

**NSDI Cooperative Agreements Program
ISO Geospatial Metadata Standards Implementation
Final Project Report**

Date: May 31, 2013

Agreement Number: G12AC20124

Project Title: Promoting ISO Metadata Standards Through Application to Puget Sound Nearshore Data

Final Report

Organization: Northwest Straits Commission, www.nwstraits.org

Principal Investigator: Suzanne Shull, 360-428-1092, sshull@padillabay.gov

Other contact personnel: Caroline Gibson, 360-428-1057, gibson@nwstraits.org

Collaborating Organizations: Washington State Department of Ecology, Terry Stevens, www.padillabay.gov; City of Bellingham, Ann Stark, www.cob.org

Executive Summary

The Northwest Straits Commission (NWSC) has completed the creation and subsequent conversion of metadata for 11 (with 9 pending) of the data layers served on the City of Bellingham server via SoundIQ. SoundIQ is a web mapping application of nearshore data layers, focused in its first phase on Island County, Washington. This online platform was designed to archive, and simplify access to, local nearshore data collected by the Island County Marine Resources Committee (MRC), Island County, and partnering state and federal natural resources agencies. SoundIQ is poised for expansion to other counties with MRCs in northern Puget Sound.

We are pleased to have achieved the goals of the project and hereby describe the process of creating/ converting metadata, as well as some of the associated technical and policy issues. A considerable portion of this process involved researching and understanding ISO standard, xml formatting, metadata acquisition, identification of existing metadata standard, access to software to facilitate the transformations and trouble shooting. There were some challenges, including inconsistencies among datasets and formats provided by the source organizations: Island County MRC, Island County, Washington State Department of Natural Resources (WDNR), Washington State Department of Transportation (WDOT), Washington State Department of Ecology (WDOE), Washington State Department of Health (WDOH), Puget Sound Nearshore Ecosystem Restoration Project (PSNERP), Washington Department of Fish and Wildlife (WDFW), United States Department of Agriculture (USDA), and the National Oceanic and Atmospheric Administration (NOAA).

The Northwest Straits Commission is securing permission to post the Island County MRC metadata in ISO format to the Washington State Geospatial Clearinghouse (WSGC), and effective May 31, 2013 it will be posted to the new SoundIQ Geoportal (see <http://www.iqmap.org/geoportal/catalog/main/home.page>).

Project Narrative

The objective of this project was to demonstrate how - with limited capacity yet extensive Geographic Information System (GIS) and other technical expertise – a regional organization can transform nearshore metadata (i.e. eelgrass, seabird; shoreline armoring) into the recently adopted ISO format. Early on, Suzanne Shull, a board member of the Washington Urban and Regional Information Systems Association, participated in its Washington GIS conference. This turned out to be pivotal first step; in addition to attending numerous informative and relevant presentations at the ESRI International User Conference, Suzanne capitalized on the opportunity to network with GIS experts from other Washington state agencies who shared insight into problems faced by implementing ISO metadata standards.

A progress report on our project was provided at one of the several CSDGM Metadata Working Group meetings attended by Suzanne, generating excellent feedback. It is hoped that this final report will inform state and other non-federal agencies and organizations as they undertake similar efforts.

The following components were fundamental to the Northwest Straits Commission project:

- Training and research: NCDDC metadata training webinars¹ and study of ISO metadata standard workbook², use of NOAA Wiki³, USGS metadata tools, and the NCDDC Online ISO Training FTP resources⁴.
- Inventory and acquisition of metadata for datasets hosted on SoundIQ.
- Transformation of existing CSDGM metadata to ISO metadata standard and validation.
- Creation of ISO metadata for previously undocumented Island MRC datasets.
- Interviews with Washington State Department of Ecology (Ecology) and Washington State Department of Transportation (WDOT) GIS staff to clarify how metadata are currently handled, trouble shoot transformation methods to get those metadata to ISO standard, document the process, and provide this information to State agencies.

Training and research:

ISO can take advantage of xml tools, modern programming platform, customization, and accommodation of multiple standards; therefore the first step to successful translation to ISO format is a basic understanding of xml and uml formatting. This may be one of the biggest obstacles to motivating practitioners' understanding of the advantage of investing in transitioning.

