

Questions and Answers

Day One, April 16, 2013: ISO Metadata Standards

Overview of the ISO Suite of Standards and How They Relate

- Purchasing the ISO standards

Q: \$30 EACH?

A: The prices for the standards vary. Check <http://webstore.ansi.org>. Check for the reduced priced 'INCITS' versions of ISO 19115 and 19115-2 (same content as ISO versions) - currently \$30.

- Which standard should I use?

Q: With 19115-1 coming, do we wait for it or make the leap now to 19115? Does 19115-1 change anything in 19115 or simply add to it?

A: Don't wait. You can begin exploring ISO 19115 now, test the transforms, and try some of the new functionality. In this way you will be past the learning curve when ISO 19115-1 is finalized. There are some true changes to the standard, Feature Catalogue and Services are more fully integrated, Responsible Party information is reorganized, etc. However, most of the changes are addressed by the transforms NOAA is developing and other changes will simply be a matter of approach.

Q: Why haven't we been told to stop creating CSDGM metadata and start creating ISO metadata? Why have two competing standards?

A: While agencies are strongly encouraged to implement ISO metadata and utilize the new capabilities there is no mandate to stop producing CSDGM metadata. There have been heavy investments in the form of training, tools, workflows, etc. by agencies in CSDGM metadata and this time should be used to explore how we build upon and transition those resources. If you are building a new system, then ISO is the obvious choice.

Organization and Components of the ISO Standard Publications

- XML Namespaces

Q: Can you set a default namespace so you don't have to specify for each element?

A: The output XML must have the namespace specified for each element to be WC3 compliant. If your software allows you to designate a default that it includes in each element, then yes. Make sure that your schema and namespace declarations are in agreement, if you have a mixture of versions validation will be affected.

Detailed Review of ISO 19115

- ISO Content

Q: Does 19115-1 include process steps as well?

A: Yes, it does. In the Lineage section

Q: Citation can be used under several different elements?

A: Yes... citation is a "supporting" section similar to the CSDGM

- Feature Catalogue Metadata

Q: Can Feature Catalogue metadata be uploaded to a Geoportal?

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

- A: Yes, you just have to build in the support for it by referencing out to the appropriate schema.
- Q: It is not clear to me if the feature catalogue is a mandatory requirement in ISO. Having the attributes defined is obviously a main point of metadata.
- A: It is not mandatory since not all geospatial resources will have a Feature Catalogue, e.g., satellite imagery. There is a spot in the content info for a "brief overview" but the 19110 is used to document detailed information. This allows you to create one feature catalogue and point to it as needed from several metadata records rather than repeating that section every time

Metadata Hierarchy

- Documentation of data collections
- Q: How can I use Aggregation to document data series and larger work citations?
- A: The default 'hierarchy level' is 'data set' but it can be changed to 'attribute', 'feature', 'series', etc. You can document the series then document the individual data set and reference the series to which it is associated and provide specific information about that association, e.g. 'cross reference', 'larger work citation', 'part of seamless database', 'stereo pair', etc. More information provided in the Metadata Hierarchy presentation @ http://www.fgdc.gov/metadata/events/ISO_Implementation_Webinar/pdf/FedISOMetaWebinar_Hierarchy.pdf
- Q: Is there an existing example of documentation of a hierarchical data set?
- A: The Census Bureau intends to establish a hierarchical structure for the TIGER/Line data if supported by data.gov

Metadata Components

- Resolving Components

Q: Can you do a partial resolve of a component?
A: Using the tools I (Mize) am familiar with, it's all or nothing.
- XLink

Q: Has there been any discussion of using the more recently developed XInclude instead of XLink?
A: XInclude was last updated in 2006 and XLink in 2010 so it does not seem there is strong interest in XInclude.
Q: Is XLink available outside of the oXygen and XMLspy?
A: Yes, GeoNetwork and ISOMorph utilize it. More at <http://www.w3.org/TR/XLink/>
Q: Can you clarify the relationship between the ISO standards and XLink?
A: XLink is not ISO specific; it is a W3C standard that the ISO standards utilize.
- DocuComp and registries

Q: How much does it cost to run/host Docucomp?
A: Docucomp is a free service provided by NOAA NGDC.
Q: Who registers the components? Is there a means of validating?
A: Yes. Anybody can utilize the components but you must register an account to create and edit components.
Q: How is the widely is the Docucomp service used?
A: Docucomp currently has 41 logins established and most are from NOAA, but there are couple other people from other agencies. It is difficult to assess exactly

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

how many components are in use, but a count of NOAA components and how often they are re-used is interesting: NOAA metadata records reference 660 components that are used 44,045 times.

