



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

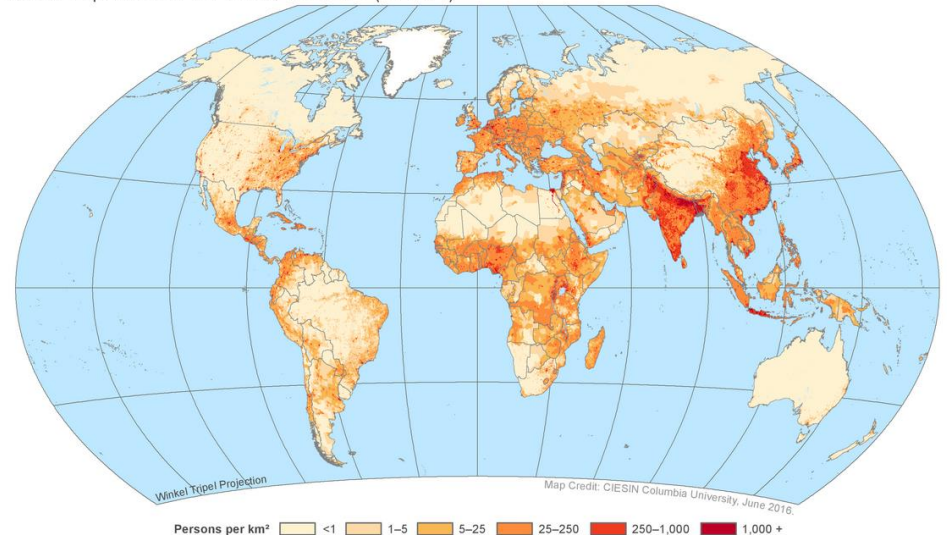
How can we improve practices for sharing and managing Earth science data?

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.

Population Density Grid, 2020: Global

Gridded Population of the World, Version 4 (GPWv4)



Gridded Population of the World, Version 4 (GPWv4) Population Density consists of estimates of human population density based on counts consistent with national censuses and population registers, for the years 2000, 2005, 2010, 2015, and 2020. A proportional allocation gridding algorithm, utilizing approximately 12.5 million national and sub-national administrative units, is used to assign population values to 30 arc-second (~1 km) grid cells. The population density grids are derived by dividing the population count grids by the land area grids. The pixel values represent persons per square kilometer.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2016. Gridded Population of the World, Version 4 (GPWv4).
Science Information Network Population Density, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <http://dx.doi.org/10.7927/H4NP22DQ>.
Earth Institute | Columbia University

© 2016. The Trustees of Columbia University in the City of New York.



This document is licensed under a
Creative Commons 3.0 Attribution License
<http://creativecommons.org/licenses/by/3.0/>

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.

GEOSS Data Management Principles

Discoverability

DMP-1: Metadata for Discovery

Accessibility

DMP-2: Online Access

Usability

DMP-3: Data Encoding

DMP-4: Data Documentation

DMP-5: Data Traceability

DMP-6: Data Quality-Control

Preservation

DMP-7: Data Preservation

DMP-8: Data and Metadata Verification

Curation

DMP-9: Data Review and
Reprocessing

DMP-10: Persistent and Resolvable
Identifiers

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



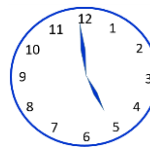
DMP-1: Metadata for Discovery

DMP-4: Data Documentation



- Data Sharing Principle 3

- Minimum time delay



DMP-6: Data Quality-Control



Derived from: GEOSS Data Sharing Principles Implementation Guidelines 2016-2025.

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

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, including persistent Id on data landing page

Recommended Citation(s)*:

Center for International Earth Science Information Network - CIESIN -
Columbia University. 2016. Global Urban Heat Island (UHI) Data Set, 2013.
Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
<http://dx.doi.org/10.7927/H4H70CRF>. Accessed DAY MONTH YEAR.



Implementing GEOSS DMP-2: Online Access with the GEOSS Data Sharing Principles



“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.

*Derived from: GEOSS Data Sharing Principles Implementation Guidelines 2016-2025.

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.

-
-
-

**Complete CIESIN Open Data Policy available at
<http://ciesin.columbia.edu/documents/CIESINDataPolicy.pdf>**

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:

Doxsey-Whitfield, E., K. MacManus, S.B. Adamo, L. Pistolesi, J. Squires, O. Borkovska and S.R. Baptista. 2015. Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. Papers in Applied Geography 1(3): 1-9. <http://dx.doi.org/10.1080/23754931.2015.1014272>.

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/>). 