The next step to ISO implementation is to understand exactly what is meant by ISO Content Standard, and how it correlates to CSDGM (APPENDIX A). ISO content can be populated by properly formatted CSDGM metadata, however the formatting requires conversion of the CSDGM metadata to xml format. In addition to the eight presentations covering Introduction

¹ <http://www.ncddc.noaa.gov/metadata-standards/metadata-training/>

² <http://www.ncddc.noaa.gov/metadata-standards/>

³ https://geo-ide.noaa.gov/wiki/index.php?title=Main_Page

⁴ ftp://ftp.ncddc.noaa.gov/pub/Metadata/Online_ISO_Training/Intro_to_ISO/

to Metadata, ISO 101, XML Basics, UML Basics, Tools for ISO Metadata, Writing Metadata Creation Methods, Validation, and Data Discovery offered by the NCDDC metadata training webinars, the NOAA Environmental Data Management Wiki serves as a valuable resource for information on the ISO Standard. Another useful document for background and comparison between CSDGM and ISO is 'Preparing for International Metadata'⁵. Please refer to APPENDIX B for additional resources.

Initially, we were encouraged that ArcGIS 10.1 (ESRI) had released tools to allow users to easily develop ISO metadata in ArcCatalog, and export them to validated ISO metadata xml format. Autosynch capabilities of ArcCatalog, automatic changes to the metadata whenever changes are made to the spatial data (i.e. change of the extent of the datalayer) is an extremely useful feature. However, we ran into problems validating the ISO records exported from ArcCatalog and in trying to identify where the errors were occurring. Ultimately we found it much less expensive to purchase software for working with xml data. ISO format does provide the use of linkages. These are called xlink (see footnote 2 XML-Basics.pptx powerpoint) to maintain links to files that can be updated regularly, separately from the metadata. When the metadata file is ready to archive, the xlinks can be inserted permanently into the metadata document.

The training webinar on Tools for ISO Metadata provided a comprehensive list of available software and associated strengths and weaknesses. After researching the online tools Mermaid and Docucomp we determined that Oxygen and Altova XML Spy would be the most cost effective to streamline transformations, check errors, and especially for validation. We purchased Altova XMLSpy software for \$500, and selected the NOAA Environmental Sensitivity Index data layers as the first set of metadata to convert. This decision was based on the assumption of high probability that we could a) find the metadata, b) identify the file format and content standard, and c) that it would be valid for that standard (i.e. properly formatted and all required CSDGM metadata fields populated). It seemed likely that if errors were found in these data layers that they would be consistent among others from the same (NOAA ESI) source.

Transformation of existing CSDGM metadata: Using the NOAA ESI data.

- 1.) We accessed FGDC CSDGM metadata records from the originator in text (.met) format. The response to our inquiry about their plans to transition to ISO standard was That they currently produce our metadata based on FGDC standards and there will come a time when they will need to produce data in ISO, but not yet. Meanwhile, they welcome our transitioning records for them and getting them posted to the GeoPortal.
- 2.) We submitted the text files to USGS mp metadata parser for validation⁶ of the CSDGM content format and conversion to xml file format. The output is saved to the mp corrected directory as source files for the next step of transforming and validating it to ISO standard. The metadata parser will provide a list of errors if the CSDGM format is

⁶ <http://www.fgdc.gov/metadata/geospatial-metadata-standards#nap>

⁶ <http://geo-nsdi.er.usgs.gov/validation/>

problematic. If the user is trying to create valid CSDGM content, these errors should be addressed before opening the file in XMLSpy for the conversion. If not, it is not usually necessary to correct before successful conversion.

- 3.) The transform file used to convert the data is provided on the NCDCC web site⁷. This file contains the links to the ISO standards based on a code (i.e. gco, gmd, gmi). These codes, called namespaces (Appendix B), are containers providing context and rules for items. This way, a definition of a term may change depending on which namespace is applied.
- 4.) Lessons learned from validation errors with ISO are:
 - a. Vector information transformed over but did not repeat at the correct "level". This is one of those buggy things in the transform that can be easily fixed. The tag "gmd:geometricObjects" should repeat and not "gco:Integer".
 - b. The geometric object type code is blank and we had to fill that in with the appropriate code list selection (Appendix C).
 - c. A validation error common to many datasets is that time period ids must be unique (each one different in the record). This too is a simple fix and not something that the transform really could have automated. The fix is to just number them. For example <gml:TimePeriod gml:id="sourceTemporalExtent2"> and <gml:TimePeriod gml:id="sourceTemporalExtent3"> etc.
 - d. Source Scale has to be a number. The NOAA ESI metadata had 1:24k:1:500k so we replaced that with the text string 'varies'.
 - e. In the USDA metadata we had to delete the repeating 'order instructions'.
- 5.) We reviewed the record and cleaned up empty tags (this is ongoing). Using the cleaned mp output file, we applied CSDGM to ISO 19115-2 Transform using XPath 2.0 January 2013. This is the XML Style sheet that transforms metadata conforming to the Content Standards for Digital Geospatial Metadata of the Federal Geographic Data Committee June 1998 FGDC-STD-001-1999 version to the ISO 19115-2:2009(E). This stylesheet can be applied to CSDGM XML to generate ISO 19115-2 XML.

We also had a detailed entity and attribute information section in the FGDC CSDGM so one needs to do a second transform. FGDC CSDGM to ISO 19110 (Methodology for Feature Cataloguing) Transform using XPath 2.0. This is the XML Style sheet that transforms metadata conforming to the Entity and Attribute Section from the Content Standards for Digital Geospatial Metadata of the Federal Geographic Data Committee June 1998 FGDC-STD-001-1998 version to the ISO 19110. This stylesheet can be applied to FGDC CSDGM XML that contains the entity and attribute section, with detailed information, to generate ISO 19110 feature catalog XML. This stylesheet should only be applied if there is an Entity and Attribute section within the CSDGM XML file.

Once we successfully used Altova XMLSpy to translate existing txt and xml metadata records, we returned to ArcCatalog to test export capabilities. The export tools in ArcGIS 10 work, but

⁷ ftp://ftp.ncddc.noaa.gov/pub/Metadata/Online_ISO_Training/Intro_to_ISO/transforms/csdgm2iso19115-2.xsl

the original metadata associated with the geodatabase were auto deleted. We recommend saving the original geodatabase with metadata before applying the following procedure:

- Use Export Metadata Model, select ARCGIS2CSDGM.xml translator located under Program Files (x86)\ArcGIS\Desktop 10.X\Metadata\Translator\ARCGIS2FGDC.xml and save results to xml format.

Creation of ISO metadata :

The project team spent several hours culling records from existing SoundIQ datasets and contributing individuals for four major projects of the Island MRC. Essentially this was who, what, where, when and how data were collected. The logistical challenges in this basic exercise proved reflective of those for any entity making a similar effort with geospatial data collected and housed by different people and entities over many years. Project managers and volunteers, nearshore survey protocols, and data storage capabilities have changed considerably over the 15 year tenure of the Northwest Straits Initiative. This further validates the need for accurate and current metadata, organized through credible and easily accessible archival systems such as SoundIQ.

The Northwest Straits Commission continues to clean up and update missing fields, and delete empty fields that are not required. We are also adding language to the maintenance note re: conversion from one standard to another and by whom. Lastly, we are contacting metadata providers to inform them that we are posting ISO versions of their metadata online.

Status of Metadata Service

The Northwest Straits Commission now has 11 datasets (with 9 pending plus their Feature Class records totaling 40 files) in validated ISO format
<http://www.iqmap.org/geoportal/catalog/main/home.page>.

Geoportal Welcome, geoportal [Logout](#) [Help](#) [About](#) [Feedback](#)

[HOME](#) [SEARCH](#) [BROWSE](#) [ADMINISTRATION](#)

Search


fish

Records shown from: This Site
Click here to select different site or configure search.

Additional Options
[Clear](#)

WHERE

Anywhere Intersecting Fully within



My Saved Searches

Results 1-1 of 1 record(s)

Expand results [Zoom To Results](#) [Zoom To Searched Area](#)

TEST: Fish Distribution
SoundIQ Fish Distribution map service. Layers: Priority fish distribution; Smelt (spawning); Pacific Herring (holding areas); Pacific Herring (spawning); Sand Lance (spawning)

[Open](#) [Preview](#) [Globe \(.kml\)](#) [ArcGIS \(.nmf\)](#) [ArcGIS \(.lyr\)](#) [Details](#)
[Metadata](#) [Zoom To](#)

See results through REST
API: [GEORSS](#) [ATOM](#) [HTML](#) [FRAGMENT](#) [KML](#) [JSON](#) [CSV](#)

This Geoportal was built using the Geoportal Server. Please read the [Disclaimer](#) and [Privacy](#) or [Contact Us](#).

Next Steps:

The metadata portion of the project is complete, and proved a useful exercise in creating and transforming metadata from our existing geospatial data layers. Presentation of and discussion about the project is planned during Summer and Fall 2013 at Island County MRC, San Juan MRC, and Jefferson MRC meetings, as well as the 15th annual Northwest Straits Initiative conference in November 2013 in Bellingham, WA.