- Q: Are EPSG codes supported?

A: Yes. European Petroleum Survey Group (<http://www.epsg.org>) maintains a freely available database with standard codes for coordinate systems, datum, spheroids, and units

ISO Metadata Resources

- Training

Q: NOAA has training underway; will it be offered again this summer?

A: Yes, schedule here:

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

- Transform

Q: Are there transformations to go from ISO back to CSDGM?

A: Yes. However, remember that there is new information in ISO that is not in CSDGM, e.g. Services, and that information will be lost.

Q: Where does the CSDGM Biological Profile species information go upon transform to ISO 19115?

A: The transforms include CSDGM Biological Profile to ISO 19115. The Biological Profile was provided as an example of how to extend ISO 19115 and NOAA built a transform based on that example.

Q: Is there an ISO 19115 transform for the CSDGM Remote Sensing extension?

A: No. CSDGM Remote Sensing extension metadata transforms to ISO 19115-2 directly.

Q: How do you extend ISO 19115?

A: You must write a schema to which the extension can be validated.

Q: Are there plans for an ISO 19115 to 19115-1 transform?

A: Yes. NOAA has developed the transform and it is in testing. See

<http://www.ngdc.noaa.gov/metadata/published/xsl/19115to19115-1.xsl>.

- NOAA wiki

Q: What is a metadata record evaluation ‘spiral’?

A: It is a method of breaking an evaluation into multiple iterations or spirals, each of which addresses a small set of critical requirements, and including a feedback cycle with users and other stakeholders prior to the next iteration.

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

Day Two, April 17, 2013: ISO Metadata Implementation

Publishing ISO Metadata

- Policy

Q: Question related to new requirements coming down via OMB Open Data memo - what is the level of effort, technical expertise needed, etc. for metadata managers and/or data stewards to implement the new JSON requirements and the proposed "new metadata standard"? How is this related to FGDC CSDGM and/or ISO 19115 metadata standards and current requirements for the agencies?

A: The Open Data Policy was published on May 9, 2013 and includes a reference to the "GitHub" site where implementation guidance and best practices reside. FGDC is proposing revisions to the guidance that would allow for agencies to include an entry in the proposed JSON data feed for each geospatial metadata collection (WAF or CSW) in lieu of duplicating the metadata entries in that feed. JSON feed management tools and the CKAN harvest will be augmented to support this capability.

- Registration

Q: For former data.gov, etc., publishers, will there be any form of confirmation/validation as to the currency of the registration information? Will those publishers be notified before you include them?

A: There are many users in geodata.gov. GSA will be sending out an email prompt to reset your password and anyone who doesn't do that over time then their login will just be dropped. Rather than starting from scratch, we knew there was at least one owner for each published source. We do have many members of organizations as users (~3000), 250 publishers. So, there are some credentials that will pass forward.

- Security Issues

Q: We are trying to test ArcGIS Online for organizations and haven't received answers about IT security approval. We are experiencing issues within USGS to get approval for ArcGIS Online to open up the site outside of the USGS network. Who did you get approval from?

A: GeoPlatform.gov is using Portal for ArcGIS, which is hosted by GSA and they hold the Certification and Accreditation (C&A) on that system. We are migrating to an ArcGIS Online Organization instance for the Platform, but not for USGS specifically. It will be approved by DOI who intends to leverage certification (in-progress) at EPA and USDA by Q1 FY 2014. Agencies may publish in their own AGOL/Portal instances which can be registered with the Platform – or agency publishers may use the Platform AGOL instance to compose and publish web maps.

Q: What is the timeline of introducing access control - to create fed only spaces, but not for public consumption? Not necessarily classified information.

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

A: Generally, as in the past, the GeoPlatform environment, data.gov environment, communities and the catalog are designed for public access or staging public access. There is an ability to have groups and isolate access to resources within ArcGIS Online group, e.g. Hazards then only people who are members of that group can see anything in Hazards. The issues are how do we expose the description of the resource in the catalog, so as a user if you're logged in associated with Hazards, you can find it. So, it gets more difficult to propagate the access control, so the design criteria are more public/private, but not designed for highly sensitive data. So, private groups exist across the whole Platform. We have had some inquiries from Department of Homeland Security, we'd like to share more sensitive or critical infrastructure data with each other and can you link to our authentication system and it gets complicated and expensive very quickly.

Q: How will the Geospatial Platform handle agency specific IT security requirements (A&A) outside the scope of the USDA certification? Have there been discussions about agencies that have stricter access requirements that may have to be added if our security officers indicate that those requirements don't meet their standards?

