Science Applications Metadata Guidance

A guide to MDEditor and ScienceBase for SA

Table of contents

Chapter 1

USFWS Science Applications Metadata Guidance

A guide to MDEditor and ScienceBase for SA

1.1 Introduction

This manual describes the requirements and best practices for the creation of high-quality and consistent metadata records for projects and products for USFWS Science Applications. These metadata requirements conform to the standardized metadata format agreed upon by the SA Data Management Working Group and meets USFWS policy requirements for data management. This metadata drives the unified SA Science Catalog, where SA projects and products are discoverable, accessible, and usable.

> **About this documentation**
>
> This is a working draft that will continue to be edited. Please refresh this manual every time you open it to ensure you are viewing the most recent version. **Last updated: 26 March, 2024**.

If you have questions/concerns regarding this guidance, please contact the FWS SA National Data Steward (sadatasteward@fws.gov) or Blake Massey (blake_massey@fws.gov).

1.2 USFWS Data Management Integration

Metadata creation is part of the Maintain step of the Data Management Lifecycle. Guidance is also available in the USFWS Data Management Handbook.

1.2.1 Metadata and mdEditor

The metadata editing tool employed to create and support metadata requirements is mdEditor. This tool was an extension of an initiative by the Alaska Data Integration Working Group (ADIwg) and adopted by SA.

This manual refers to the creation of SA metadata specifically. If you want information on the creation of other metadata records using mdEditor, or more information in general, please refer to the mdEditor User Manual.

1.3 Who Should Use this Manual

This guide is for anyone creating or updating metadata for SA-related projects and products. The primary purpose is to describe in detail how to develop metadata for items for inclusion in the SA Science Catalog, but the tool can also be used to track any project or funding.

1.4 How to Use this Manual

Directly accessing this manual from the internet is the recommended way to use this manual. By doing so, you will guarantee that you are using the current version since it is a live document and any updates can be made instantly. The online version will also have the best functionality and graphical display.

1.5 Disclaimer

The mdEditor application is under constant development. Due to this, there may be minor discrepancies between screenshots in the manual and the current production version of mdEditor.

Chapter 2

Data Management Goals

Science Applications Program

> We envision a well-functioning data management, integration, and sharing process that integrates, at minimum, the products of SA-funded projects with those from regionally-funded projects, and that support the data lifecycle in a coordinated manner. – Conservation Science Plan v.1

In the beginning, Science Applications was leading the Landscape Conservation Cooperatives: A network of individual, geographically distinct partnerships coordinated by SA. The 22 Landscape Conservation Cooperatives (LCCs) developed uniquely and independently based on the needs of their regions and partnerships. As a result, there was not a single workflow for data management and sharing among the LCCs. Instead there were multiple workflows, varying formats, and different use cases for managing and sharing the results of LCC-funded projects. Further, because LCCs placed varying levels of importance on data management, LCCs varied greatly in their staff capacity for data management.

While the individual workflows met the needs of the individual LCCs and their partner organizations, they were not enough to achieve the goal of a network-wide, integrated dynamic project and product database. Through collective agreement, LCCs agreed they needed to modify their workflows to create comparable project and product metadata.

Procedurally, the LCC Network has committed to using ScienceBase for cataloging all funded projects and associated products. As per the LCC Network's response to a DOI Office of the Inspector General 2017 report, individual LCCs must contribute their project and product metadata to the ScienceBase repository by the end of 2017 and mid-2018 respectively. In tandem with this ScienceBase requirement, the LCC Network is required to post data products to data.gov and this data conduit will be in place with products cataloged by end of 2018.

While SA has refocused on FWS priorities, the need for data management, integration, and sharing process remains essential to mission of FWS and SA. This guidance was developed by the SA Data Management Working Group (DMWG) with the leadership of the Architecture Subgroup (ASG).

2.1 Science Catalog Goals and Objectives

2.1.1 Goal 1: Provide an easily searchable repository for SA-funded science projects and products through the SA Science Catalog.

