Bibliographic freedom and the future direction of map cataloging

Terry Reese
Oregon State University
121 Valley Library
Corvallis, OR 97331
tel. 541.737.6384
email: terry.reese@oregonstate.edu

Abstract:

With more and more cartographic data being created and released only as GIS datasets, map catalogers will eventually find themselves moving away from Machine Readable Cataloging (MARC) as their primary bibliographic description schema and moving towards the Federal Geographic Data Committee (FGDC). This move will be necessitated in large part by the needs of the cartographic user community, who are standardized on the FGDC format, and the realization that the FGDC provides a higher level of granularity than other metadata schemas when describing cartographic materials. At the same time, the library community's need for quality MARC records for use within legacy ILS systems will not immediately disappear. This dual need will require map catalogers of the future to straddle two very different user communities and provide bibliographic description in multiple metadata schemas. This paper will discuss the use of metadata crosswalks as one method of bridging this gap between user communities.

Keywords: Metadata interoperability, crosswalks, FGDC

INTRODUCTION

The life of a technical services librarian used to be so simple – items were described in MARC using Anglo-American Cataloging Rules (AACR2) and the Library of Congress Rule Interpretations (LCRIs) with relatively few exceptions. This isn't to say that bibliographic description within this framework was or always is and easy endeavor, but the MARC framework that we used to create and maintain bibliographic data seemed to never change. This is no longer true today. Metadata librarians, as many are called, must navigate through a litany of metadata schemas, each designed to meet the needs of very specialized user communities.

When working with these specialized communities, the question is no longer, how would I describe this resource in MARC, but in rather, in what metadata schema or schemas do I describe it? For those doing map cataloging – this concept is nothing new. In fact, in some ways, I like to think of map catalogers as the very first metadata librarians, in part, because of the cartographic user community. Since the early 1990's, map catalogers have had to have some familiarity with not only the Federal Geographic Data Committee's content standard (FGDC), but also its use and its application within the cartographic user community. And it is because of the FGDC that the profession of cartographic bibliographic description is quickly finding itself nearing a cross-road. With more and more cartographic data being released solely in the form of Geographic Information Systems (GIS) data, it has become increasingly important for map catalogers to move outside of the traditional, MARC-centric bibliographic view of the world and jump feet first into the realm of the FGDC. In fact, for GIS data, the cartographic user community is beginning to demand FGDC data for all available datasets as the current technology continues to find different ways of utilizing metadata within GIS systems. At the same time, the traditional need for MARC bibliographic data will still exist within the library into the foreseeable future.

The challenge for map catalogers then, is the ability to generate bibliographic records regardless of the preferred metadata schema from a single master bibliographic source. This will allow catalogers to create bibliographic records to meet the needs of the current cartographic and library communities as well as meeting the needs of new and future user communities. Likewise, there is a mutual benefit to the map cataloger as bibliographic data provided by an alternative data source, i.e., a GIS data provider, can be utilized to generate the MARC records needed by the library community. To do this, map catalogers must be able to "free" bibliographic data so that it is independent of the metadata schema used to create it; that is, catalogers must now consider how their work can aid in this movement towards metadata interroperability. This paper will discuss one such method that could be used to achieve this level of bibliographic freedom through the use of eXtensible Stylesheet Language Transformations (XSLT) crosswalks. This paper will discuss how crosswalks could be used to move data between the MARC and FGDC formats, as well as look at crosswalking into other metadata formats like Dublin Core or Metadata Object Description Schema (MODS) and some of the issues associated with doing data crosswalking. It will look at some of the current research and explore some of the tools available today that can be used to move between competing metadata schemas. In the end, the purpose of this paper is to provide map catalogers with some of the tools and information that they need to free their bibliographic content to meet the current and future needs of their many user communities.

METADATA CROSSWALKS

There is a common misconception among librarians that metadata crosswalking isn't a technical services issue, but rather, is a systems issue because of the very technical nature of the process. To be sure, metadata crosswalking is a very technically oriented endeavor, but I believe that it should be a technical services endeavor nevertheless. In my humble opinion, metadata crosswalking represents the future of technical services and bibliographic description. As organizations continue to strive for greater interoperability between diverse systems, technical services librarians will be needed to build and manage the metadata crosswalks needed for these systems to interact. As map catalogers, we have the opportunity to lead in this movement, simply because our user community's data needs are not met within the library's traditional MARC-centric approach. Moreover, as data producers begin to take a greater role in metadata creation, technical services librarians will need to be able to find ways to reformat and reuse existing metadata to meet the needs of their own user communities.

So what is metadata crosswalking? Very simply, metadata crosswalks facilitate the mapping of bibliographic data from one metadata schema to a different metadata schema. Their use essentially allows bibliographic data to exist independent of any particular metadata schema and thus, in theory, potentially be available in every metadata schema. This process then, allows catalogers the flexibility to generate their bibliographic description based on the material type and the metadata schema that will provide the greatest descriptive granularity.

Crosswalks are developed by examining the similarities and the differences between differing metadata schemas. At a very basic level, each metadata schema is made up of three core components:

- Syntax: the way in which ideas are formulated.
  i.e., structure of Author's first/last name or structure of taxonomies
- Structure: the actual container used to store the bibliographic data.
  i.e., an xml, rdf or even binary data structure
- Lexicon: the keywords or labels used to bring meaning to the metadata schema
  i.e., MARC uses numeric field tags (1xx for authors, 245 for titles) while the FGDC utilizes textual labels like origin, bounding, etc.

Crosswalks are developed through the examination of these three core components to identify how syntax and lexicon can be reproduced within a new metadata schema. Structure on the other hand, is replaced by the structure of the target schema. This process breaks down data transfer barriers, allowing different systems to reliably share data. Moreover, metadata crosswalks currently represent the best mechanism to promote interoperability between systems. Crosswalks can be used to combine catalogs of data, facilitate cross database searching and enable the assimilation of differing datasets. Crosswalks can also be used to perform data maintenance tasks and reduce metadata creation costs through the promotion of shared content or by moving bibliographic data out of an obsolete metadata schema as new standards continue to emerge. However, there are a number of challenges associated with the use of crosswalks that cannot be overlooked. Several key issues like granularity, consistency and unmappable bibliographic data must be considered when working with a crosswalking solution.

When evaluating the potential success of a metadata crosswalking solution, the granularity of both the source and target schemas must be carefully considered. One must remember that data crosswalks do not add to the bibliographic description of a record – but rather, re-form a set of bibliographic data from one metadata schema to a different metadata schema. This means that the success of a metadata crosswalk will be directly related to the relative granularity of the source and target metadata schema. A metadata crosswalk will be most successful when moving between two metadata schema of equal granularity or moving from a metadata schema with a lot of granularity to a metadata schema with less granularity. Crosswalks that attempt to move data into a more granular metadata schema are more likely to be viewed as less successful. For example, crosswalking data between FGDC and MARC or MARC to FGDC would generally be viewed as a successful crosswalking opportunity even though the FGDC format does tend to be more granular than MARC when dealing with cartographic elements. Since both of these metadata schemas require a great deal of granularity, the potential mapping conflicts remain relatively few.   However, crosswalking from Dublin Core to FGDC data would make little sense, when considering the very low level of granularity offered by the Dublin Core metadata schema. In this case, it is very likely that the data found in the Dublin Core record would be insufficient to generate a valid metadata record within the target metadata schema, limiting the usefulness of the re-formed data.

Another concern related to this issue of granularity is how a particular crosswalk deals with many to one and one to many mappings. Very often, a one to one equivalent will not exist between a source and a target schema, so how a crosswalking system handles these potential conflicting mappings will also greatly determine its level of success. In general, one to many mappings occur when moving to a target schema of greater granularity. For example, when moving bibliographic data between the Dublin Core and MARC metadata schemas, the creator element can map to eight different MARC elements.