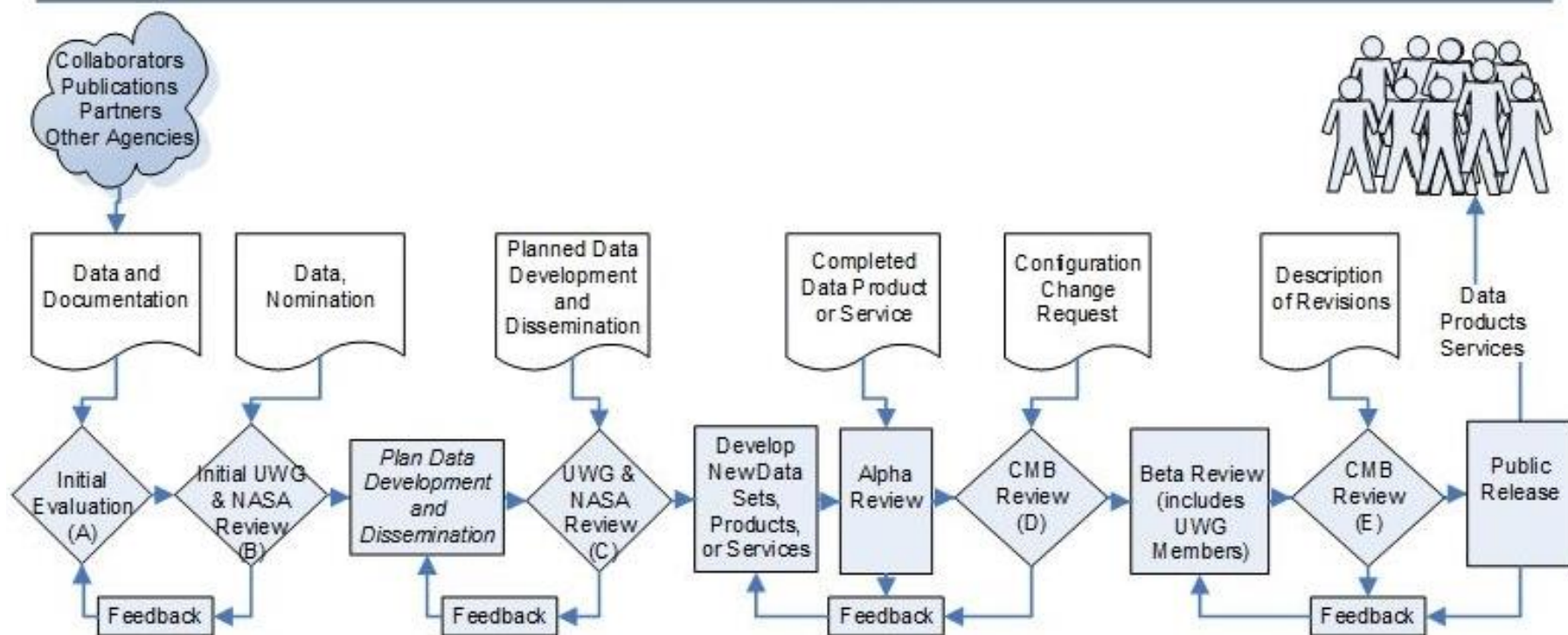
Implementing GEOSS DMP-6: Data Quality-Control with the GEOSS Data Sharing Principles



SEDAC conducts regular review procedures that include internal reviews and inviting community experts to complete reviews throughout data lifecycle

SEDAC Review For Type 1 Data and Type 2 Data - Flowchart

Revised September 27, 2013



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)*:

Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Global Urban Heat Island (UHI) Data Set, 2013. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <http://dx.doi.org/10.7927/H4H70CRF>. Accessed DAY MONTH YEAR.

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

DOI for Software in Data Documentation

Recommended citation for the source code: Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. Source code for the India annual winter productive cropped area 2001-2016. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H47D2S3W>. Accessed DAY MONTH YEAR.

DOI for Data Documentation

Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Documentation for the Gridded Population of the World, Version 4 (GPWv4). Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). <http://dx.doi.org/10.7927/H4D50JX4> Accessed DAY MONTH YEAR.

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

Originator: Center for International Earth Science Information Network - CIESIN - Columbia University
Publication Date: 2017
Title: U.S. Census Grids (Summary File 1), 2010
Edition: 1.00
Geospatial Data Presentation Form: raster, map
Publication Information:
Publication Place: Palisades, NY
Publisher: NASA Socioeconomic Data and Applications Center (SEDAC)
Online Linkage: <https://doi.org/10.7927/H40Z716C>

DataCite Metadata XML <identifier identifierType="DOI">10.7927/H4C53HSR</identifier>

Thank You

Communicate and Collaborate with GEO:

