ServCat Metadata and Dataset Guidelines

Introduction

The Service Catalog, or ServCat, is an online application designed to centralize and preserve information about the Service or used by the U.S. Fish and Wildlife Service (Service). Historically important and management relevant documents- such as reports, management plans, geospatial data, and survey protocols- are now easily accessible. ServCat enables employees to discover information between programs, regions, and refuges, which promotes collaboration and minimizes redundancy across the Service.

Metadata and dataset guidelines were developed by the Natural Resource Program Center (NRPC) for creating and managing metadata to accompany datasets to upload into ServCat. A primary goal of the NRPC's ServCat project is to increase accessibility of natural resource inventory and monitoring information to managers and biologists engaged in planning, management, and decision-making. The following guidelines will ensure long-term usability of the U.S. Fish and Wildlife Service's (FWS) data. ***Please see the Guide to Entering Survey Protocol Frameworks and Guide to Entering Site-Specific Survey Protocols under Guides to Entering Common Documents of the Help section for specific instructions on *PRIMR related datasets and documents*.

Overview

Metadata is defined as data that provides information about one or more aspects of the dataset. Metadata provides context for the data and supports its effective use. Metadata answers the following questions:

- Who developed the data?
- When was the data collected?
- Where was the data collected?
- How was the data collected?
- How was the data processed?
- How are the attributes defined?
- In what formats are the data available?
- How does someone obtain the data?

Metadata allows datasets to be cataloged within Intranet and Internet systems making the data discoverable and available to a broad range of potential users.

Types of Data and Information

Categories of Data	Examples
Raw data	Field forms Photographs GPS files
Processed data	Tabular data files Relational databases GIS layers, maps Species checklists
Documentation	Data collection and analysis protocols Data design documentation Specimen catalogs

Data Quality

It is important to consider the following issues before entering a record into ServCat to ensure the quality, integrity and relevance of information:

Quality: Datasets will be used by FWS staff to inform management decisions about natural resources; it is important that data be accurate and complete.

Clarity: Confusing and cryptic datasets are of little use and can be easily misinterpreted. Datasets need to be accompanied by complete documentation (metadata) so that users will understand the applicability and limitations of the data.

Longevity: Datasets are documented and organized so nothing is lost over time.

Relevancy: Datasets are relevant to ongoing and future monitoring and managing of natural resources.

Availability: A dataset can only be useful to staff if it is discoverable and accessible in a useful form.

Core Components of the Metadata Record

The core components are required when generating a metadata file and are compliant with the Federal Geographic Data Committee (FGDC) standards (http://www.fgdc.gov/index_html). These should be included with any dataset that is uploaded to FWS's ServCat.

- **Metadata Information** includes the language the file was written in, a unique file identifier for the metadata, the standard used to organize the metadata, a point of contact for the metadata, and the date the metadata file was written.
- **Identification Information** is citation-level information about the data includes the title, abstract, purpose for creation, status, keywords (theme and place), and extent (temporal and spatial)
- **Constraints Information** includes the legal and security limitations to data access and use.
- **Data Quality Information** includes the processes and sources used to collect and develop the data and the positional and/or accuracy assessments performed.
- **Maintenance Information** includes the scope and frequency of data updates.
- **Spatial Representation** includes the mechanism(s) used to represent spatial data (grid, point, vector).
- **Reference System Information** includes the reference systems used to represent geographic position and time.
- **Content Information** includes the dataset entities and attributes.
- **Symbology Information** includes the symbols used to represent spatial features.
- **Distribution Information** includes the data distributors and methods for obtaining the data.
- **Metadata Extension Information** includes custom, user-based, changes to the elements, domains or conditionality of the metadata standard.
- **Application Schema Information** about the plan or data models used to structure the data.

ServCat Reference Types

The following reference types are available in ServCat for datasets. Core attributes of these reference types are defined in the ServCat Help Manual Chapter 6.3, Reference Attribute Definitions.

Generic Dataset: Applied to any type of spatial or tabular data that cannot be more specifically defined (e.g., as a tabular dataset). No interpretation of the data is within the dataset.

Examples: Survey results, Access database files, graphical datasets

Geospatial Dataset: Any dataset that is inherently geospatial, including grids, points, lines, polygons and their associated attribute information.

Examples: ESRI shapefile, coverage shapefile, ERIS GRID, ERDAS IMG, geodatabase

Protocol: A rule, guideline, or document which guides how an activity should be performed. The guidelines are not associated with a specific time or point.

Examples: Standards and policies associated with an activity, a plan for quality assurance

Standard Operating Procedure: Detailed written instructions to achieve uniformity of the performance of a specific function. The dataset is among a full description of methodology.

Examples: Step-wise instructions, survey methods

Tabular Dataset: A collection of data. Data are presented in tabular form, where each column represents a particular variable, either as a text string or geography (i.e., points, lines and polygons), and each row corresponds to a given member of the dataset. There is no analysis associated with the table; it contains only raw numbers/variables.

Examples: Survey results, Excel file, printed tables

Unpublished Report: Often used to present the result of an experiment, investigation, or inquiry. An unpublished report includes some level of analysis or discussion.

Example: technical reports with dataset, methods, and conclusion, survey reports

ServCat Dataset Core Attributes

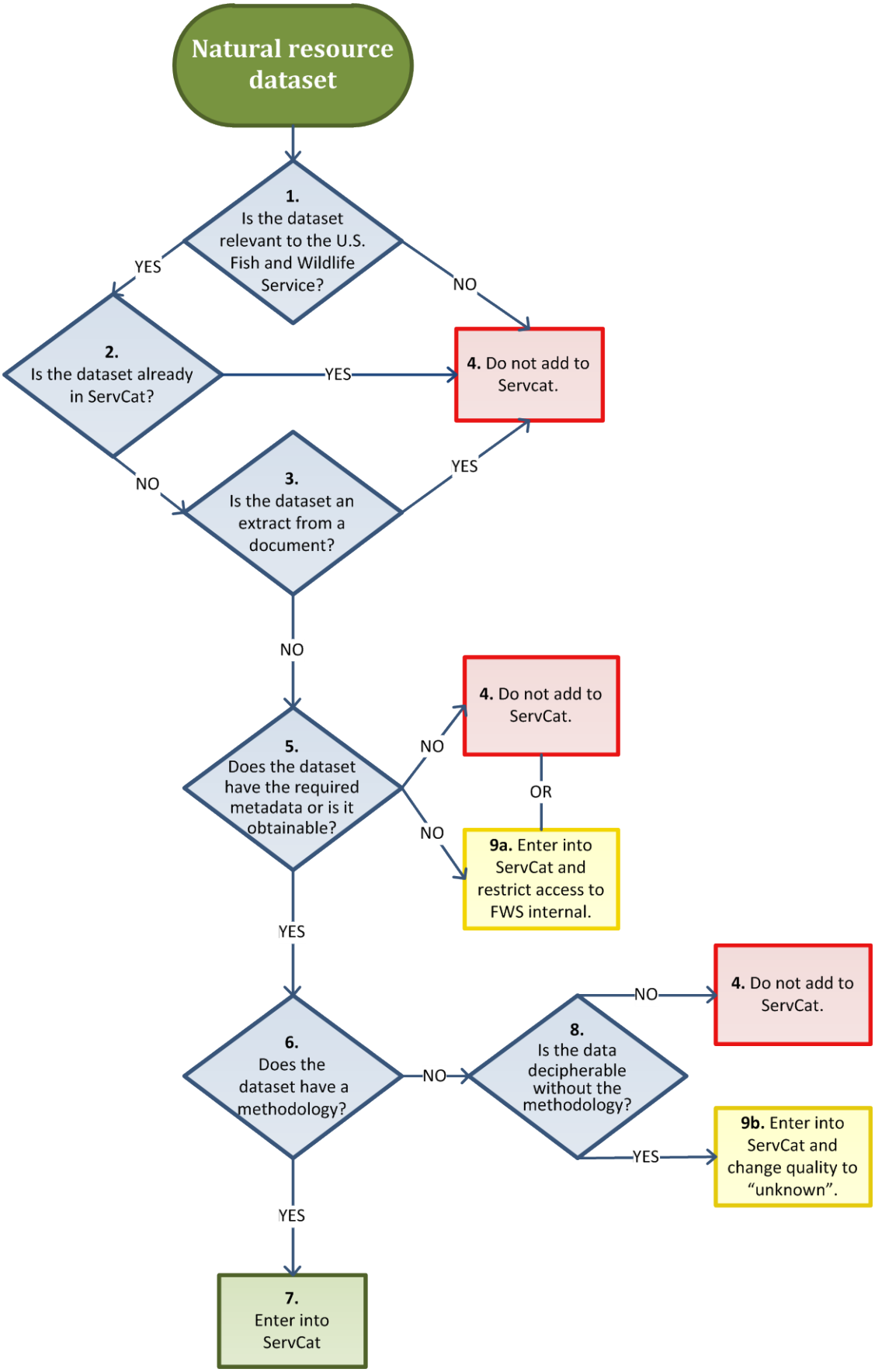
The following core attributes are required when entering a ServCat dataset record.