A: Metadata and data are owned by the respective agencies and served through systems that are accredited by an agency. The Geospatial Platform website, the AGOL for Organizations instance, and the GeoCloud hosting environment will be accredited by the USGS/FGDC for community use. Individual systems operating in the GeoCloud environment but bearing a different agency domain will need to seek their own (virtual) system accreditation based.

- Harvesting Metadata

Q: In the test catalog, I found a record from the USGS Water Node with a bunch of xml records in a folder that geodata.gov has harvested for years. It's a good example of a bare bones approach whereas you showed examples with services and other features. When I did a search on GIRAS, the old name for some land use land cover data, I got many html links one a web page, a link to the html presentation of metadata and some other ones that were related, but there were other publications and no map services (we don't have those). So, how does that work, does the metadata get harvested from an ISO record or from a link?

A: The metadata are harvested from wherever they sit as xml files so it could be a catalog interface or web accessible folder (WAF). So, the old WRD NSDI Node set up by Ken Lanfear many years ago, that has about 90 different resources and they have links for data download or an information page. Also, there should be a harvest date, but I'm not sure if it's visible. One thing we are discovering in this fresh harvest is that it's one thing to know the metadata are still available, but it's another thing to be sure that the links referenced in the metadata are still live. So, one of the to-do list items in CKAN is to initiate a scheduled process to check links and inform the publisher that your metadata has a dead link here and if you don't remedy it

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

within some period of time, we'll drop the record. So, we want to make sure things are active and accessible and provide a better quality of service.

Q: If we have a homogeneous collection of CSDGM records, some of which use the standard CSDGM and others that incorporate the Biological Data Profile, the shoreline profile, etc., can we validate against multiple schemas or will the collection we need to separate the metadata records and place them in separate WAFs on our end for proper validation and harvest in CKAN or can they be in one WAF?

A: You can use auto detect and it will do its best guess. So, CKAN will do minimal CSDGM, ISO 19139 based on the NOAA public schema validation. Most records won't be showing species taxonomy information and your original metadata will come across. Actually rendered and provided to the user when they click on view full metadata, it isn't taken apart and then put back together as a record. For example, you might have a series of biological metadata records and apply minimal CSDGM validation. It's not going to detect if your taxa is correct, but it will deliver the full metadata record. So, this is more about how rigorous you want the checks to be relative to every record in the collection. You can do an auto-detector a multi directory WAF. This would be one directory for ISO, for CSDGM, and CSDGM BDP and you would have to register three harvest sources in that case. There is an advantage to auto detect, because it looks at the first few lines and then tries to do its best job.

- Apps Catalog

Q: The Raw catalog and Geo Data catalog are going to be moved to CKAN/Data.gov, what about the Apps data (entered through the DMS) like widgets, gadgets, tools, RSS feeds. Will these entries be migrated to CKAN? Our organization has Geospatial metadata for 'applications' and have the same submitted via the APPs section through the DMS using different accounts.

A: The Apps catalog in data.gov. Those records/descriptions won't be submitted any differently than another raw data resource, so it will be the same format for describing an application. You can either describe it through a metadata record such as ISO, which can describe an application, service or dataset or using raw JSON feed which is going to be proposed by OMB soon. Either way, you would get it into the system and be able to filter by resource type, which would retrieve the pure applications (e.g., mobile, web. etc.) The plan for all raw entries including the metadata entries that were done in DMS, there will be a DMS dump in new format with editing tools to every agency. So, each agency will receive a massive JSON file (for example, to NOAA) along with some editing tools that will allow you to edit the content, and then republish it into a new JSON file to post at the site.

Follow up Q: For my specific entries into DMS for applications, I can de-conflict and remove the applications that favor the geospatial metadata that we already have.

A: Yes. We prefer that the community continue to prepare detailed metadata to tell the whole story to end-users rather than reduce the reporting only to the proposed

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

Common Core metadata of data.gov. The fields will be mapped from one to the other, but the Full metadata will be available for the user, whenever created.

- Technical Resources Needed

Q: For some of us that are still learning the more technical aspects of ISO and XML, what is the level of effort needed to implement a JSON feed, and how does it impact us as data stewards and metadata managers that are not at that level of technical expertise?

A: It shouldn't require significant resources. JSON can be converted to and from XML. The latest proposal is that agencies simply create a single JSON entry to describe their geospatial metadata catalog. The JSON feed would then contain all 'raw' metadata entries and a few 'catalog' ones. These will provide all the clues necessary for the CKAN harvester to do its job. GSA/OMB will be distributing a web-based tool to create/edit the JSON feed.