Table 1: Dublin Core to MARC

Creator	100 main entry – personal name, 110 main entry – corporate name, 111 main entry – meeting name, 130 main entry – uniform title, 700 added entry – personal name, 710 added entry – corporate name, 711 added entry – meeting name, 720 uncontrolled name 730 added entry – uniform title

In MARC, each 1XX MARC element is used to illustrate a distinction in the type or level of authorship responsibility for a particular piece. When this bibliographic data is moved from Dublin Core to MARC this context data is lost in the translation. In many cases such as these, one will need to create an artificial one to one mapping between the source and target schemas. So in this case, one might map all Dublin Core creator elements to MARC tag 700 (meaning "Added Personal Author"), with the realization that some human intervention would be required to "finish" the crosswalking process for this data. On the other hand, many to one mappings generally occur when moving to a target metadata schema of lesser granularity. For example, when moving bibliographic metadata from FGDC to MARC, one will find that there are basically two FGDC concepts for "source of acquisition geographic reference data" that can be expressed using multiple FGDC elements but that can only be mapped to a single MARC element.

Table 2: FGDC to MARC

6.4.2.1.1 format name, 6.4.2.1.2 format version number, 6.4.2.1.3 format version date, 6.4.2.1.7 transfer size, 6.4.2.2.2.2.1 recording density, 6.4.2.2.2.2.2 recording density units, 6.4.2.2.2.3 recording format	037g

In this case, the MARC element 037 (Source of Acquisition) can potentially be mapped to seven FGDC elements that provide additional information regarding the format characteristics. These characteristics range from the format name and versioning information to information relating to the actual creation of the data. However, because only one of these elements can be mapped into the corresponding MARC element, a choice needs be made regarding how this mapping will take place. Does one crosswalk just the general information in FGDC element 6.4.2.1.1 or does one format these seven elements into a single string and map the combined data into the single MARC element and would this be allowable in the MARC21 metadata schema? In this case, I would argue for doing the former, since it would map the bibliographic data that most closely reflects data usage in MARC subfield 037g.

When working with crosswalks, one will find that rarely will all bibliographic data be mappable from one metadata schema to another. This means that in most cases, there will always be some level of data loss when generating new bibliographic data sets via crosswalking. The trick is to understand why particular data elements cannot be mapped into a targets schema and evaluate if that data can be "lost". Will users need the missing bibliographic elements or is that data used by a particular user community? If the answer is yes, then the crosswalk needs to be modified and the unmappable bibliographic data will need to be accounted for in some fashion. If the answer is no, then the bibliographic data can be "lost" during the translation without inconveniencing the user. The most important thing is to understand why the data loss is occurring and to consider the impact that it will have on the targeted user community.

Moreover, data loss isn't necessarily limited to a loss in bibliographic data. Data loss could also take the form of reduced granularity or the loss of object relationships expressed as part of a metadata schema. An example of this can be seen when crosswalking Encoded Archival Description (EAD) data to MARC. The EAD metadata schema facilitates hierarchical relationships between described objects. Since most metadata schemas cannot express this type of hierarchical structure, the context would be lost even though all of the bibliographic data from the record may be crosswalked into a generated MARC record. While this type of data loss is often times less tangible than the loss of bibliographic data, its absence is no less important.

CURRENT RESEARCH

As mentioned earlier, map catalogers have been thinking about these issues for some time. While doing a literature review, I discovered an article written in January, 2000 by Adam Chandler, Dan Foley and Alaaeldin Hafez of the University of Louisiana Lafayette entitled: "Mapping and Converting Essential Federal Geographic Data Committee (FGDC) Metadata into MARC21 and Dublin Core – Towards an Alternative to the FGDC Clearinghouse". In this paper, the research group examined the feasibility of utilizing FGDC metadata to generate Dublin Core and MARC21 records with the primary interest of investigating alternative methods to the current FGDC clearinghouse system. The group argued that the current FGDC clearinghouse system was flawed having "discover[ed] in 1998 that more than 50% of the queries directed at the National Biological Information Infrastructure [hosted by the University of Louisiana Lafayette] (NBII) node of the FGDC Clearinghouse retrieve zero (0) hits for the user." (Chandler, A, Foley, D., and Hafez, A, 2000). Chandler, et. al worked with OCLC and their now defunct CORC product to develop a web-based tool that could be utilized to generate MARC and Dublin Core metadata records from an existing FGDC record. As Chandler, et. al explain, the tool was written in C and handled a small subset of the 1998 FGDC specifications. The tool marked the first and, from my research, only public attempt to create a reliable crosswalking mechanism between the FGDC and multiple metadata schemas. At the time, only FGDC offered a tool to map data between FGDC and Dublin Core through a utility called mp. This tool however only generated an HTML output with Dublin Core data encoded in the META tags of the document. While Chandler's tool provides a much more robust toolset, it suffers from two obvious limitations. First, the tool was created in C meaning that it could not be easily modified by those individuals lacking some programming background (Chandler, A, Foley, D., and Hafez, A, 2000). Today, metadata crosswalks are primarily being generated using XSLT specifically because they do not require the same level of programming expertise. Moreover, XSLT is quickly becoming the defacto method for processing eXtensive Markup Language (XML) and since most new and emerging metadata schema are XML-based, the use of XSLT as a crosswalking language makes for a good fit. And second, the Chandler tool was designed to handle only these two very specialized crosswalks and allowed only for the translation of data from FGDC to Dublin Core or MARC, not the reverse. As a result, the tool has a very narrow use, but does provide a good demonstration of metadata crosswalking concepts.

In addition to the creation of their tool, Chandler et al. provided the mappings that they used to crosswalk bibliographic data between FGDC, MARC and Dublin Core. These mappings, coupled with the mappings done by Elizabeth Mangan of the Geography and Map Division at the United States Library of Congress provide a very good initial set of mappings between the FGDC and MARC format. These mappings simplify the crosswalking process considerably by providing a tested set of mappings from which other crosswalks can be derived. However, these mappings don't do all of the work for us, since these mappings were developed prior to the FGDC 2.0 specification. So while most of the mappings will remain valid, all the mappings still need to be reviewed and updated to accommodate the new FGDC 2.0 framework.

More recently, J. Norgueras-Iso et. al. examined the feasibility of metadata interoperability within the field of GIS. J. Norgueras-Iso et. al. looked as how metadata stored in ISO 19115 could be utilized by web services which may have different metadata needs. The scenario that the researchers at the University of Zaragoza outlined involved an ISO 19115 structured database which would be queried by a tourism information provider (Norgueras-Iso, J., Zarazaga-Soria, F.J., Lacasta, R. Béjar, Muro-Medrano, P.R., 2004). Within this scenario, Norgueras-Iso et. al. argued that the use of dedicated metadata crosswalks could be used to provide dedicated data streams for particular metadata types. These crosswalks, which would be written using XSLT, would be managed by a metadata crosswalk broker which would examine incoming requests and locate the crosswalk on the system if supported (Norgueras-Iso, J., Zarazaga-Soria, F.J., Lacasta, R. Béjar, Muro-Medrano, P.R., 2004). However, the research team notes that this approach is inherently difficult because each dedicated crosswalk assumes that the developer has a deep understanding of both the source metadata schema and the target metadata schema. Unlike the Chandler tool, Norgueras-Iso's research team proposes the use of XSLT as the crosswalking language for their system to allow for more rapid development of data crosswalks. By using this approach, Norgueras-Iso et. al. database application would provide greater flexibility in adding future metadata schemas. However, like the Chandler tool, their database would also suffer from one of its limitations as well. By requiring direct metadata mappings, Norgueras-Iso et. al. crosswalking tool wouldn't be able to make use of indirect metadata mappings – whereas, metadata is crosswalked through a two known schemas to produce the output of an unknown metadata schema.

Research into crosswalking in general remains a fruitful field of study, in part, because the problem of data interoperability affects nearly every discipline in some way. And while a discussion of current crosswalking theory is really outside of the scope of this particular paper and audience, I feel that I would be remiss if I didn't mention one work that has had a hand in shaping how I look at the problem of data interoperability today. Written in 1998, John Ockerbloom's, "Mediating Among Diverse Data Formats" offers a forward thinking examination of metadata crosswalking and the development of a "smart" metadata crosswalking systems. Ockerbloom suggests a peer network model that utilizes a type-broker system to facilitate crosswalking between heterogeneous systems and offers suggestions on how to build metadata crosswalks that facilitate multiple metadata conversions. Ockerbloom's type broker system is designed around the idea of indirect metadata mapping and crosswalk discovery. According the Ockerbloom's model, a user could request VRA (Visual Resource Association) Core metadata from a native FGDC system if the system supported crosswalks between VRA and a shared crosswalk with FGDC (Ockerbloom, 1998). So for example, if the system supported crosswalks between FGDC=>MARCXML, FGDC=>Dublin Core, FGDC=>MODS as well as a MODS=>VRA crosswalks, then within Ockerbloom's type broker system, an indirect crosswalk could be constructed using FGDC=>MODS=>VRA. Ockerbloom's model persumes the creation of a smart metadata crosswalking system that is able to discover metadata crosswalks from a systems shared knowledge-base. And while I won't spend any more time discussing Ockerbloom's theories within this paper, I would recommend that anyone working on the problems associated with metadata crosswalking should read this work.

BUILDING CROSSWALKS

Before work on the FGDC crosswalks could begin, I had to generate a table that defined how I was going to make MARC data correspond to the FGDC standard and vise versa. As I noted above, both the Chandler et al. article and the initial mappings done by Elizabeth Mangan provided a good starting point – but since each of these mapping predate the FGDC 2.0 specification, even these initial mapping tables needed to be re-mapped to the current FGDC metadata schema. In doing this mapping, I encountered a number of MARC elements that could be mapped to multiple FGDC elements and in those cases I always choose the most generic FGDC element to artificially create a one to one mapping. An example of this can be seen in the example cited earlier in this paper dealing with MARC field 037g. To simplify crosswalking, the 037g was mapped to 6.4.2.1.1, since it most closely matches the type of data that would likely be present within MARC field 037g. In general, I have tried to use the mappings provided by Elizabeth Mangan, with changes made only to simplify the mapping of FGDC data to MARC fields 037 and 342. These two MARC fields store information about source of acquisition (037) and geographic reference data (342). Within the FGDC, these two concepts can be represented over a span of approximately 20 and 39 FGDC elements respectively, so in building my crosswalk, I had to evaluate what FGDC elements best fit the MARC field usage. Otherwise, mappings were only modified when some bibliographic data could be dropped or to update mappings so that they fit with the current FGDC 2.0 specifications.

And while the harmonization of many to one and one to many relationships can be difficult, the most challenging aspects of crosswalking building are likely related to the handling of controlled vocabularies between metadata schemas. This is particulary a problem when moving from one metadata schema to MARC. The difficulties arise when trying to output the proper subfield coding for personal main entry and subject headings. The reason these two classes of data are so difficult relate to the unorthadox subdivisions found within the field data and the lack of support for subdividing subject terms within many XML-based metadata schemas. However, while these two data classes are both problematic, how one handles them is generally quite different.

Personal Main Entry:
When building a crosswalk between FGDC and MARC, two issues are readily apparent. First, most FGDC metadata isn't generated using the same set of controlled vocabulary. So an FGDC metadata record referencing the U.S. Geological Survey probably would not use the controlled form of the name that would appear within a standard MARC record. Moreover, FGDC records offer little information 1.1.1 (Origin) that would identify the dataset creator as an organization or an individual – though individuals are generally constructed in standard citation format (lastname, first). With this information, one could do one of two things: 1) generate a generic translation that maps all 1.1.1 elements to MARC field 700 (added personal main entry) without

any data processing or 2) generate a generic translation that maps all 1.1.1 elements to MARC fields 700 and 710 using the expected data format as a method of determining main entry type. Using the data processing method, one would simply need to generate a template that could process a given set of data using punctuation and data placement to recreating the fields proper MARC coding. For this simple FGDC example, I've chosen to simply map all data as a 700 field, with the understanding that these elements will need to be reviewed if they are used within our ILS. However, for automatic data processing, this mapping would generally be sufficient. For those interested in what the XSLT personal name template may look like, I've included an example in Appendix E.

Subject Headings:

In general, how one handles subject heading will always be a problem when crosswalking data into MARC – though the level of difficulty will vary depending on how the source metadata schema handles subject elements. FGDC for example doesn't utilize subject headings, but instead makes use of short list of defined keywords. As a result, there is very little data processing that can be done on the subject elements – and for records that require authorized headings, some manual interaction will always be necessary to control these headings. For other metadata schemas like EAD for example, LCSH subject headings are generally output using the following format: Geography—Oregon—Archives and are parseable using the same type of method as found in Appendix E. For the purposes of my FGDC=>MARC crosswalk, all subject elements are mapped to 650$a or 651$a so manual review would need to occur on the generated record if controlled headings were desired.

Once the data mappings had been established, it was time to start generating crosswalks. As noted above, the crosswalks that I have developed were created in XSLT. For XML-based metadata schemas like the FGDC 2.0 schema, XSLT provides the most straightforward method of parsing XML data. Using XSLT, one has the ability to extract and rearrange data using XML Path Language (XPATH), eXstensive Stylesheet Language (XSL) script extensions and proprietary XSLT extensions of an individual XSLT parser. XSLT also benefits from having a very low learning curve, allowing individuals with little to no programming experience to quickly start working with the language. This is probably partly due to the fact that an XLST crosswalk is simply an XML file itself which can be opened and read in most web browsers like any other XML file. But more importantly, XSLT is not a niche standard, allowing individuals or organizations to take advantage of expertise that exists outside of one's organization to support or help build these metadata crosswalks. To illustrate the simplicity of an XSLT-based crosswalk, I've provided a small snippet of the MARC to FGDC crosswalk below.

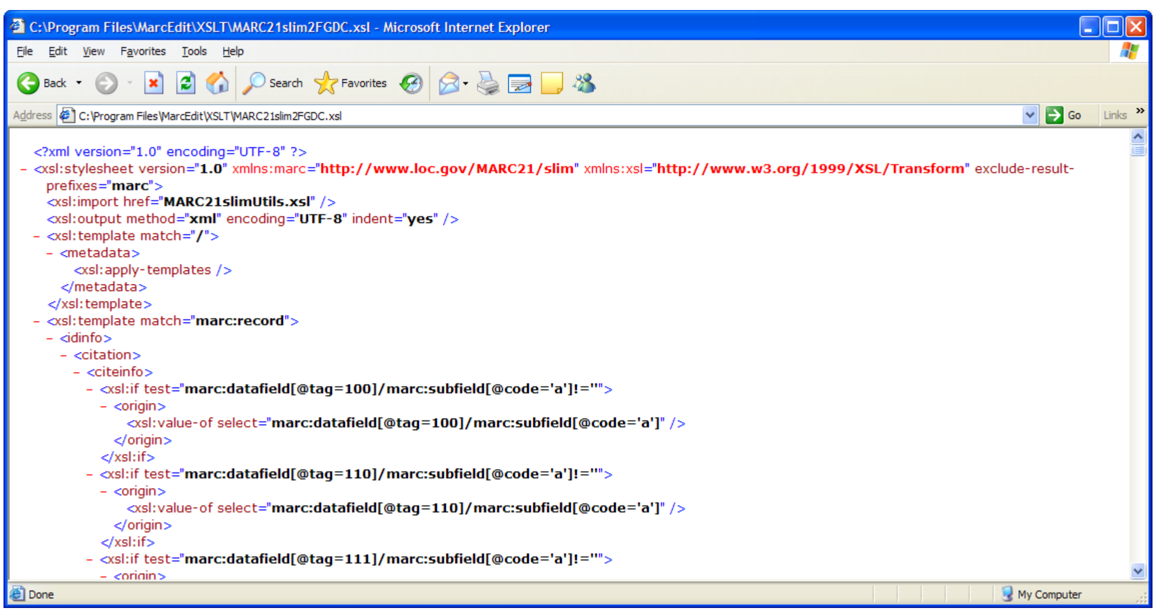


Figure 1: MARC=>FGDC crosswalk viewed in Microsoft Internet Explorer

Moreover, each of my crosswalks has been designed using MARCXML as the controlled translation schema. Exactly what does that mean? It means that two FGDC crosswalks were developed – one crosswalking FGDC data to MARCXML and one crosswalking MARCXML to FGDC. A controlled translation schema is utilized to provide a single translation point that can then be used to map FGDC bibliographic data to multiple metadata formats. For example, if a user wanted to map data from FGDC to MODS, the user would not need to generate a crosswalk defining the direct relationships between FGDC and MODS, but rather, would only need to directly map the relationships between MARCXML and MODS. The use of a controlled translation schema simplifies the crosswalking process, since neither the source or target metadata schemas need to be related to the other, which opens up an infinite number of crosswalking opportunities and simplifies the actual process of mapping bibliographic data between metadata schemas. In the above example, the user would not need to know how the FGDC elements specifically relate to the MODS format, but through the use of MARCXML as the controlled translation schema, the relationships could be inferred. Using MARCXML as a translation schema, the FGDC data would be translated to MARCXML and then crosswalked to MOD3. The reason MARCXML was chosen as the controlled translation schema is that it represents the most granular, general purpose metadata schema that supports the least amount of data loss when moving between specialized metadata formats. Unlike the Chandler et al. tool, which provided a very rigid set of predefined crosswalks, an XSLT-based crosswalk approach provides the user limitless flexibility to create new crosswalks and to customize current crosswalks. In the Chandler et. al tool, users could not add new mappings or modify how mappings were made to match local practice, while using an XSLT-based approach, one could. Finally, utilizing an XSLT-based crosswalking system allows one to choose the tools set that they would like to use in implementing their own crosswalking solution. While the examples that I show in this paper will be done using a tool called MarcEdit, one could just as easily utilize the MARC.pm/MARC::FILE::XML Practical Extraction and Report Language (PERL) modules and the referenced XSLT files and achieve the same results.

PUTTING IT ALL TOGETHER

Once a set of crosswalks has been developed, one will need to locate or build a set of tools that will help facilitate these crosswalks. At Oregon State University, we use a small application that I developed called MarcEdit. MarcEdit is a free, downloadable Windows-based application designed specifically for the purpose of facilitating metadata management. The program was originally developed primarily as a MARC editing software suite, but has changed through the years as the needs in technical services have changed. Currently, MarcEdit includes a set of XML API that allows users to register and execute new metadata crosswalks within the application. The only caveats being that registered metadata crosswalks must be XSLT-based and they must use MARCXML as their base translation layer. Currently, MarcEdit includes five XSLT-base metadata crosswalks that allow metadata to be taken from a one metadata schema to another and back. For the remainder of this paper, I'll be using MarcEdit as my metadata crosswalking tool.

Scenario One: Cataloging GIS datasets without metadata

Unfortunately, map catalogers often receive GIS datasets without any metadata associated with the files. At Oregon State University (OSU), the map cataloger traditionally evaluated the GIS dataset and then generated a single MARC record for the library's Integrated Library System (ILS). This allowed patrons to locate the materials within the library's catalog, but was of little use to actual GIS researchers once the dataset had been found. Since the only available metadata for the resource was in the library catalog, and in a format that wasn't necessarily familiar to most in the GIS community, many in the OSU GIS community needed to generate their own simple set of metadata for the particular dataset. Since this metadata was created for their own personal use, it was rarely shared within the greater GIS community. The ideal solution would have been for the map cataloger to generate both a MARC record and a FGDC metadata record for the particular dataset, thus meeting the needs of both the library and GIS communities. However, many times this type of dual record creation would not be possible because of issues such as time, cost or the cataloger's unfamiliarity with the FGDC format.

In this scenario, the use of a MARC=>FGDC crosswalk can be used to overcome the issues listed above. In this case, the map cataloger must only generate a single MARC record for the dataset, and then can use this single record to generate the other needed recordsets. Moreover, this allows the cataloger to capture the dataset's bibliographic data in the metadata schema with which they are most proficient. Take the following record for example:

```
=LDR  01565nem  2200349Ia 4500
=001  49389519
=003  OCoLC
=005  20020325094243.0
=006  e\\\\\\\a\\l\\0\\\
=007  az\czzzn
=007  co\cg---------
=008  020325s2002\\\\oru\\\\\\\e\\ls\0\\\eng\d
=034  1\$aa$b31680
=040  \\$aORE$cORE
=043  \\$an-us-or
=049  \\$aOREU
=052  \\$aG4294$bC6
=090  \\$aG4294 C6A4 2002$b.C521
=090  \\$aG4294 C6A4 2002$b.C521
=110  1\$aCorvallis (Or.).$bPublic Works Dept.
=245  10$aGeographic information system$h[electronic resource] :$bmapping & GIS services
```

```
/$cCity of Corvallis, Public Works Department.
=246  3\$aMapping and GIS services.
=255  \\$aScale [ca.1:31,680] 1 in. = 2,640 ft.
=260  \\$aCorvallis, Or. :$bPublic Works Dept.,$c[2002]
=300  \\$aOrthophotomap on disk ;$c4 3/4 in.
=500  \\$aHandwritten on disk: City of Corvallis -- GIS; 18 March 2002.
=500  \\$a"This compact disk contains City of Corvallis feature layers that are referenced to their
geographic location."--CD case.
=500  \\$a"City GIS shapefiles and the topographic/planimetric map library are based on North
American Datum (NAD) of 1927, State Plane Coordinates, Oregon North Zone, Clarke 1866
Spheroid, U.S. feet."--CD case.
=538  \\$aSystem requirements: Intel Pentium 133; Windows 98/2000/NT/XP; ArcExplorer
(Environmental Systems Research Institute, Inc.) or other GIS viewing software.
=651  \0$aCorvallis (Or.)$vMap, Topographic.
=650  \0$aOrthophotography$zOregon$zCorvallis.
```

This is a record that was generated for a set of GIS data that the City of Corvallis sends to the OSU Libraries on an annual basis. However, FGDC metadata rarely accompanies these datasets so traditionally the only metadata record that would be created for this dataset would be a MARC record for use within the Library's ILS. So while a record would be in the Library's catalog, no metadata would still accompany the datasets. Using the MARC=>FGDC crosswalk, however, we can generate a brief, but valid, FGDC metadata record that can then be distributed with the GIS data. Using MarcEdit, crosswalking the dataset is as simple as pushing a couple of buttons. To crosswalk this record, a cataloger would simply need to select the file to crosswalk, a save location and select the MARC=>FGDC crosswalk from the known crosswalk list.

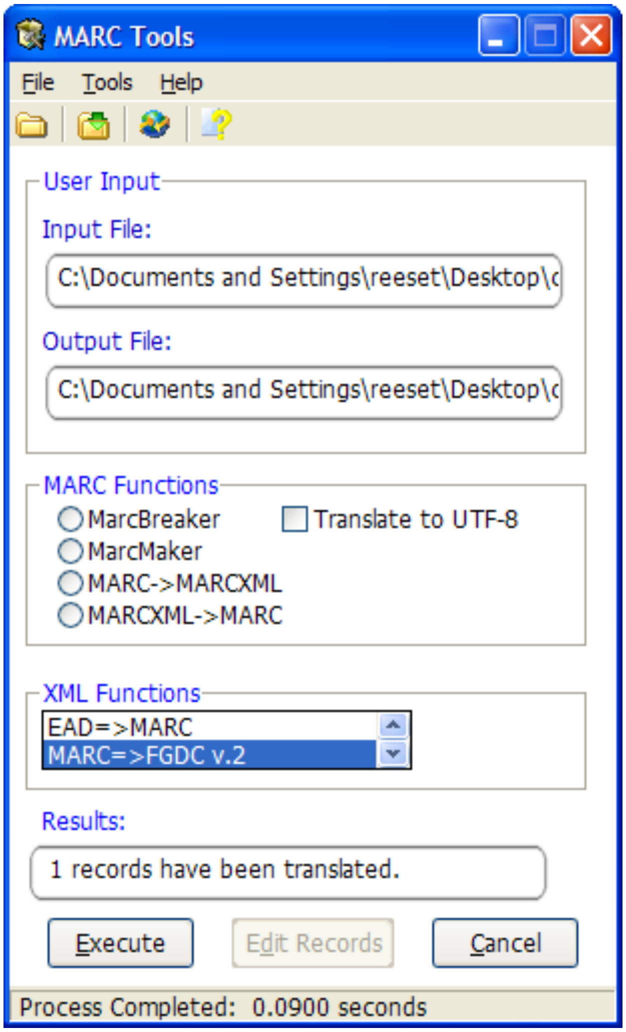
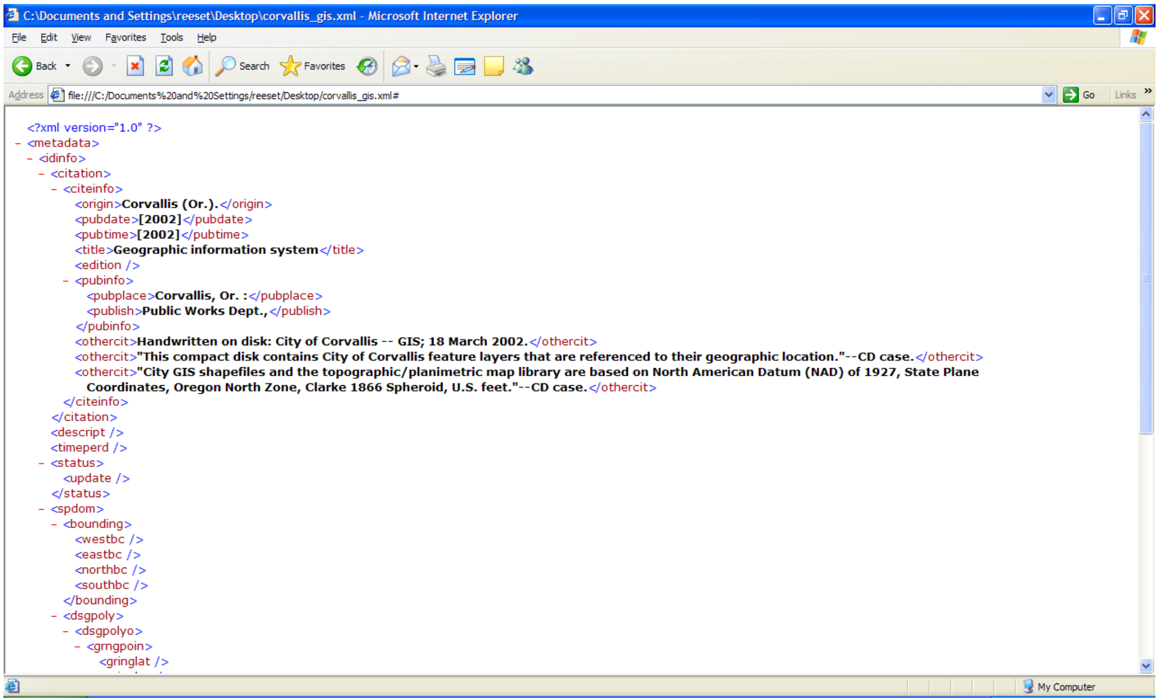


Figure 2: MarcEdit's MARC Tools window

Once the files have been selected, the cataloger simply needs to select the Execute button and the FGDC record is created. MarcEdit facilitates the crosswalking process by automatically taking the bibliographic data out of MARC and putting it into MARCXML. Once the data has been reformated to MARCXML, the MARC=>FGDC XSLT crosswalk is applied to the data. The resulting output is a valid FGDC record. Using the source record noted above, the generated FGDC record would look like the following:




```xml
<?xml version="1.0" ?>
- <metadata>
- <idinfo>
- <citation>
- <citeinfo>
<origin>Corvallis (Or.).</origin>
<pubdate>[2002]</pubdate>
<pubtime>[2002]</pubtime>
<title>Geographic information system</title>
<edition />
- <pubinfo>
<pubplace>Corvallis, Or. :</pubplace>
<publish>Public Works Dept.,</publish>
</pubinfo>
<othercit>Handwritten on disk: City of Corvallis -- GIS; 18 March 2002.</othercit>
<othercit>"This compact disk contains City of Corvallis feature layers that are referenced to their geographic location."--CD case.</othercit>
<othercit>"City GIS shapefiles and the topographic/planimetric map library are based on North American Datum (NAD) of 1927, State Plane Coordinates, Oregon North Zone, Clarke 1866 Spheroid, U.S. feet."--CD case.</othercit>
</citeinfo>
</citation>
<descript />
<timeperd />
- <status>
<update />
</status>
- <spdom>
- <bounding>
<westbc />
<eastbc />
<northbc />
<southbc />
</bounding>
- <dsgpoly>
- <dsgpolyo>
- <grngpoin>
<gringlat />
```

Figure 3: Snippet of the generated FGDC metadata record. A more detailed version of this record can be found in Appendix A.

As one will notice, the record created is very brief, but that is because the MARC record itself is brief – lacking the mathematical data that would generally be found in the 034 or geospatial reference and coordinate data that is generally coded in MARC fields 342 & 343 respectively. However, in less than a second, the map cataloger has provided his or her GIS community with a useable FGDC record, no matter how brief, that can be used within their local GIS systems.

Scenario Two: Cataloging a GIS dataset with FGDC metadata

One of the benefits of the current crop of GIS data systems like ESRI's ARCGIS, is that metadata creation has been built into the workflow of creating GIS datasets. ESRI's ARCGIS for example, requires users to create a minimal FGDC metadata record for a dataset before it can be used within the software. Since GIS data creators should be creating and distributing more and more FGDC metadata with their datasets, it makes sense that map catalogers should try to find a way to make use of the provided FGDC metadata to simplify the creation of metadata records in whatever metadata schema the library currently utilizes without the need to re-key large amounts of data. At OSU for instance, a single FGDC dataset may need metadata records created in Dublin Core, MARC and potentially MODS, so finding a way to utilize the provided FGDC metadata in an automated sort of way could greatly simplify the record creation process.

Using the FGDC metadata record for the dataset, *Modern Average Global Sea-Surface Temperature (Appendix B)*, we can easily generate all three metadata records (MARC, Dublin Core and MODS) from this single FGDC source. Again, we can do this without actually creating crosswalks from FGDC to Dublin Core or MODS because of the FGDC crosswalk map to MARCXML – which is being used as the controlled translation schema. So, as long as we have crosswalks from MARCXML to Dublin Core and MARCXML to MODS, then crosswalking data from FGDC to either of these formats would be possible. Fortunately, MarcEdit includes crosswalks for each of these two metadata schemas. Again, crosswalking from FGDC to these various metadata formats is simply a process of selecting the desired crosswalking option from MarcEdit. Using the crosswalks developed at OSU, our test FGDC record would generate the following MARC record:

```
=LDR  07557nem  2200481Ia 4500
=008  040801u9999\\\\xx\\\\\\\\\u\\\\\\\\und\d
=034  1\$d-180.0$e180.0$f72.0$g-66.0
=037  \\$cnone$gMRLE, $nContact address below: Book and Open-File Report Sales U.S. Geological Survey Denver Federal Center Box 25286 Denver, CO 80225-0286
=100  10$aSchweitzer, Peter N.
=245  10$aModern Average Global Sea-Surface Temperature
=250  \\$a1
=260  \\$aReston, Virginia$bU.S. Geological Survey
=270  \\$aMail Stop 955 National Center U.S. Geological Survey 12201 Sunrise Valley Drive$bReston$cVA$dUSA$e22092$k(703) 648-6533$mpschweitzer@usgs.gov$pPeter N. Schweitzer
=310  \\$cIrregular
=342  \\$bDecimal degrees$c0.01757812$d0.01757812
=352  \\$aRaster$bPixel$d512$e512
```

=500 \\$aThe purpose of this data set is to provide paleoclimate researchers with a tool for estimating the average seasonal variation in sea-surface temperature (SST) throughout the modern world ocean and for estimating the modern monthly and weekly sea-surface temperature at any given oceanic location. It is expected that these data will be compared with temperature estimates derived from geological proxy measures such as faunal census analyses and stable isotopic analyses. The results can then be used to constrain general circulation models of climate change.
=500 \\$aAs ancillary data, the ETOPO5 global gridded elevation and bathymetry data (Edwards, 1989) were interpolated to the resolution of the SST data; the interpolated topographic data are included. The images are provided in three formats: a modified form of run-length encoding (MRLE), Graphics Interchange Format (GIF), and Macintosh PICT format. Also included in the data set are programs that can retrieve seasonal temperature profiles at user-specified locations and that can decompress the data files. These nongraphical SST retrieval programs are provided in versions for UNIX, MS-DOS, and Macintosh computers. Graphical browse utilities are included for users of UNIX with the X Window System, 80386- based PC's, and Macintosh computers.
=500 \\$aSmith, E., 1991, A user's guide to the NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data set produced by the University of Miami/Rosenstiel School of Marine and Atmospheric Science: Distributed by the Distributed Active Archive Center, Jet Propulsion Laboratory, Pasadena, California. 10 p.
=500 \\$aremote-sensing image
=500 \\$aAugust sea-surface temperature in the Western North Atlantic Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/na/gif/m_augna.gif
=500 \\$aAugust sea-surface temperature in the Eastern North Atlantic Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/nae/gif/m_augnae.gif
=500 \\$aAugust sea-surface temperature in the South Atlantic Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/sa/gif/m_augsa.gif
=500 \\$aAugust sea-surface temperature in the Southeastern Atlantic and Southwestern Indian Oceans. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/ag/gif/m_augag.gif
=500 \\$aAugust sea-surface temperature in the Indian Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/io/gif/m_augio.gif
=500 \\$aAugust sea-surface temperature in the Northeast Pacific Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/nep/gif/m_augnep.gif
=500 \\$aAugust sea-surface temperature in the Northwest Pacific Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/nwp/gif/m_augnwp.gif
=500 \\$aAugust sea-surface temperature in the Southeast Pacific Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/sep/gif/m_augsep.gif
=500 \\$aAugust sea-surface temperature in the Southwest Pacific Ocean. This file is available electronically using the URL http://geochange.er.usgs.gov/pub/magsst/cdrom/core/raw/swp/gif/m_augswp.gif
=506 \\$anone
=514 \\$aCalculation of sea-surface temperature values from AVHRR data is described extensively in the references and in the user's guides of the source data products.$dIncluded in the data set are the UNIX shell scripts and makefiles that were used to derive the averages and create

the tables and images. Source code is provided for all programs written by the author.$eIncluded in the data set is a table enumerating the days for which sea-surface temperature data were available in the source material. In general, images were available every week during the time period from 811001 through 891231.$fThe grid provided in the source data (supplied by PO.DAAC) is described in the user's guide of the source product.
=520 \\$aThe data contained in this data set are derived from the NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data (AVHRR MCSST), which are obtainable from the Distributed Active Archive Center at the Jet Propulsion Laboratory (JPL) in Pasadena, Calif. The JPL tapes contain weekly images of SST from October 1981 through December 1990 in nine regions of the world ocean: North Atlantic, Eastern North Atlantic, South Atlantic, Agulhas, Indian, Southeast Pacific, Southwest Pacific, Northeast Pacific, and Northwest Pacific. This data set represents the results of calculations carried out on the NOAA data and also contains the source code of the programs that made the calculations. The objective was to derive the average sea-surface temperature of each month and week throughout the whole 10-year series, meaning, for example, that data from January of each year would be averaged together. The result is 12 monthly and 52 weekly images for each of the oceanic regions. Averaging the images in this way tends to reduce the number of grid cells that lack valid data and to suppress interannual variability.
=538 \\$aThis data set is provided on a single CD-ROM in ISO 9660 format with Macintosh-compatible HFS directory information as well. The data set was developed on a Data General AViiON computer system running DG/UX version 5.4.2R01 (UNIX); additional software products provided on the CD-ROM were developed on a Macintosh IIci and on an IBM-compatible personal computer made by ZEOS, Inc., running an Intel 80486 DX microprocessor.
=552 \\$cSea-surface temperature grid cell value$dThe average temperature, in degrees Celsius, of the ocean surface water in the location indicated by the pixel., Smith, E., 1991, A user's guide to the NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data set produced by the University of Miami/ Rosenstiel School of Marine and Atmospheric Science: Distributed by the Distributed Active Archive Center, Jet Propulsion Laboratory, Pasadena, California. 10 p.
=650 \0$aOceanography
=650 \0$aSea Surface Temperature
=651 \0$aNorth Atlantic Ocean
=651 \0$aSouth Atlantic Ocean
=651 \0$aIndian Ocean
=651 \0$aPacific Ocean
=651 \0$aMediterranean Sea
=653 \\$aPresent
=856 40$zConnect to this dataset online.$uhttp://geochange.er.usgs.gov/pub/magsst/magsst.html

Since FGDC represents a much more granular format, crosswalking bibliographic data into MARC or other metadata formats should result in a much more complete metadata record. As we can see from the generated MARC record, this created record is much more complete than the generated FGDC record from the first scenario largely because of the issue of granularity between the metadata formats. What's more, utilizing the FGDC record allows the cataloger to generate records that contain references to data quality and geospatial reference information – information that could only be provided if an FGDC metadata record was present. And while the record would still need to be edited (i.e., controlling headings, validating punctuation, etc.), the cataloger is able to both save time and create a richer bibliographic record.

So what about the Dublin Core and MODS crosswalks? Again, this is as easy as selecting the crosswalk from the list of known crosswalks and running the source file against the XSLT crosswalks. For those interested, I've included the generated Dublin Core and MODS records in Appendix C and D respectively.

FGDC, ISO AND LOOKING TOWARDS THE FUTURE

While I believe an eventual move to the FGDC and its XML-centric metadata schema are inevitable for map catalogers – I'm not sure when this move will take place. For some institutions, like OSU, the move has happened or will happen in the near future simply because their user communities require FGDC metadata with their record sets. However, I think that the volatility of the current FGDC format will play a role in delaying the move to the FGDC standard for many map and GIS catalogers. In this paper I have provided crosswalks, and analysis for crosswalking data between FGDC v.2 and MARCXML. However, in a few short years much of this work will need to be modified to support the FGDC's current work on the Version 3 framework. Version 3 of the FGDC will be significantly different from Version 2 because for the first time the FGDC metadata schema will be based on the International Organization for Standardization (ISO) 19115 standard. This will be a major revision to the current FGDC framework, including enhanced granularity in the data quality, spatial data, and metadata elements of the FGDC. Currently, the Digital Geographic Information Working Group is taking the lead role in the harmonization efforts and in the creation of a crosswalk between FGDC v.2 and ISO 19115, and as of July 2003 had published Version 4 of the crosswalk specification. This specification remains in discussion as the FGDC continues to deliberate on how and when best to incorporate the ISO 19115 schema with the FGDC framework. So what does this mean for catalogers? In terms of the information presented in this paper, when the ISO 19115 schema is adopted crosswalks will need to be modified to reflect the changes. Fortunately, this won't be difficult to do. In terms of practical application, it means that those doing current GIS cataloging in the United States need to master both the FGDC v.2 while keeping an eye on these future ISO developments.

CONCLUSION

For a long time, librarians seemed to be the only professionals obsessing over the creation of metadata. However, this has all seemed to change as more and more data is being produced digitally. Today, and into the future, more and more metadata will exist for materials and it will be up to librarians to develop ways to capture and utilize this metadata to support its many missions. Using metadata crosswalks, librarians have the ability to utilize the best metadata schema to capture an item's bibliographic data, while still maintaining the ability to crosswalk the data into alternative schemas. And this is really one of the primary goals of building data crosswalks – to allow metadata creators the ability to capture an item's bibliographic description in the metadata schema that will provide the greatest level of granularity. While this paper focused primarily on the capture of FGDC data, there's no practical reason why the concepts discussed couldn't be applied to other metadata schemas.

Appendix A: *MARC=>FGDC crosswalk results*


```
 <?xml version="1.0" ?>
- <metadata>
  - <idinfo>
    - <citation>
      - <citeinfo>
          <origin>Corvallis (Or.).</origin>
          <pubdate>[2002]</pubdate>
          <pubtime>[2002]</pubtime>
          <title>Geographic information system</title>
          <edition />
        - <pubinfo>
            <pubplace>Corvallis, Or. :</pubplace>
            <publish>Public Works Dept.,</publish>
          </pubinfo>
          <othercit>Handwritten on disk: City of Corvallis -- GIS; 18 March
             2002.</othercit>
          <othercit>"This compact disk contains City of Corvallis feature layers that
             are referenced to their geographic location."--CD case.</othercit>
          <othercit>"City GIS shapefiles and the topographic/planimetric map library
             are based on North American Datum (NAD) of 1927, State Plane
             Coordinates, Oregon North Zone, Clarke 1866 Spheroid, U.S. feet."--
             CD case.</othercit>
        </citeinfo>
      </citation>
      <descript />
      <timeperd />
    - <status>
        <update />
      </status>
    - <spdom />
      - <bounding>
          <westbc />
          <eastbc />
          <northbc />
          <southbc />
        </bounding>
      - <dsgpoly>
        - <dsgpolyo>
          - <grngpoin>
```

```
              <gringlat />
              <gringlon />
            </grngpoin>
          </dsgpolyo>
        </dsgpoly>
      </spdom>
    - <keywords>
      - <theme>
          <themekt />
          <themekey>Orthophotography</themekey>
        </theme>
      - <place>
          <placekt />
          <placekey>Corvallis (Or.)</placekey>
        </place>
      </keywords>
      <native>System requirements: Intel Pentium 133; Windows 98/2000/NT/XP;
        ArcExplorer (Environmental Systems Research Institute, Inc.) or other GIS viewing
        software.</native>
      <crossref />
    </idinfo>
  + <dataqual>
  + <spdoinfo>
    <spref />
    <eainfo />
    <distinfo />
  + <metainfo>
</metadata>
```

Appendix B: *Modern Average Global Sea-Surface Temperature [truncted to just show info]*
Full Record URL: http://hdl.handle.net/1957/16


```
<?xml version="1.0" ?>
<metadata>
    <idinfo>
        <citation>
            <citeinfo>
                <origin>Schweitzer, Peter N.</origin>
                <pubdate>1993</pubdate>
                <title>Modern Average Global Sea-Surface Temperature</title>
                <edition>1</edition>
                <geoform>remote-sensing image</geoform>
                <serinfo>
                    <sername>U.S. Geological Survey Digital Data Series</sername>
                    <issue>DDS-10</issue>
                </serinfo>
                <pubinfo>
                    <pubplace>Reston, Virginia</pubplace>
                    <publish>U.S. Geological Survey</publish>
                </pubinfo>
                <onlink>http://geochange.er.usgs.gov/pub/magsst/magsst.html</onlink>
            </citeinfo>
        </citation>
        <descript>
            <abstract>
The data contained in this data set are derived from the NOAA
Advanced Very High Resolution Radiometer Multichannel Sea
Surface Temperature data (AVHRR MCSST), which are obtainable
from the Distributed Active Archive Center at the Jet
Propulsion Laboratory (JPL) in Pasadena, Calif.  The JPL tapes
contain weekly images of SST from October 1981 through
December 1990 in nine regions of the world ocean: North
Atlantic, Eastern North Atlantic, South Atlantic, _gulhas,
Indian, Southeast Pacific, Southwest Pacific, Northeast
Pacific, and Northwest Pacific.

This data set represents the results of calculations carried
out on the NOAA data and also contains the source code of the
programs that made the calculations.  The objective was to
derive the average sea-surface temperature of each month and
```

week throughout the whole 10-year series, meaning, for
example, that data from January of each year would be averaged
together. The result is 12 monthly and 52 weekly images for
each of the oceanic regions. Averaging the images in this way
tends to reduce the number of grid cells that lack valid data
and to suppress interannual variability.
</abstract>
[...idinfo Truncated…]
</idinfo>
[…Trucated…]

Appendix C: *FGDC to Dublin Core [Truncated output]*
Full Record URL: http://hdl.handle.net/1957/16


```xml
<?xml version="1.0" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dc="http://purl.org/dc/elements/1.1/">
	<rdf:Description>
		<dc:title>Modern Average Global Sea-Surface Temperature</dc:title>
		<dc:creator> Schweitzer, Peter N. </dc:creator>
		<dc:type>cartographic</dc:type>
		<dc:publisher>Reston, Virginia U.S. Geological Survey</dc:publisher>
		<dc:language>und</dc:language>
		<dc:description>The data contained in this data set are derived from the NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data (AVHRR MCSST), which are obtainable from the Distributed Active Archive Center at the Jet Propulsion Laboratory (JPL) in Pasadena, Calif. The JPL tapes contain weekly images of SST from October 1981 through December 1990 in nine regions of the world ocean: North Atlantic, Eastern North Atlantic, South Atlantic, Agulhas, Indian, Southeast Pacific, Southwest Pacific, Northeast Pacific, and Northwest Pacific. This data set represents the results of calculations carried out on the NOAA data and also contains the source code of the programs that made the calculations. The objective was to derive the average sea-surface temperature of each month and week throughout the whole 10-year series, meaning, for example, that data from January of each year would be averaged together. The result is 12 monthly and 52 weekly images for each of the oceanic regions. Averaging the images in this way tends to reduce the number of grid cells that lack valid data and to suppress interannual variability.</dc:description>
[...Truncated...]
```

Appendix D: *FGDC=>MODS3 [Truncated output]*
Full Record URL: http://hdl.handle.net/1957/16


```
<?xml version="1.0" ?>
<modsCollection xsi:sch_u109 ?aLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-0.xsd"
        xmlns:xlink="http://www.w3.org/1999/xlink" xmlns="http://www.loc.gov/mods/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
        <mods version="3.0">
                <titleInfo>
                        <title>Mod_u114 ?n Average Global Sea-Surface Temperature</title>
                </titleInfo>
                <nam_cf5  typ_cf2 ="p_u114 ?sonal">
                        <nam_u80 ?art>Schweitzer, Peter N</nam_u80 ?art>
                        <rol_cf2 >
                                <rol_u84 ?erm authority="marcrelator" typ_cf2 ="t_u120
?t">creator</rol_u84 ?erm>
                        </rol_cf2 >
                </nam_cf2 >
                <typ_u79 ?fResource>cartographic</typ_u79 ?fResource>
                <originInfo>
                        <place>
                                <placeTerm typ_cf2 ="code" authority="marccountry">xx\</placeTerm>
                        </place>
                        <place>
                                <placeTerm typ_cf2 ="t_u120 ?t">Reston, Virginia</placeTerm>
                        </place>
                        <publisher>U.S. Geological Survey</publisher>
                        <dat_u73 ?ssued encoding="marc" point="start">9999</dat_u73 ?ssued>
                        <dat_u73 ?ssued encoding="marc" point="end">\\\\</dat_u73 ?ssued>
                        <edition>1</edition>
                        <issuance>monographic</issuance>
                        <frequency></frequency>
                </originInfo>
```

[…Truncated…]

Appendix E: *XSLT Personal Name Template (supports subfield a, d and q placement)*

```
<xsl:template name="persname_template">
        <xsl:param name="string" />
        <xsl:param name="field" />
        <xsl:param name="ind1" />
        <xsl:param name="ind2" />
        <marc:datafield>
                <xsl:attribute name="tag">
                        <xsl:value-of select="$field" />
                </xsl:attribute>
                <xsl:attribute name="ind1">
                        <xsl:value-of select="$ind1" />
                </xsl:attribute>
                <xsl:attribute name="ind2">
                        <xsl:value-of select="$ind2" />
                </xsl:attribute>
                <xsl:if test="contains($string, '(')!=0">
                        <marc:subfield code="a">
                                <xsl:value-of select="substring-before($string, '(')" />
                        </marc:subfield>
                        <xsl:variable name="q_del" select="substring-after($string, '(')" />
                        <xsl:if test="contains($q_del, ',')!=0">
                                <marc:subfield code="q">
                                        <xsl:text>(</xsl:text>
                                        <xsl:value-of select="substring-before($q_del, ',')" />
                                        <xsl:text>, </xsl:text>
                                </marc:subfield>
                        </xsl:if>
                        <xsl:if test="$q_del !='' and contains($q_del, ',')=0">
                                <marc:subfield code="q">
                                        <xsl:text>(</xsl:text>
                                        <xsl:value-of select="$q_del" />
                                </marc:subfield>
                        </xsl:if>
                        <xsl:if test="contains($string, '-')">
                                <xsl:variable name="str1" select="translate(substring-before($string, '-'), ' ', '')" />
                                <xsl:variable name="str2" select="translate(substring-after($string, '-'), ' ', '')" />
                                <marc:subfield code="d"><xsl:value-of select="substring($str1, string-length($str1)-3,4)" />-<xsl:value-of select="substring($str2, 1,4)" />.</marc:subfield>
```

```
                                        </xsl:if>
                              </xsl:if>
                              <xsl:if test="contains($string, '(')=0">
                                        <xsl:variable name="tmp1" select="substring-before($string, '-')" />
                                        <xsl:variable name="tmp2" select="string-length($tmp1)-6" />
                                        <xsl:if test="contains($string, '-')!=0">
                                                  <marc:subfield code="a"><xsl:value-of select="substring($tmp1, 1,
$tmp2)" />, </marc:subfield>
                                        </xsl:if>
                                        <xsl:if test="contains($string, '-')=0">
                                                  <marc:subfield code="a">
                                                            <xsl:value-of select="substring($tmp1, 1, $tmp2)" />
                                                  </marc:subfield>
                                        </xsl:if>
                                        <xsl:if test="contains($string, '-')">
                                                  <xsl:variable name="str1" select="substring-before($string, '-')" />
                                                  <xsl:variable name="str2" select="substring-after($string, '-')" />
                                                  <marc:subfield code="d"><xsl:value-of select="substring($str1, string-
length($str1)-3,4)" />-<xsl:value-of select="substring($str2, 1,4)" />.</marc:subfield>
                                        </xsl:if>
                              </xsl:if>
                    </marc:datafield>
          </xsl:template>
```

NOTES

AACR2 is an acronym that stands for the Anglo-American Cataloging Rules, 2nd edition.

LCRI is an acronym that stands for the Library of Congress Rule Interpretation. This text documents how the U.S. Library of Congress interpretes AACR2 in its cataloging practice.

MARC stands for MAchine Readable Cataloging. The United States Library of Congress currently oversees the development of the MARC21 format. <http://www.loc.gov/marc>

XSLT stands for eXtensible Stylesheet Language Transformation. The XSLT standard, one part of the eXtensible Stylesheet Language (XSL) family, is maintained by the World Wide Web Consortium. http://www.w3.org/Style/XSL/

ISO 19115 defines the International Standards Office schema for describing geographic information and services. Within the United States, the FGDC standard is used by governmental agencies for describing geographic information services. While different, the FGDC standard was created as a subset of the ISO 19115 schema and as of 2005, a process is underway to harmonize the two schemas so that ISO 19115 can be formally adopted and used by FGDC entities.

REFERENCES

Chandler, Adam, Dan Foley and Alaaeldin Hafez. 1999. "DC-7 case study: mapping essential Federal Geographic Data Committee (FGDC) metadata into Dublin Core." <http://cuadra.nwrc.gov/pubs/crosswalk/dc7casestudy.htm > (Last Modified October 4, 1999; Viewed February 16, 2005)

Chandler, Adam, Dan Foley and Alaaeldin Hafez. 2000. "Mapping and Converting Essential Federal Geographic Data Committee (FGDC) Metadata into MARC21 and Dublin Core – Towards an Alternative to the FGDC Clearinghouse." *DLIB*. 6:1. <http://www.dlib.org/dlib/january00/chandler/01chandler.html> (Last Modified January 2000; Viewed February 16, 2005)

Digital Geographic Information Working Group: <http://metadata.dgiwg.org/> (Last Modified: August 12, 2004; Viewed February 16, 2005)

Dublin Core Metadata Initiative: <http://www.dublincore.org/> (Last Modified February 7, 2005; Viewed February 16, 2005)

EE-IR Center FGDC -> MARC21 Metadata Converter: <http://cuadra.nwrc.gov/converter/> (Viewed February 16, 2005)

Encoded Archival Description: <http://www.loc.gov/ead/> (Last Modified: August 12, 2004; Viewed February 16, 2005)

FGDC: <http://www.fgdc.gov> (Last Modified: February 16, 2005; Viewed February 16, 2005)

FGDC (2000). *FGDC Metadata Workbook: version 2.0.* <http://www.fgdc.gov/publications/documents/metadata/workbook_0501_bmk.pdf> (Last Modified: ; Viewed February 16, 2005)

FGDC/ISO Metadata Standard Harmonization: <http://www.fgdc.gov/metadata/whatsnew/fgdciso.html> (Last Modified: January 18, 2005; Viewed February 16, 2005)

FGDC=>MARC21XML and MARC21XML=>FGDC XSLT crosswalks. < http://hdl.handle.net/1957/16 > (Viewed February 16, 2005)

Mangan, Elizabeth U. 1996. "USMARC to FGDC." *Alexandria Digital Library*. <http://www.alexandria.ucsb.edu/public-documents/metadata/marc2fgdc.html> (Last Modified: September 3, 1996; Viewed February 16, 2005)

MARC.PM/MARC::FILE::XML: <http://marcpm.sourceforge.net/> (Viewed February 16, 2005)

MARC21XML Schema: <http://www.loc.gov/standards/marcxml/> (Last Modified February 11, 2005; Viewed February 16, 2005)

MarcEdit: <http://oregonstate.edu/~reeset/marcedit/html> (Last Modified: February 11, 2005; Viewed February 16, 2005)

Metadata Object Description Schema: < http://www.loc.gov/standards/mods/>. (Last Modified: December 29, 2004; Viewed February 16, 2005)

Nogueras-Iso, J. et al. "Metadata standard interoperability: application in the geographic information domain." *Computers, Environment and Urban Systems*. (28):2004. 611-634.

OCLC. 2003. *OCLC Bibliographic Formats and Standards*. Dublin, Ohio: OCLC. <http://www.oclc.org/bibformats/en/> (Last Modified: November 2003; Viewed February 16, 2005)

Ockerbloom, John. 1998. *Mediating among diverse data formats.* Pittsburgh, Pa.: School of Computer Science, Carnegie Mellon University. CMU-CS-98-102. <http://tom.library.upenn.edu/pubs/thesis/> January 1998. (Viewed February 16, 2005)

XML Path Language: <http://www.w3.org/TR/xpath>. (Last Modified: November 16, 1999; Viewed February 16, 2005)