1. Title
2. Description
3. Date Issued/Released
4. Purpose
5. Contacts
6. Keywords
7. Organization(s)

Dataset Tools

The dataset destination flowchart and dichotomous key on the following pages are aids to determine if or how a dataset should be entered into ServCat. The flowchart is a visual aid and quick reference guide that can be used with the dichotomous key. The dichotomous key provides detailed written explanations for each step shown in the flowchart. Use these tools together or separately to show if or how a dataset should be entered into ServCat.

Dataset Destination Flowchart



Dataset Destination Dichotomous Key

1. Is the dataset relevant to the U.S. Fish and Wildlife Service? For example, Land Conservation Cooperative Reports, tract documents from the FWS Division of Realty, species inventory from the National Wildlife Refuge System.

Yes 2
No 4

2. Is the dataset already in ServCat? Perform a preliminary search to confirm that the dataset *and* related documents are not in ServCat. Related reports, protocols, standard operating procedures, and related datasets may exist in ServCat which contain important metadata.

Yes 4
No 3

3. Is the dataset an extract from another document? Is the dataset stand-alone, or was it extracted from a document which makes the dataset incomplete without the rest of the document?

Yes, the dataset is an extract 4
No, the dataset is stand-alone 5

4. Do not add the dataset to ServCat. The dataset does not meet the requirements to be in ServCat. Maintain dataset in Refuge files as necessary or convenient.

5. Does the dataset have the required metadata? Or is the required metadata obtainable? Documents entered into ServCat require the following metadata: title (or derived title), author (or staff contact), purpose, date, brief description, location, and keywords.

Yes 6
No ..... 4 or 9a

6. Does the dataset have a methodology? Can the questions: who, what, when, where, why, how be answered about the dataset?

Yes 7
No 8

7. Enter the dataset into ServCat. Choose the appropriate reference type based on definitions in the metadata guidelines. If the dataset has a methodology, but it is a separate document, or used by multiple datasets, enter the methodology separately into ServCat. Then, create a relationship between the two documents using the cross-referenced-by relationship, by putting “cross-references the following [Reference Codes: ####,####]” in the notes field. For more instructions on cross-referenced-by relationships see section 6.9.5 of the ServCat Help Manual.

8. Is the data decipherable or interpretable without the methodology? Are units, locations, acronyms, and formulas explained?

Yes 9b
No 4

9a. If desired enter into ServCat and restrict access to FWS internal. Without complete metadata the reference needs to be restricted from public access.
9b. Enter into ServCat and change quality to “unknown”. If the data does not have a methodology, then the quality of the data collection is unknown.