Geospatial Data Preservation Plan Worksheet

This worksheet was developed based on templates from data management plans, data preservation strategies and FAIR data principles. It is meant to assist in capturing the information needed to think about data preservation and how it can be implemented in your work group.

Developed by Denise Bleakly, 2025, d3bleakly@gmail.com

Administrative	*Basic information about the project, data collected and the context for which the data was developed*
Basic information	Project title, contact name, contact details
Summary	Summary of data set, explain the purpose of the dataset
Guidance Documents	• Identify Organizational requirements for data preservation, • list any organizational guidance documents • list federal government guidance documents (if applicable)
Legal	*Identify any legal requirements for the preservation of your data*
	Are there any regulatory requirements for your data? (such as permit requirements, other Federal Government Requirements, state requirements, etc.) Sharing requirements will allow for planning for periodic "archives" (every 10 + years)
	Identify and address any data ownership and data distribution restrictions
	Is privacy protection in place? Have you removed personal or sensitive information from your data to ensure privacy protection?
	Have you established who owns the copyright of your data?
	Have you documented how the data should be institutionally credited or cited?
Documentation & Metadata	*Documentation of the data to enable contextual understanding and long-term usability*
	List standards and guidance used • FGDC Content Metadata standard • ISO standards • File format standards
	Inventory content and location of descriptive metadata of data sets
	Do you have an indexing system for your metadata?
	Geospatial content – have you described the map projection, datum, and other geospatial parameters for your project in general? Are the details in the dataset specific metadata?
	Have you documented software, models or codes used in the development of the data?

File Formats	*This describes the file types, data structures and naming conventions to aid long-term preservation and reuse*
	Inventory the file formats in use
	Identify repository specific file formats required (NARA/LOC/Project)
	Plan for data set conversion to required file formats
	Have you documented any naming conventions used within your files?
Data Prioritization	*This describes how the project decides what data is prioritized for preservation activities*
	Identify program/project data priorities
	Identify organizational records retention requirements
	Identify data needing to be preserved to meet legal requirements
Storage and Geographic Location	*Storage systems, locations, and multiple copies of data/metadata to prevent loss of data*
	Have you identified where your data and metadata will be stored for the long term?
	Has the data and metadata been moved from hard drives, workstations and other external media to the organization data storage system?
	Have you identified the long term repository for your data?
	Do you have at least two complete copies of your data and metadata? Is one of them stored in a different location?
Data Integrity	*Procedures to prevent, detect and recover from unexpected or deliberate changes to data*
	Describe any processes you have in place to check that the data has been unchanged.
	Describe any QA/QC processes you have for data integrity.
	Describe any processes you have in place for identifying and correcting any changes to data.
Information Security	*Procedures to prevent human-caused corruption of data, deletion and unauthorized access*
	Identify who has read, write, move and delete authorization to remove individual files and data
	Document access restrictions for content
	Maintain logs of who performed what actions on files, including deletions and preservation actions
Financial	*Identify long term costs for data preservation, data refreshing and data archiving*
	Costs for data preservation can be planned for as part of the data lifecycle
	Planning for periodic data refreshing and associated costs (every 5-10 years)

FAIR Data Principles	
Findable	
	Are the data assigned a globally unique and persistent identifier
	Are the data described with complete metadata
	Does the metadata clearly and explicitly includes the identifier of the data it describes
	Are the data and metadata registered or indexed in a searchable resource?
Accessible	
	Are the data and metadata retrievable by their identifier using a standardized communications protocol?
	Is the communications protocol, open, free and universally implementable?
	Does the protocol allow for an authentication and authorization procedure, when and where necessary
	Is the metadata accessible, even when the data are no longer available?
Interoperable	
	Does the data and metadata use a formal, accessible, shared and broadly applicable language for knowledge representation?
	Does the data and metadata use vocabularies that follow FAIR principles
	Does the data and the metadata include qualified references to other applicable metadata/data
Reuseable	
	Does the metadata richly describe the data with complete, accurate and relevant attributes
	Are the data and metadata released with a clear and accessible data usage license
	Are the data and metadata associated with detailed provenance
	Does the data and metadata meet domain-relevant community standards?

Resources for More Information

National Archives and Records Administration (NARA)

- Digital Preservation Strategy 2022-2026 *https://www.archives.gov/preservation/digital-preservation/strategy*
- The National Archives and Records Administration (NARA) maintains a current list of acceptable file formats for all data required to go into a federal data repository. *https://www.archives.gov/records-mgmt/policy/transfer-guidance-tables.html*

The Library of Congress (LOC) *also maintains a list of currently acceptable file formats if your content will be subject to storage at the Library of Congress*. *https://www.loc.gov/preservation/resources/rfs/geo-carto.html*

FAIR Data Principles

- Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
- Thompson, P.T., Ojha, S., Powell, C.D. *et al.* A proposed FAIR approach for disseminating geospatial information system maps. *Sci Data* 10, 389 (2023). https://doi.org/10.1038/s41597-023-02281-1

National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation

- Bailey, Jefferson, et al. "The NDSA Levels of Digital Preservation: An Explanation and Uses." (2013). https://digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf
- **NDSA website**: https://ndsa.org/publications/levels-of-digital-preservation/

Digital Preservation Strategies

Shimray, Somipam R., and C. Kodanda Ramaiah. "Digital preservation strategies: an overview." *11th National Conference on Recent Advances in Information Technology-2018*. 2018.

https://www.researchgate.net/profile/Somipam-Shimray/publication/327221006_Digital_Preservation_Strategies_An_Overview/links/5b80da10a6fdcc5f8b6592f4/Digital-Preservation-Strategies-An-Overview.pdf

Data Management Plan Templates

- **United States Geological Survey (USGS) Data Management Plan** website: https://www.usgs.gov/data-management/data-management-plans
- **Data Management Plan (DMP) Tool Templates** (A public collection of a wide variety of data management plan templates including Department of Defense (DOD) and Department of Energy (DOE): https://dmptool.org/public_templates