - Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) HTML (also in XML and text); at GCMD in ISO 19115 (XML) and DIF 10 (XML); and DataCite Metadata Schema (XML),
- Standardized rights declarations and licenses in metadata
- Recommended citation, including persistent Id on data landing page

Use Constraints:

This work is licensed under the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0). Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

Recommended Citation(s)*:

“Without charge or restrictions” … “specify any access conditions in metadata”*

24/7 Free Online Access to Open Data from SEDAC Website

- Negotiate permissions for access to open data: Attribution only required for maps and many datasets
- Landing page for each data collection: Collection Overview, Datasets, Map Gallery, Map Services, Citations, FAQs, Acknowledgements
- Landing page with overview for each dataset: Title, Purpose, Abstract, Recommended Citation, Available Formats
- Dataset landing page tabs: Dataset Overview, Download, Maps, Map Services, Documentation, Metadata

Data created by CIESIN are licensed under the CC By license.

VIII. Use Constraints

This work is licensed under the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0). Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

Open Data Policy

Center for International Earth Science Information Network (CIESIN), Columbia University

This CIESIN Open Data Policy has been established to promote the open and free exchange of data and information in support of research, decision making, education, and other applications. CIESIN intends this policy to support and comply with relevant open data policies, guidelines, and initiatives for scientific, government, and sustainable development data. CIESIN’s policy is to make data “open by default,” with only narrow exemptions for genuine security, privacy, or legal concerns.

Implementing GEOSS DMP-4: Data Documentation with the GEOSS Data Sharing Principles

Outline of the SEDAC Data Documentation Template

Documentation for <Dataset Title>
<Documentation Publication Date>
<Authors>
Abstract
Data set citation
Suggested citation for documentation
Contact to provide feedback on documentation
Table of Contents
I. Introduction
II. Data and Methodology
III. Data Set Description(s)
IV. How to Use the Data

IV. How to Use the Data
V. Potential Use Cases
VI. Limitations
VII. Acknowledgments
VIII. Disclaimer
IX. Use Constraints
X. Recommended Citation(s)
XI. Source Code
XII. References
XIII. Documentation Copyright & License
Appendix 1. Contributing Authors & Documentation Revision History
Appendix 2. Data Revision History

Scientific Publication:

XIII. Documentation Copyright and License
Copyright © 2017. The Trustees of Columbia University in the City of New York. This document is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).
SEDAC conducts regular review procedures that include internal reviews and inviting community experts to complete reviews throughout data lifecycle.
Implementing GEOSS DMP-10: Persistent and Resolvable Identifiers with the GEOSS Data Sharing Principles

NASA SEDAC Digital Object Identifier (DOI) Examples

Policy and procedures to routinely assign and maintain DOIs

DOI for Data Set on Data Landing Page

Recommended Citation(s)*:


DOI for Software in Data Documentation


DOI for Data Documentation


DOI for Data Set in Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) HTML

DataCite Metadata XML

<identifier identifierType="DOI">10.7927/H4C53HSR</identifier>

DOIs link data to documentation, related data products, and publications
Thank You

Communicate and Collaborate with GEO: