University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, UNL Libraries

Libraries at University of Nebraska-Lincoln

10-1-2002

Distributing non-MARC metadata: The CUGIR metadata sharing project

Adam Chandler
Cornell University, alc28@cornell.edu

Elaine L. Westbrooks
University of Nebraska-Lincoln, elainelw@email.unc.edu

Distributing non-MARC metadata:
The CUGIR metadata sharing project

Adam Chandler [1] and Elaine L. Westbrooks [2]

[1] Central Technical Services, 110c Olin Library, Cornell University, Ithaca, NY 14853-5301, USA
(Corresponding author; email alc28@cornell.edu)

[2] Albert R. Mann Library, Technical Services, Cornell University, Ithaca, NY 14853-5301, USA

Abstract
Technical Service departments throughout libraries today are increasingly in the position to create and maintain non-MARC metadata that may be shared among institutions. The Cornell University Geospatial Information Repository (CUGIR) Metadata Sharing Project is an example of a system that manages heterogeneous metadata such as MARC, Dublin Core, and Content Standard for Digital Geospatial Metadata (CSDGM). The mechanics and policies required to exchange MARC records are mature and well understood. In contrast, non-MARC metadata resource sharing is largely uncharted territory. The authors present and discuss the mechanisms they implemented to enable heterogeneous metadata sharing. Now that CUGIR geospatial metadata records are widely distributed, searchers who might not otherwise have access to geospatial information can more easily discover what they need.

Keywords: metadata, CORC, OAI, MARC, Dublin core, CUGIR

Metadata is not monolithic. Instead, it is helpful to think of metadata as multiple views that can be projected from a single information object [1].

1. Introduction

1.1. MARC network effects

The library literature is filled with definitions and examples about the latest metadata standards that are used in libraries today [2–8]. Two of the best and earliest articles of this type are Stuart Weibel's "Metadata: the foundations of resource description" [9] and "Re-

source description in the digital age" by Jennifer Younger [5]. Another noteworthy contribution is Lorcan Dempsey and Rachel Heery's article, "Metadata: a current view of practice and issues," which lists various metadata standards and categorizes them according to a typology [10]. Ohio State's Magda El-Sherbini reports on her literature search, e-mail survey, and the challenges of mapping Government Information Locator Service (GILS), Dublin Core, Colorado Digital Library Project, and the Content Standard for Digital Geospatial Metadata (CSDGM) to MARC21 [11]. Sherry Vellucci, from the School of Library and Information Science at St. John's University, discusses some of the problems using MARC to describe electronic resources and encourages us to embrace a world of heterogeneous standards [12].

Librarians are already aware of the use of non-MARC metadata in libraries today. From this literature review it is clear that library technical services departments need to proactively get involved in not only creating and maintaining non-MARC metadata, but more importantly, developing the means for widely sharing metadata with libraries that need it for resource discovery and access.

When it comes to discussions of metadata in libraries, Norm Medeiros, the coordinator for Bibliographic and Digital Services at Haverford College, identifies three factions: "XML-only proponents," "MARC traditionalists," and what we call the "modular metadata camp [13]." While each faction has some convincing arguments to offer, the variety of metadata types in use today all have a role to play. XML-only proponents overlook the enormous costs of switching the standard for sharing bibliographic resources from MARC to XML. Carl Shapiro and Hal Varian, from the University of California-Berkeley, describe the economics of these issues in their text, *Information Rules* [14]. "Switching costs" are the cost of changing from one information technology standard to another. When switching costs are very high, the user(s) of the standard may be said to be "locked-in." "Lock-in arises whenever users invest in multiple, complementary, and durable assets specific to a particular information technology system" [15]. Lock-in, they say, occurs at the level of the individual, company, or even society. Shapiro and Varian also discuss "network effects." The value of a network is proportionate to the number of users who utilize the same standard or platform. When one considers that libraries depend on MARC tools, database systems, and bibliographic utilities to share metadata, it is clear that librarians and libraries reap tremendous value from the MARC standard. As of February 2002, OCLC served over 40,000 member institutions sharing over forty-five million unique bibliographic records [16].

Thus, MARC's network effects assure its continued prominence in the future. OCLC's CORC (Cooperative Online Resource Catalog) [Connexion] system's actual use by libraries is a good illustration of the power of MARC network effects. Although CORC offers templates and displays output in both Dublin Core and MARC, evidence suggests that librarians overwhelmingly prefer creating MARC records rather than Dublin Core. Chandra Prabha, from the OCLC Office of Research, provides a good description of this phenomena in her presentation "CORC users: what are they cataloging?" [17]

MARC traditionalists, "who judge AACR2 and MARC as the only credible instruments of bibliographic description," often cite MARC's richness and ubiquity as reasons to embrace it over other metadata standards. However, the traditionalists tend to minimize the fact that MARC is costly, labor-intensive, and requires special technical knowledge about

AACR2 rules and MARC encoding [18]. Although the addition of bibliographic description fields such as Electronic Location and Access (MARC 856) describe networked information, there is continuing dissatisfaction with the flat structure of MARC and the limitations that the flatness put on handling versioning and hierarchical relationships in documents [19].

The "modular metadata camp" embraces aspects of traditional MARC as well as emerging tools, standards, and architectures that can work alongside MARC. First and foremost, the modular metadata perspective recognizes that Dublin Core is a standard to aid resource discovery, not to replace rich, complex, and reliable standards such as MARC or CSDGM. Carl Lagoze, Jessica Milstead, Susan Feldman, and Norm Medeiros are a few who acknowledge MARC and non-MARC schemas in the spirit in which they were conceived [1, 13, 20, & 21]. The CUGIR Metadata Sharing Project adopts this modular metadata approach.

2. The challenge before us

We cannot overlook the tremendous network effects we enjoy by using the MARC standard throughout libraries. One challenge before us, then, is how to link non-MARC metadata such as Dublin Core and CSDGM with metadata stored in MARC/AACR2. A second challenge is designing efficient methods for maintaining related metadata.

In early 2001, the authors planned to improve access to the Cornell University Geospatial Information Repository (CUGIR) [22]. CUGIR, established at the Albert R. Mann Library in 1998, is an online clearinghouse that provides unrestricted access to geospatial data and metadata with special emphasis on those natural features relevant to agriculture, ecology, natural resources, and human-environment interactions in New York state [23]. CUGIR is a node of the Federal Geographic Data Committee (FGDC) Clearinghouse and is one of 250 spatial data servers that have digital geographic data primarily used for geographic information systems (GIS) [24]. CUGIR metadata are stored as text, HTML, SGML and XML files. Online users may view any metadata record in any format of their choice.

Converting CUGIR's metadata into MARC grants students, faculty, and staff access to geospatial data via the Cornell University Library OPAC and beyond. The co-existence of geospatial metadata with traditional resources in the OPAC is essential to making geospatial data files known and accessible beyond the narrow world of GIS (geospatial information systems). What started as a "simple" geospatial information access project, became known as the "CUGIR Metadata Sharing Project" [25].

In an effort to make geospatial information resources more accessible to users who might not otherwise encounter them, the team [26] implemented the CUGIR Metadata Sharing Project to move geospatial metadata (CSDGM) into MARC systems (CORC and OPAC) as well as non-MARC metadata systems (CORC and the Open Archives Initiative). The CUGIR Metadata Sharing Project was designed to address two access-related problems. The first problem was that individual metadata records are not accessible in CUGIR because they reside in the "Deep Web." In the white paper "The Deep Web: Surfacing Hidden Value," Michael Bergman of BrightPlanet Inc. defines Deep-Web as being:

> … content that resides in searchable databases, the results from which can only be discovered by a direct query. Without the directed query, the database does

not publish the result. When queried, deep Web sites post their results as dynamic Web pages in real-time. Though these dynamic pages have a unique URL address that allows them to be retrieved again later, they are not persistent [27].

The second problem was due to the fact that Cornell University Library's core constituency, faculty, students, and staff are not fully utilizing CUGIR's geospatial resources. The log files indicate that most users come into CUGIR with commercial and governmental IP addresses [28]. As a result of this finding, we set out to increase access to CUGIR by making it easier for core constituents to discover geospatial information via the OPAC; OCLC's CORC and WorldCat databases; and the Open Archives Initiative, or OAI.

CORC is an example of a working system that can manage Dublin Core, XML, and MARC. The CORC system provides functionality needed to create and store Dublin Core in XML as well as MARC. CORC, launched in 1999, is an online service that provides a Web-accessible shared database and automated tools to help libraries manage and provide intellectual access to Web materials. CORC includes descriptions of Web resources that may be cataloged or displayed in either MARC or Dublin Core. David Allen, Karen Calhoun, Ingrid Hsieh-Yee, and Michael Smith are just a few authors who have described CORC and documented how it is used in libraries [29–31].

OAI develops and promotes interoperability standards aiming to facilitate the efficient dissemination of content [32]. There are two standards that make up OAI: the Metadata Harvesting Protocol (MHP) and Dublin Core. The former standard is HTTP-based and the latter is transported in XML. Since OAI became public in January 2001, it has attracted a great deal of attention due to its interoperability, openness, and potential. OAI is an important component of the CUGIR Metadata Sharing Project because it allows members of the OAI community to harvest CUGIR metadata in DC and thereby increases access to CUGIR.

3. Project methods and mechanisms

This article recounts, in a non-technical way, how the project team devised an open digital library architecture using existing standards to address the maintenance of related MARC and non-MARC metadata. We plan to report on the digital library architecture of the project elsewhere.

In an interview appearing in *OCLC Systems & Services,* Andy Powell, creator of the DC-Dot metadata generator, [33] advises library professionals who prepare metadata projects to re-use existing metadata elements sets, build open systems, and use open standards [34]. Pursuant to Powell's advice, the CUGIR Metadata Sharing Project did not attempt to reinvent the wheel. Rather, open standards (DC, XML, RDF—the Resource Description Framework, HTTP, OAI) were used to build an open, simple, and well-documented system.

An effective method for increasing the use of CUGIR data are to convert its geospatial metadata to MARC and then add the newly formed records to the OPAC and WorldCat. The second way to increase access involves making CUGIR a metadata provider through the rapidly growing OAI community. The decision to increase access forced the project team to figure out an effective mechanism for centralizing and organizing digital items that are accessible from multiple systems. The solution is to create one persistent URL that can be embedded in different metadata records.

Figure 1. CSDGM metadata record in CUGIR. Patrons are lead to this complete record from surrogates in the local library OPAC, WorldCat (via FirstSearch) or OAI service providers. From this CSDGM record, patrons may download the actual dataset.

All data files are cataloged in accordance with Federal Geographic Data Committee (FGDC) standards of the CSDGM geospatial metadata and made available in widely used geospatial data formats. Figure 1 is a brief excerpt from the CUGIR record: "Minor Civil Divisions, Albany County." From the field "Online Linkage," users can easily download the binary data files. As we will demonstrate, the persistent URL mechanism we implemented to support sharing of CUGIR's geospatial metadata allows searchers to eventually get to the record shown in Figure 1, whether they are searching the Cornell University Library OPAC, OCLC's FirstSearch, or OAI-harvested metadata.

We opted to create MARC records from CUGIR SGML files using an online converter [35]. The online converter was designed to map or transfer the maximum amount of information possible from the CSDGM to MARC. The newly formed MARC records were imported into CORC and enhanced with standard headings and classification numbers by a student employee trained to "catalog" in CORC. The records were cataloged according to local and national practices at the PCC (Program for Cooperative Cataloging) core level [36]. CORC proved to be a valuable time saving tool for authority control and the addition of access points, which significantly adds value to the MARC and corresponding DC records.

After importing the MARC records into CORC, we used CORC to convert them to Dublin Core records. The CORC Dublin Core versions were then imported into a local relational database system though a custom locally created Web interface. The creation of the MARC and Dublin Core records is one of the easiest and quickest steps in the CUGIR Metadata Sharing Project. Discovering a simple method to create, maintain, and centralize a persistent link for the 856 field (Electronic Location and Access) of the MARC record proved to be the challenge.

Often, discussions about persistence begin with the fundamental and obvious problems inherent to Uniform Resource Locations (URLs). Sandy Payette of the Digital Library Research Group at Cornell University characterizes the difficulties as follows: "[T]he URL is simply an address masquerading as an identifier, and relying on … [the URL] to identify a unique digital resource is analogous to using a home address in place of a social security number" [37]. This perspective on object identifier persistence starts from the premise that all digital objects should have a Uniform Resource Name (URN), like Handle or Digital Object Identifier (DOI) systems, rather than a URL. That is, digital objects need an identity that will remain stable while the location of the physical object changes over time. Uniform Resource Names separate the identity of an object from a physical network location. Another way to create digital object persistence is through the use of a local Persistent Uniform Resource Location (PURL) resolver. A PURL resolver is technology developed by OCLC in the mid-1990s to create more stable links to digital objects [38]. Numerous libraries use the OCLC PURL resolver software to create persistent links to electronic resources.

For the CUGIR project, we sought out a solution that is more flexible and dynamic than URNs or PURLs. A digital library identifier resolver that de-couples the identity of the object from the location of the object, but which provides more functionality, is needed. The solution for the CUGIR Metadata Sharing Project is a complex resolver that we refer to as a "bucket."

The term "bucket" is borrowed from Michael Nelson's research on digital library architecture [39]. The bucket is the glue that holds the CUGIR Metadata Sharing Project together because it enables identifier persistence across the heterogeneous metadata surrogates of CUGIR records that are distributed in the OPAC, OCLC's WorldCat, CORC, and OAI. Nelson's "bucket" solution is a way of grouping everything (i.e., metadata) into a common place, then building a small container around it. Implementing Nelson's complete bucket architecture is not practical for the CUGIR Metadata Sharing Project at this time because it would require a new directory structure scheme. Presently, we do not have the option of moving the CUGIR metadata and data files into a new directory. Furthermore, the software needed to create Nelson's buckets is incomplete. Therefore, we arrived at a system that borrows the simplicity of a PURL for resolving identifiers, but which adds the capacity of linking related objects together into a coherent framework, like Nelson's buckets. We refer to our system as the CUGIR Simple Bucket Digital Object Model, as illustrated in Figure 2.