- Search

Q: How will searches work with collection metadata or the parent and child metadata records?

A: As an example, Doug showed the collection metadata from USGS. Here in the USGS Topo Map Collection there is one record, but within that there are 62,353+ records, which are all virtually identical titles except for their location and date. It is important to note that this capability is intended ONLY for homogeneous product collections, like imagery or product files of the same type – it is not appropriate for more heterogeneous collections of varied data. You can create a harvest source and you can identify the URL of the harvest source using Web Accessible Folder (WAF) Collection and when you do that, it asks for the URL. This is the parent record and should contain information true to the entire series such as range of dates, footprint of everyone, abstract, process of collection. In the URL, you put the WAF URL to where that collection can be harvested (only supporting WAF at this time). The WAF needs to be in its own directory with no subdirectories, and then CKAN creates a sub-index, or search facet, within the index. You need an individual WAF folder for each homogeneous collection. As an example, for the Census Bureau we used the 25 themes of TIGER and so there are 25 records for census instead of 200,000 records.

- Links

Q: I see for the record on my publication, so I can add links to related publications. Is this best done by maintaining links in my source metadata instead of editing using a login on the site?

A: The ability to process linked metadata records based on internal hyperlinks is not supported yet in CKAN – other than the parent-child relationship listed above. To assert other associations, use the CKAN tools.

- Hierarchical Structure

Q: Will the ISO hierarchical structure be supported?

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

A: ISO dataset, service, and collection type metadata records are recognized, but each ISO entry must be complete and not rely on multiple parts to be found or displayed. With respect to homogeneous collections, a best practice should be that in your parent record you should include the online resource and resource description and the network resource should point to the WAF. So, you're leaving breadcrumbs in the parent record to the child. You may have multiple links in the parent record to help perform the search, but also to provide more context in the metadata.

- ISO Standards Used

Q: When you retrieve a metadata record from the system, will it be delivered as ISO 19115-2?

A: The metadata is being passed forward in ISO 19115-2. We are converting into -2 from CSDGM based on the NOAA developed XSLT, so that should be for any of the converted records the target. If you submit a 19115-1 metadata record that's what is in there. We are trying to preserve the originality of the record rather than making changes. That's why if we get CSDGM we keep it around as well as its processed version. If we get ISO, we won't change it.

- Development

Q: What is a development spiral?

A: A spiral is an AGILE development term. You will get a series of tasks completed in a few weeks. Every three or four weeks, the CKAN team starts a spiral with items to be worked on, sets prioritization, and gets them completed.

- Metadata Best Practices

Q: How can I make comments on the Metadata best practices document?

A: For the metadata practices document, see

http://www.fgdc.gov/metadata/events/ISO_Implementation_Webinar/index.html.

ISO Metadata Implementation Workflow Model

- Selecting an ISO standard

Q: Agencies seem to have the option of picking and choosing their own ISO standards. It seems there are some core common ones we all need and then agencies can extend as needed. Is this true? If so, what are they?

A: Yes. In Day One we presented a decision route for selecting your standards ([Overview of the ISO Standards and How They Relate](#)). Basically you will start with ISO 19115 and add additional standards for specific data types: imagery and instrumentation (19115-2), GIS and classified data with Feature Catalogs (19110) geospatial services (19119), etc.

ISO Metadata Implementation Case Studies

- USGS Case Study

Q: In USGS, the data is created and published by scientists with no metadata experience. How do we educate others so that the person that signs off on a product or publication has confirmed that we've done a good job with the data and the metadata

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

before it is published?

A: This is a challenge and a change from the way things have been done in the past. Perhaps basic guidance could be drafted that makes clear the value and role of metadata and explicitly outlines how scientist and publishers can assess and use the metadata.

- EPA Case Study

Q: How does the EPA metadata editor interface with GIS tools?

A: EPA Metadata Editor works directly with ArcCatalog in the ESRI desktop products. It is an extension but can also be run standalone used to edit data created outside of ArcGIS. The EPA editor does utilize 'components' to which you can add your own organizational content to auto populate some fields.

- NOAA Case Study

Q: A selling point of ISO is the flexibility with different data types out of the box. What are those examples?