Metadata materials and lessons learned will be shared with colleagues using GIS at Ecology and other state agencies and organizations. We will continue to solicit feedback from the originators of the data on selection of theme keywords. ISO standard topic categories for “Theme Keywords” are limited to: farming, biota, boundaries, climatologyMeteorologyAtmosphere, economy, elevation, environment, geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary, inlandWaters, location, oceans, planningCadastre, society, structure, transportation, utilitiesCommunications. These concatenated terms with their unusual capitalization appear odd because they reflect the UML origins of the standard applications. Metadata editors and other applications can present the terms in a more easily read style but the content should remain as listed above. (excerpted from NAPGuidance_Transition_Final)

Metadata materials and presentation will also be delivered to the Northwest Straits Commission and Ecology, and archived on SoundIQ and the Northwest Straits Initiative website www.nwstraits.org. This report and all related information will remain available for use by members of the Northwest Straits Initiative and other interested parties.

Our regional archive of nearshore metadata will continue to grow as SoundIQ expands to include San Juan, Jefferson, and other counties around Puget Sound.

Feedback on Cooperative Agreements Program

The CAP support staff and online ISO training resources are excellent. Jennifer Carlino and Brigitta Urban-Mathieux scheduled conference calls, facilitated discussions, posted presentations and updates, and provided valuable assistance with information available online and between grant recipients. We would also like to acknowledge the stellar support provided by Jacqueline Mize, NOAA/ NCDDC, whose timely and thorough responses to our many questions were monumental in carrying out this project.

The program would greatly benefit from increased capacity to provide metadata trainers for other key projects among the Puget Sound restoration community, in which extensive geospatial nearshore data of regional and national import are being gathered during any given year.

Where did it make a difference?

The online ISO training resources on Introduction at ISO and Transitioning to ISO are critical. The webinar presentations provided the information necessary to understand and translate the data.

Was the assistance you received sufficient or effective?

Yes, both. Our appreciation for ISO standards and its ability to take advantage of xml tools, programming manipulations, customization, and accommodation of multiple standards has grown significantly since embarking on this project. Demonstrating this functionality when introducing newcomers to ISO should help alleviate some of the resistance.

What would you recommend that the FGDC do differently?

Because the majority of geospatial data are stored in ESRI format, it would be a huge boost to implementing this new standard if ESRI would put more effort into resolving issues concerning metadata formats and transformations. For example, having to purchase a \$500 software package (as we did) in addition to ESRI licensing can be prohibitive.

Are there factors that are missing or are there additional needs that should be considered?

The development of ArcGIS Online and Cloud computing will introduce yet another issue for this standard, in that online datasets do not maintain full metadata descriptions. This will need to be addressed in cooperation with ESRI.

Are there program management concerns that need to be addressed, such as the time frame?

No.

If you were to do the project again, what would you do differently?

Our recommendation to others tackling the conversion of metadata is to take advantage of any introductory training that can be provided on the sequence of necessary steps and available resources. Ideally funds could be allocated to sending metadata trainers to places like Washington state. At the outset of our project much time was spent in simply identifying where to find existing resources, and understanding what ISO standard *is* before beginning to transform.

APPENDIX A

Executive Summary Excerpt from NSDI 'Preparing for International Metadata' (October 20, 2011):

Executive Summary

The Federal Geographic Data Committee has endorsed several new geospatial metadata standards that support the International Organization for Standardization ISO 19115: Geographic information – Metadata standard. Most significant among the endorsed standards are:

- ISO 19115:2003 – the base ISO geospatial metadata standard; a UML representation that specifies the content, conditionality and interrelations of the data documentation elements
- ISO 19139:2007 – an XML implementation of ISO 19115 that specifies metadata record format
- ISO 19115-2:2009 Geographic information – Metadata – Part 2: Extensions for imagery and gridded data – an extension of the base standard that adds element for describing imagery and gridded data
- North American Profile (NAP) of ISO 19115: Geographic information – Metadata – a U.S. and Canada specified implementation of the base standard that increases some conditionality, extends some domains and specifies best practices for populating the metadata record.

A revised version of the base standard, ISO 19115-1: Geographic information – Metadata – Part 1: Fundamentals, is expected to reach final draft status in 2012. The affiliate standards, ISO 19139, ISO 19115-2 and NAP, are expected to be updated to reflect the changes implemented in ISO 19115-1. In addition, the U.S. community has requested that the NAP more fully integrate the geospatial database documentation specified by another affiliate standard, ISO 19110: Geographic information: Feature Catalogue.

In the same way that the existing FGDC Content Standard for Digital Geospatial Metadata (CSDGM) codified geospatial data documentation for the U.S. geospatial data community in 1994, ISO geospatial metadata extends standardization across national borders. Key features include:

- fewer mandatory elements and more optional elements
- extended elements and new elements to capture more specific information
- a hierarchical structure that creates 'packages' of metadata that can be reused and combined to form new metadata records
- support for the documentation of new geospatial data topologies and technologies including geodatabases, web mapping applications, data models, data portals, ontologies, etc.
- suggested best practices for populating metadata elements in a manner that enhances the quality and usefulness of the metadata (NAP feature).

All levels of government, non-governmental organizations (NGOs), and the private sector are strongly encouraged to prepare for the transition to this new, international, approach to metadata by:

- editing existing metadata records to add new content and convert some 'free text' elements to a ISO designated code list using either a manual or automated process,
- testing available transform tools for converting records,
- informing management and technical staff of pending changes and
- planning a strategy for transition to international metadata.

This document provides an overview of ISO 19115 and NAP geospatial metadata and specific guidance on preparing for the transition. Additional guidance documents, similar in scope to the CSDGM Workbook, are currently under development and will include detailed explanations of the record structure, individual elements and best practices. GIS vendors and Federal Agencies are actively developing new applications to transform, create, validate, publish and distribute ISO 19115 and NAP metadata. With these resources in hand, geospatial data and service providers can be fully prepared to update and enhance their geospatial metadata to better support data management, discovery, distribution, application and archive both within, and external to, their organization.

APPENDIX B

Excerpt from Appendix C of the NSDI 'Preparing for International Metadata' (October 20, 2011) ISO/NAP References and Resources

ANSI publication: *North American Profile of ISO 19115:2003 Geographic Information - Metadata*
<http://webstore.ansi.org/RecordDetail.aspx?sku=INCITS+453-2009> (accessed October 14, 2011)

CSDGM to ISO/NAP Conversion xslt
<http://www.ncddc.noaa.gov/metadata-standards/metadata-xml/> (accessed October 14, 2011)

EPA Metadata Editor (EME)
<https://edg.epa.gov/EME/> (accessed October 14, 2011)

FGDC ISO/NAP Webpage
<http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/NAP-Metadata> (accessed October 14, 2011)

FGDC ISO Metadata Editor Review
<http://www.fgdc.gov/metadata/iso-metadata-editor-review> (accessed October 14, 2011)

FGDC Online Training Materials
<http://www.fgdc.gov/training/training-materials>(accessed October 14, 2011)

GeoCommunities Metadata Listserver
<http://lists.geocomm.com/mailman/listinfo/metadata> (accessed October 14, 2011)

ISO Publication: ISO 19110: Geographic Information – Methodology for Feature Cataloging Factsheet
http://www.isotc211.org/Outreach/Overview/Factsheet_19110.pdf (accessed October 14, 2011)

APPENDIX C

Namespace links for definition of terms when using ISO 19115-2 standard

```
<gmi:MI_Metadata xmlns="http://www.isotc211.org/2005/gmi"
xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gmi="http://www.isotc211.org/2005/gmi"
xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns:gmx="http://www.isotc211.org/2005/gmx"
xmlns:gsr="http://www.isotc211.org/2005/gsr"
xmlns:gss="http://www.isotc211.org/2005/gss" xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.isotc211.org/2005/gmi
http://www.ngdc.noaa.gov/metadata/published/xsd/schema.xsd">.
```

APPENDIX D

Sample correction for repetitive error in the Geometric Objects section of the FGDC metadata.

The ISO namespace tag "gmd:geometricObjects" should repeat and not "gco:Integer". Also the geometric object type code was blank so we had to fill that in with the appropriate codelist selection

```
ex:      <gmd:MD_VectorSpatialRepresentation>
          <gmd:geometricObjects>
            <gmd:MD_GeometricObjects>
              <gmd:geometricObjectType>
                <gmd:MD_GeometricObjectTypeCode
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodellists.xml#MD_GeometricObjectTypeCode" codeListValue="" codeSpace="" />
              </gmd:geometricObjectType>
              <gmd:geometricObjectCount>
                <gco:Integer>2851</gco:Integer>
              </gmd:geometricObjectCount>
            </gmd:MD_GeometricObjects>
          </gmd:geometricObjects>
```

</gmd:MD_VectorSpatialRepresentation>