The CUGIR digital object has three components. Working from left to right in Figure 2, the "CUGIR" box represents the original FGDC metadata files, in all four manifestations (SGML, HTML, XML, and ASCII text) plus the data file described by the metadata. The bucket in the middle binds the whole digital object together. First and foremost, it contains the location of the different CUGIR metadata files. It does not need to contain the lo-
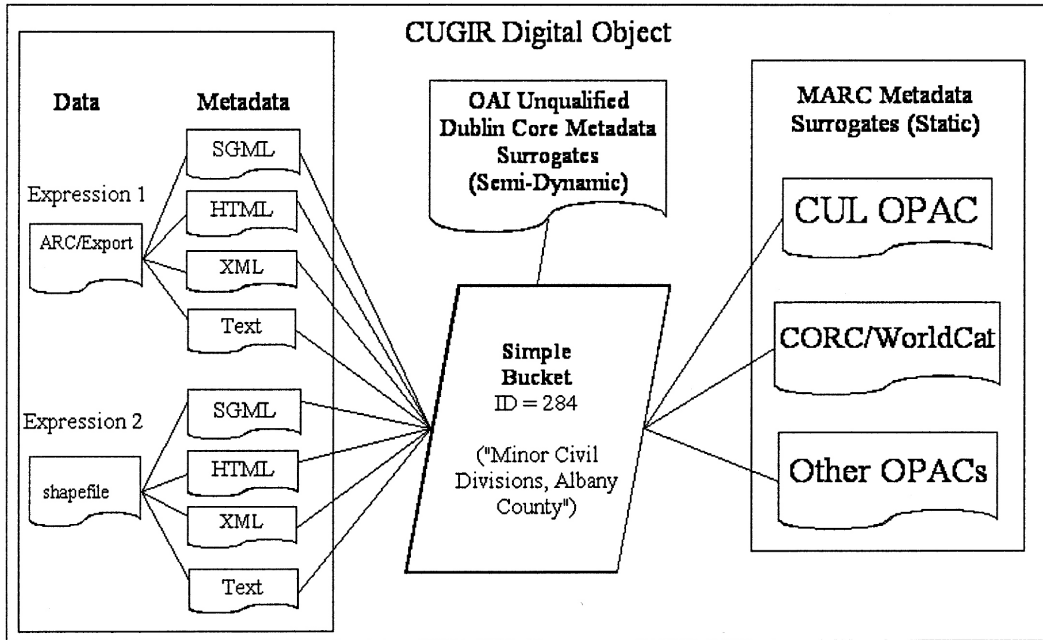
Figure 2. CUGIR Simple Bucket Digital Object Model.

cation of the data file, since the data file location is maintained within the CUGIR meta-data file. The relationship between the bucket and the four CUGIR metadata files is the key feature of this system. The bucket location on the network is by design stable and per-sistent, like a URN or PURL, thus creating the possibility of linking to CUGIR metadata from MARC surrogates placed in the Cornell Library OPAC, CORC, and OCLC's World-Cat. Dublin Core surrogates may be harvested from our system through the OAI meta-data harvesting protocol, further increasing access to the CUGIR metadata and data files. These records also point to the bucket.

Figure 3 is an example of a bucket that is displaying the Dublin Core for "Minor Civil Divisions, Albany County." From this bucket the user has access to the full geospatial metadata record (HTML and SGML) in CUGIR, the MARC record, as well as access to buckets that are related to "Minor Civil Divisions, Albany County." Figure 4 demon-strates how users searching for "Minor Civil Divisions, Albany County" in OCLC World-Cat will arrive at the same bucket when they click on the hyperlink in the WorldCat re-cord. The bucket leads patrons to the CUGIR descriptive metadata (as seen in Figure 1).

4. Project outcomes

The CUGIR Metadata Sharing Project was successful in reaching its overall goal: in-creasing access. Currently, data are being collected to track how many users are gaining

Address [icon] http://cugir2.mannlib.cornell.edu/buckets/Display.jsp?id=284   ▼   ↻ Go

title **Minor Civil Divisions, Albany County** ▲

Arc Info :

CSDGM HTML	http://cugir.mannlib.cornell.edu//Isite/CUGIR_METADATA/001/001mca.html
CSDGM SGML	http://cugir.mannlib.cornell.edu//Isite/CUGIR_METADATA/001/001mca.sgml
MARC	http://cugir2.mannlib.cornell.edu/userDir/bw47-cornell.edu/284.dat
topical : subject	Administrative and political divisions.
geographic : subject	Albany County (N.Y.)--Administrative and political divisions--Maps.
class : subject	G3803.A3.F7
coverage	n-us-ny
coverage	Scale not given Universal Transverse Mercator
type	Map data
type	[electronic resource].
language	eng
date	1998
identifier	http://cugir2.mannlib.cornell.edu/buckets/Display.jsp?id=284
publisher	Washington, DC :
contributor	United State. Bureau of the Census.
contributor	United State. Dept. of Commerce
publisher	Bureau of the Census,
coverage	a W0734037 W0721554 N0424922 N0422425

Related Links :

Shape File   http://cugir2.mannlib.cornell.edu/buckets/Display.jsp?id=293 ▼

[icon] http://cugir2.mannlib.cornell.edu/buckets/Display.jsp?id=293   🌐 Internet

Figure 3. Bucket #284 as it appears to the user.

access to CUGIR through OAI, the OPAC, and WorldCat. Although the data for this analysis are currently not available, the findings will certainly be considered as we determine the future of the CUGIR Metadata Sharing Project.

The conversion of CSDGM to MARC and then MARC to Dublin Core grants students, faculty, and staff more access to geospatial data via the OPAC as well as OCLC's online reference system, FirstSearch. CUGIR is finally integrated into the OPAC with other print and online collections of the Cornell University Library. Another positive outcome of the CUGIR project is the contribution of over 1,000 records to OCLC that can be shared with member institutions. Similarly, there are approximately sixty institutions that can share or harvest CUGIR metadata via OAI.

During the process of implementing the CUGIR Metadata Sharing Project, the team encountered a number of problems, one of which is worth mentioning here. While OCLC has made significant and positive changes to CORC since the beginning of the CUGIR Metadata Sharing Project in May 2001, CORC's ability to deal with cartographic materials is limited. For example, one of the most fundamental uses of maps is to specify a location on the surface of the earth using a locational system, or bounding coordinates [40]. In effect, the Dublin Core Metadata Initiative DCMI recommends a standard to categorize the bounding coordinates in Dublin Core, but, CORC does not yet recognize it [41 and 42]. We are hopeful that OCLC will eventually reconcile its Dublin Core implementation to the standard elements and qualifiers endorsed by the Dublin Core Metadata Initiative.

Minor Civil Divisions, Albany County

U.S. Department of Commerce. Bureau of the Census.

1998
English ▦ Map ▦ Computer File ⊛ Internet Resource
Washington, DC : Bureau of the Census,

Mark: ☐

Database: WorldCat

Ownership: **FirstSearch indicates your institution owns the item.**
• Libraries that Own Item: 1
• Connect to the catalog at Cornell University Library

Title: **Minor Civil Divisions, Albany County**

Publication: Washington, DC :; Bureau of the Census,

Year: 1998

Language: English

Abstract: These files are an extract of selected geographic and cartographic information from the 1995 TIGER/Line files detailing county subdivisions. This dataset includes minor civil divisions and other statistical entities.

Access: ➲ http://cugir2.mannlib.cornell.edu/buckets/Display.jsp?id=284

Figure 4. MARC record in WorldCat. The 856 field leads back to bucket #284.

5. Conclusion

This paper began with the following statement from Carl Lagoze: "Metadata is not monolithic. Instead, it is helpful to think of metadata as multiple views that can be projected from a single information object." We are now at a point in this paper where it is possible to decipher Lagoze's statement. CUGIR geospatial metadata now exists beyond the confines of bits residing on servers in Ithaca, NY and even beyond the world of geospatial information systems. By implementing the CUGIR Simple Bucket Model, we have been able to share, reuse, and repurpose the CSDGM metadata in information systems based on MARC/AACR2, Dublin Core, and OAI protocols. The "view" of our digital object is optimized for each particular system.

The massive investment in the MARC standard works in the library community's favor: the "network effects" that libraries accrue from this international network should be preserved and enhanced while we are making the transition to a heterogeneous metadata environment. The way forward is to build open systems and methods that embed the ability to link between different types of metadata. The CUGIR Metadata Sharing Project is an example of a working system utilizing a heterogeneous linking methodology to enhance resource discovery and to share both MARC and non-MARC metadata in a variety of information systems.

