Sufficiency of Self-Documenting Data Exchange Standards

Felimon Gayanilo
Harte Research Institute, Texas A&M University Corpus Christi, Corpus Christi, TX 78412-5869 (Email: fgayanilo@tamu.edu)

Matthew K. Howard
Department of Oceanography, Texas A&M University, College Station, TX 77843-3146 (Email: mkhoward@tamu.edu)

Abstract

Scientific data warehouses aggregate data and information to facilitate data use and reuse. Data provenance metadata describes the data’s origins and any transformations that may have occurred as data move from raw forms to the processed forms available from the warehouse. Such metadata establishes usability and credibility for the data by providing attributions for people, institutions, methods and more. Metadata standards specify what auxiliary pieces of information are required and how to encode them. Many metadata standards exist and it may take more than one to completely express all that needs to be expressed and often, there is considerable overlap and duplication. The International Organization for Standardization (ISO) produced the ISO 191xx series of geographic information standards which are widely-used to document data that has strong geospatial content. The Open Geospatial Consortium (OGC) developed widely-used standards for publishing heterogeneous sensor data of all types for consumption primarily by automated data system. The U.S. Integrated Ocean Observing System (IOOS) has as its goal the automated and largely-unattended transfer of data from sensors to desktops. To support interoperability among various regional data systems and the Federal backbone, IOOS adopted two standards from the OGC Sensor Web Enablement (SWE) framework: the Sensor Observations Service (SOS) web interface and the Observation & Measurements (O&M) encodings. The SOS defines what parameters are available, their temporal coverage, the methods used to collect them, and the data originator. Because much of the IOOS data has a strong geospatial content, efforts are made to automate transformation of OSS ISO to the ISO 191xx family of standards. This poster evaluates sufficiency of the OGC SOS inputs to Digital Object Identifier (DOI) requirements, ISO 191xx and information needed in data landing pages.

Discussion and Conclusion

Exterproic Style Sheet Language (XSL) can be used (see below for an example from NOAA IOOS Program) to map and transform elements of the OGC SOS to ISO 191xx. A qualitative assessment of this mapping using the Gulf of Mexico Research Initiative Information and Data Cooperative’s (GRIIDC) core elements (see table), using the following indicators was performed.

- SOS provide more details than required
- Equivalent SOS data is available
- Alternate SOS data is available but may be insufficient or questionable
- No equivalent SOS data

This exercise reveals the quality of the possible transformation from the self-describing data services OGSS SOS to an ISO 191xx standard. An assessment was also made for mapping requirements to maintain a Digital Object Identifier (DOI) and for required elements of a landing page such as those displayed using GRIIDC’s evolving landing page format (e.g. https://data.gulfresearchinitiative.org/data/RI.x1.340730.014/) and NOAA National Oceanographic Data Center (NODC) dataset landing pages.

The resulting table indicates that self-documenting data service standards, such as the OGC SOS GetCapabilities, is sufficient to be in compliance to ISO standard. However, the table also reveals sections of the ISO standard that are insufficient or may require review and revisions to completely meet the requirements of ISO. Most recently, the ISO has started to introduce changes to the ISO 191 and family and may affect the results presented in this document.

The results of this mapping study also leads to discussions as to how organizations, who serves near real time data via the OGC SOS and other self-documenting standards such as SensorML, should submit data to national data centers for archiving or to data centers such as GRIIDC.

References and Links of Interest

IOOS Working XSL for SOS -> ISO
https://github.com/ioos/metadataTransformations/blob/master/SOS/XS1S2ISO_M.xsl

NOAA EDM SOS-ISO

Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC)
https://data.gulfresearchinitiative.org/

GCOOS SOS GetCapabilities in ISO

ISO 19115 x Open Geospatial Consortium (OGC) SOS Map Checklist and Quality Assessment

<table>
<thead>
<tr>
<th>ISO 19115 Core Elements</th>
<th>SOS 3.1.x Pals</th>
<th>OGC SOS 2.0.0</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>OpenGeoSpatialDate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact</td>
<td>if there is no documented contact, there is no point in using an ISO 19115 (my opinion)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td>if there is no documented contact, there is no point in using an ISO 19115 (my opinion)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>References and Links of Interest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- SOS provide more details than required</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Equivalent SOS data is available</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Alternate SOS data is available but may be insufficient or questionable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- No equivalent SOS data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Acknowledgement

This publication was made possible in part by a grant from BP/The Gulf of Mexico Research Initiative (GoMRI) and in part by NOAA Integrated Ocean Observing System (IOOS) and the Gulf of Mexico Coastal and Ocean Observing System (GCOOS).