A: The smaller number of required elements and the modular nature allow you to highly customize your content. Some examples include:

- The NOAA Extended Continental Shelf (ECS) uses Rolling Deck to Repository (R2R) technology to develop templates for various data types and to implement standard metadata, vocabularies, documentation rules, and best practices.
- Metadata components reduce maintenance by enabling you to make edits to the source component that may be referenced by multiple metadata records
- Because xml, vs. html or text, is machine readable, it makes automation, transformation and publication easier

- Census Bureau Case Study

Q: A selling point of ISO is the flexibility with different data types out of the box. What are those examples?

A: The flexibility of ISO has a smaller core and then has a lot of flexibility of what you can add to it. See the NOAA case study.

Promoting ISO Metadata Implementation (ROI, etc.)

Q: With ISO metadata adoption, has information access and interoperability improved or are you seeing that down the road.

A: From EPA: we're not there yet, but we are expecting it to improve with the adoption of ISO

Q: What are the selling points to make this transition to ISO? What is the return on investment? What can we present to our agencies?

A: From NOAA: If you have CSDGM records and you publish them to the Geoplatform/Data.gov, the records are transformed to ISO on the fly to facilitate search and discovery. You may want to edit those records and take ownership of that process. Also, the functionality within the standard, such as the ability to use components may be an asset as well because if you can make edits more easily that can be valuable as well in terms of saving time. Maybe consider a pilot project using ISO rather than trying to convert everything all at once. Learn about the functionality and how it can improve what you're doing and then making the switch.

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

ISO Metadata in ESRI

Q: Can anyone on the panel comment on the creation of metadata in ESRI ArcCatalog, specifically in the absence of extensions such as the EPA metadata editor)? Does the ArcGIS 'style' of metadata translate well to an ISO record? Does ArcCatalog provide or expose all the elements of say 19115-2? Would anything be left out? Is post-processing of these records with other xml software necessary to truly take advantage of all that ISO offers? In the panel's expert opinion, are ArcCatalog and the available transforms enough?

A: None of the panelists actively uses ESRI software to produce ISO metadata and do not have the expertise to address the questions. However, we did contact ESRI after the webinar and they provide the following information:

ArcGIS 10 - Esri's ArcGIS Desktop applications store metadata using a proprietary (but publicly documented) schema. The software uses "styles" and XSLT transformations to support import, editing, synchronization, and export of metadata conforming to multiple standards, including FGDC CSDGM, ISO 19115, and others. At 10.x, ArcGIS metadata does not support XLink, although it does allow saving and restoring of "snippets" of content in the editor. ArcGIS Desktop search takes advantage of metadata records associated with datasets and tools. For more information, read: [A new approach for Metadata with ArcGIS 10.](#)

Q: Since ArcGIS doesn't support XLink, which is one of the useful things within ISO because it allows metadata to be more consistent and smaller files. Is it an option to use XSLT transformations to replace the standard contact information instead of XLink/XInclude?

A: From Peter Schweitzer: That would be doable - write an XSLT to make a change as you need it.

ISO Metadata for Non-geospatial Data

Q: The conversion to ISO is more extensive than I thought. If we make the effort, we would need to address both geospatial and non-geospatial data. Is there need for a coordinated effort, perhaps through the data.gov metadata group, to address the transformation of all data to meet ISO metadata standards?

A: From EPA: We use the Dublin Core metadata standard for non-geospatial data and have established some technical specifications and procedures to bring it together in one catalog. As far as other types of metadata, we're working with our Data Standards Branch as well as the documents staff to figure out how do we interact and integrate with each other.

Tools to Transform CSDGM to ISO

Q: Listening to Doug's presentation, when we submit a CSDGM record to data.gov, it will be transformed to ISO. I am interested in seeing what the ISO record looks like. Would we be able to submit CSDGM records and have them transformed because then it becomes a template to help you get started?

A: Agencies are encouraged to explore the use of data.gov to transform their metadata, report back issues and create templates from the output. Also:

- o Altova has a free processor that has more functionality (http://www.altova.com/components_processors.html)

Federal Metadata Coordinator Webinar: ISO Geospatial Metadata Implementation

- ArcGIS Desktop has a cross-walk tool as well and there is a CSDGM to 19115 style sheet and you can add your own
- NOAA's DocuComp service is at
<https://www.ngdc.noaa.gov/docucomp/recordServices>

Creating Interagency Cooperation and Community: Next Steps

Q: How can we establish a 'community' space to keep best practices for the implementation related efforts and share experiences and examples?

A: NOAA shares a lot of information at their NOAA Environmental Data Management Wiki. The FGDC MWG could explore options for building upon the NOAA effort.

Q: How are you going to tie into the FGDC Strategic Plan that is underway?

A: The Standards and Metadata leads within FGDC will follow up.