References

1. C. Lagoze , Keeping Dublin Core simple: cross-domain discovery or resource description? Online. *D-Lib Magazine* (January 2001). http://www.dlib.org/dlib/january01/lagoze/01lagoze.html

2. L. Hwey Jeng , A converging vision of cataloging in the electronic world. *Information Technology and Library* **15** 4 (1996), pp. 222–230.

3. J. K. H. Lee Smeltzer , Finding the needle: controlled vocabularies, resource discovery, and Dublin Core. *Library Collections, Acquisitions & Technical Services* **24** 2 (2000), pp. 205–215.

4. J. Darzentas , Sharing metadata: enabling online information provision. *OCLC Systems & Services* **15** 4 (1999), pp. 172–178.

5. J. A. Younger , Resources description in the digital age. *Library Trends* **45** 3 (1997), pp. 462–488.

6. M. H. Szunejko , Description of Internet resources: a consideration of the relationship between MARC and other metadata schemes. *Technical Services Quarterly* **18** 3 (2001), pp. 1–9.

7. N. Medeiros , Report from the trenches—the 8th International Dublin Core Metadata Initiative Workshop. *OCLC Systems & Services* **17** 1 (2001), pp. 15–18.

8. H. R. Lange and B.J. Winkler , Taming the Internet: metadata a work in progress. *Advances in Librarianship* **21** (1997), pp. 47–72.

9. S. Weibel , Metadata: the foundation of resource description Online. *D-Lib Magazine* (May 19, 1995). http://www.dlib.org/dlib/July95/07weibel.htm

10. L. Dempsey and R. Heery , Metadata: a current view of practice and issues. *Journal of Documentation* **54** 2 (1998), pp. 145–172.

11. M. El-Sherbini , Metadata and the future of cataloging. *Library Review* **50** 1 (2001), pp. 16–27.

12. S. L. Vellucci , Options for organizing electronic resources: the coexistence of metadata. *Bulletin of the American Society for Information Science* **24** 1 (1997), pp. 14–17.

13. N. Medeiros , Peering over the fortress walls: the metadata invasion begins. *OCLC Systems & Services* **17** 4 (2001), pp. 154–156.

14. C. Shapiro and H. R. Varian *Information rules: a strategic guide to the network economy*, Harvard Business School Press, Boston (1999).

15. C. Shapiro and H. R. Varian In: *Information rules: a strategic guide to the network economy*, Harvard Business School Press, Boston (1999), p. 12.

16. OCLC, WorldCat® (the OCLC® Online Union Catalog) is the world's most comprehensive reference resource, with over 45 million bibliographic records representing 400 languages. WorldCat covers information dating back to the 11th century, and includes holdings information from libraries in 45 countries. http://www.oclc.org/oclc/man/6928fsdb/worldcat.htm

17. C. Prabha , CORC users: what are they cataloging? *CORC Users Group Meeting* (April 29, 2001) http://www.sil.si.edu/staff/CORC-USER-GROUP/chandra.ppt

18. J. K. H. Lee Smeltzer , Finding the needle: controlled vocabularies, resource discovery, and Dublin Core. *Library Collections, Acquisitions & Technical Services* **24** 2 (2000), pp. 209–210.

19. J. A. Younger , Resources description in the digital age. *Library Trends* **45** 2 (1997), pp. 474–478.

20. J. Milstead and S. Feldman , Metadata: cataloging by any other name. *Online* **23** 1 (1999), pp. 22–27.

21. N. Medeiros , Making room for MARC in a Dublin Core world. *Online* **23** 6 (1999), pp. 57–61.

22. CUGIR *Cornell University Geospatial Information Repository (CUGIR). Online*, Albert R. Mann Library, Cornell University (1998). http://cugir.mannlib.cornell.edu/

23. P. Herold, T. D. Gale and T. P. Turner, Optimizing web access to geospatial data: the Cornell University Geospatial Information Repository (CUGIR). *Issues in Science and Technology Librarianship* (Winter 1999). http://www.library.ucsb.edu/istl/99-winter/article2.html

24. USGS, According to the United States Geological Survey, GIS is a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information, i.e.

data identified according to their locations. Practitioners also regard the total GIS as including operating personnel and the data that go into the system. (February 9, 2002). http://www.usgs.gov/research/gis/title.html

25. The entire process of converting metadata in numerous systems in order to increase access to geospatial information will be known as the "CUGIR Metadata Sharing Project."