Objective 1.1: Enable efficient and consistent searching and filtering across all SA projects and products

Objective 1.2: Ensure resource managers and partners can continue to find and use LCC information and products for decision-making and conservation action on the landscape.

Objective 1.3: Support users (e.g., resource managers, scientists, auditors, graduate students) in identifying remaining/future science needs and preventing future duplication of work by displaying who and what SA has funded.

Objective 1.4: Ensure the SA Science Catalog instance will remain accessible through a portable design that can be hosted on fws.gov and elsewhere, if needed.

2.1.2 Goal 2: Increase transparency of SA project funding information.

Objective 2.1: Document how FWS and other sources of funding were used to fund science projects.

Objective 2.2: Enable Science Applications and other key partners to show how FWS and other sources of funding were allocated to partners (e.g., states, tribes, federal agencies).

Objective 2.3: Allow data managers to use the Science Catalog as an internal project tracking system.

2.1.3 Goal 3: Demonstrate the impact and value of the investment in landscape-scale conservation.

Objective 3.1: Enable Science Applications and other key partners to show how investments by FWS and other federal programs leveraged additional funding by partners.

Objective 3.2: Demonstrate the depth and breadth of the projects and products produced by SA and the partners and collaborators involved.

2.1.4 Goal 4: Illustrate the link between SA-funded applied science projects and today's conservation challenges.

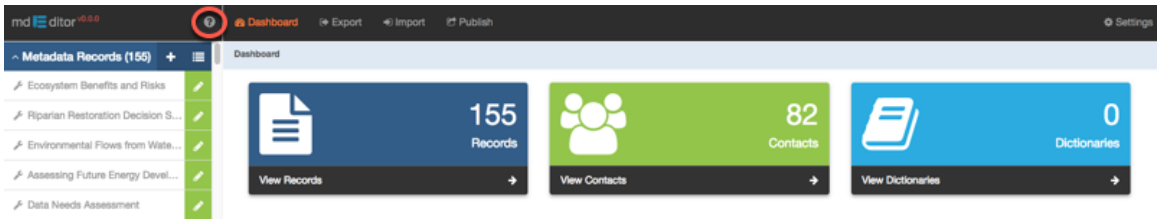
Objective 4.1: Link SA projects and products to overarching Conservation Issues.

Chapter 3

Help

3.1 mdEditor Help

Clicking the question mark icon will open the manual, which can be read online or downloaded as a PDF document.



3.2 Reporting Bugs

If you encounter bugs or errors when using mdEditor, please report them at the mdEditor Issues page on GitHub and specify the mdEditor version you are using (available in Settings). You must have a GitHub account in order to post.

Chapter 4

Glossary

4.1 ADIwg

Alaska Data Integration working group

4.2 Auto-Save

A feature in mdEditor settings that allows information to be automatically saved as it is entered. Consult the **Settings** section of this manual for more information.

4.3 Customization

The ability afforded by open-source code to edit the code of an application (in this case mdEditor) according to the needs of the users.

4.4 FGDC

Federal Geographic Data Committee https://www.fgdc.gov/

4.5 FGDC CSDGM

Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata - FGDC-STD-001-1998 Includes Biological Data Profile https://www.fgdc.gov/metadata/csdgm/

4.6 HTML

HTML stands for Hyper Text Markup Language. It is the standard markup language for creating Web pages. HTML is the 'human-readable' and printable report of the metadata content

4.7 ISO

International Organization for Standardization - ISO is an independent, non-governmental international organization with a membership of 162 national standards bodies.

Through its members, it brings together experts to share knowledge and develop voluntary, consensus-based, market relevant International Standards that support innovation and provide solutions to global challenges.

4.8 ISO 19110

International Standards Organization Geographic Information - Feature Catalogue 19110:2005

ISO 19110 defines the methodology for cataloguing feature types. It may be used as a basis for defining the universe of discourse being modelled in a particular application, or to standardize general aspects of real world features being modelled in more than one application. (International Organization for Standardization (2016). **ISO 19110:2016**. Retrieved from: https://www.iso.org/standard/57303.html)

4.9 ISO 19115-1

Defines the schema required for describing geographic information and services by means of metadata. It provides information about the identification, the extent, the quality, the spatial and temporal aspects, the content, the spatial reference, the portrayal, distribution, and other properties of digital geographic data and services. (International Organization for Standardization (2014). **ISO 19115-1:2014**. Retrieved from: https://www.iso.org/standard/53798.html)

4.10 ISO 19115-2

International Standards Organization Geographic Information - Metadata 19115-2:2009

Extends the existing geographic metadata standard by defining the schema required for describing imagery and gridded data. It provides information about the properties of the measuring equipment used to acquire the data, the geometry of the measuring process employed by the equipment, and the production process used to digitize the raw data. This extension deals with metadata needed to describe the derivation of geographic information from raw data, including the properties of the measuring system, and the numerical methods and computational procedures used in the derivation. The metadata required to address coverage data in general is addressed sufficiently in the general part of ISO 19115. (International Organization for Standardization (2009). ISO 19115-2:2009. Retrieved from: https://www.iso.org/standard/39229.html)

4.11 JSON

Javascript Object Notation, a general purpose format like CSV.

4.12 Keywords

Words used in an information retrieval system to indicate the content of a document.

4.13 localStorage Cache

localStorage Cache allows an application to store data locally, in a user's browser. Storing information on the browser's local storage cache (instead of a normal file cache) means that if you use a different browser to access the mdEditor, it will not show the data you've imported from your first browser. It also means that if you clear your browser's cache, it will generally not clear your mdEditor records. However, depending upon your browser settings (E.g., in Chrome, if the "**cookies and other site data**" option is checked), clearing your browser cache may still clear your mdEditor data.

4.14 mdEditor

Web application for authoring and editing metadata, for both projects and datasets.

4.15 mdEditor File

A mdJSON file created by mdEditor that contains all of the information contained in mdJSON, along with mdEditor settings. This can be exported and shared with collaborators, imported into another record set, or saved to a local workstation as a backup or archival copy.

4.16 mdJSON

ADIwg standard for encoding project and data metadata, based on JavaScript Object Notation (JSON).

4.17 mdJSON File

An mdJSON file that is proprietary to the Metadata toolkit developed by the Alaska Data Integration Working Group (ADIWG), learn more at [https://adiwg.github.io/mdTools/].

4.18 mdTranslator

Open-source Ruby software application for translating between metadata standards. Metadata is input in one of the supported 'reader' formats and output in one of the supported 'writer' formats. Available as Ruby gem or Command-Line-Interface.

4.19 Metadata

Metadata is a set of data that describes and gives information about other data.

4.20 Metadata Repositories

A server where metadata is published to.

4.21 ParentID

Identifier for a folder on a database where records will be stored upon publishing.

4.22 sbJSON

U.S. Geological Survey's standard for documenting records ingested into ScienceBase Catalog. The format used to define the attributes of ScienceBase items.

4.23 ScienceBase

A USGS collaborative scientific data and information management platform used directly by science teams. ScienceBase provides access to aggregated information derived from many data and information domains, including feeds from existing data systems, metadata catalogs, and scientists contributing new and original content. ScienceBase architecture is designed to help science teams and data practitioners centralize their data and information resources to create a foundation needed for their work. ScienceBase, both original software and engineered components, is released as an open source project to promote involvement from the larger scientific programming community both inside and outside the USGS. (USGS (2018). *About ScienceBase.* Retrived from: https://www.sciencebase.gov/about/content/about-sciencebase).

4.24 URI

Uniform Resource Identifier is a string of characters used to identify a resource. A URL is a type of URI.