26. The CUGIR Metadata Sharing Project Team consists of the Metadata Librarian- Elaine Westbrooks, Information Technology Librarian—Adam Chandler, a student programmer—Vivek Uppal, and Bo Wang, a student assistant who cataloged the records using CORC.

27. M. K. Bergman , The deep Web: surfacing hidden value (White Paper) *BrightPlanet* (July 2000). http://www.brightplanet.com/deepcontent/tutorials/DeepWeb/deepwebwhitepaper.pdf

28. A variety of statistics are collected, and organized on a monthly basis. The statistics are stored in a relational database for analysis.

29. K. Calhoun , CORC and collaborative Internet resource description: a new partnership for technical services, collection development and public services. *Journal of Internet Cataloging* **4** 1/2 (2001), pp. 131–142.

30. D. Y. Allen , Using Dublin Core with CORC to catalog digital images of maps. *Journal of Internet Cataloging* **4** 1/2 (2001), pp. 163–177.

31. I. Hseih-Yee and M. Smith , The CORC experience: survey of founding libraries. Part I. *OCLC Systems & Services* **17** 3 (2001), pp. 133–140.

32. OAI, Mission Statement. Online. *Open Archives Initiative* (no date). http://www.openarchives.org/organization/index.html

33. UKOLN, Dublin Core metadata editor, also known as DC-dot was developed by the UK office for Library and Information Networking (UKOLN), and is used to generate Dublin Core Metadata that can be inserted into Web pages or converted to USMARC, TEI, and GILS. Online. *UKOLN* (2000). http://www.ukoln.ac.uk/metadata/dcdot/

34. N. Medeiros , A craftsman and his tool: Andy Powell and the DC-dot metadata editor. *OCLC Systems & Services* **17** 2 (2001), p. 64.

35. A. Chandler, D. Foley and A. Hafez , Mapping Essential Federal Geographic Data Committee Metadata to MARC21 and Dublin Core: Towards an Alternative to the FGDC Clearinghouse. *D-Lib Magazine* (January 2000). http://www.dlib.org/dlib/january00/chandler/01chandler.html

36. The core record standard for cartographic materials has not been fully established, however, the CUGIR Metadata Team is closely following the draft in anticipation of its acceptance at the end of 2002.

37. S. Payette , Persistent identifiers and digital terrain Online. *RLG DigiNews* (1998). http://www.rlg.org/preserv/diginews/diginews22.html

38. OCLC, Persistent URL Homepage Online. *OCLC* (N. D.). http://purl.oclc.org/

39. M. Nelson , Buckets: smart objects for digital libraries. *Thesis (p. ii)*, Old Dominion University, Norfolk (2000).

40. FGDC, According to the Content Standard for Digital Geospatial Metadata, CSDGM, the bounding coordinates 1.5.1. are the limits of coverage of a data set expressed by latitude and longitude values in the order western-most, eastern-most, northern-most, and southern-most. Online. *FGDC. (N.D.)* http://www.fgdc.gov/metadata/csdgm/

41. S. Cox , DCMI point encoding scheme: a point location in space, and methods for encoding this in a text string. Online. *DCMI* (2000). http://www.dublincore.org/documents/dcmi-point/

42. S. Cox , DCMI information technology, and libraries box encoding scheme: specification of the spatial limits of a place, and methods for encoding this in a text string Online. *DCMI* (2000). http://www.dublincore.org/documents/dcmi-box