



Guidance on the Structure, Content, and Application of Metadata Records for Digital Resources and Collections

Report of the IFLA Cataloguing Section Working Group on the Use of Metadata Schemas

**Draft – for Worldwide Review
27 October, 2003**

1.0 Introduction

- 1.1 Background to the report
- 1.2 Working Group's terms of reference
- 1.3 Working Group membership
- 1.4 Working Group activities and accomplishments

2.0 Using Metadata in Libraries or Other Information-Intensive Organizations

- 2.1 Initiating a Metadata Project
- 2.2 Selecting a Metadata Schema or Schemas
- 2.3 Choosing between Metadata and Bibliographic Records

3.0 Metadata Core Record Framework

- 3.1 Deriving the Framework
- 3.2 Core record elements and definitions
- 3.3 Potential applications for the common core record

Bibliography

Appendices:

- I. Proposed core record metadata elements and functional requirements (FRBR)
- II. Common core record elements and their presence in other metadata schemas

Guidance on the Structure, Content, and Application of Metadata Records for Digital Resources and Collections

Report of the IFLA Cataloguing Section Working Group on the Use of Metadata Schemas

1.0 INTRODUCTION

1.1 Background to the report

Over the past decade, the proliferation of electronic texts, images, sounds, and objects stored in Internet- and Intranet-accessible knowledge bases or other digital repositories has increased the potential range and quantity of readily-accessible multimedia information. It has also resulted in what Levy (1990)¹ has referred to as a “second flood”, threatening to drown the engaged searcher in massive amounts of material, both useful and irrelevant. The need to delimit electronic resources more precisely in order to facilitate access has intensified activity in the development of metadata schemas – generally defined as “data about data”. This “Metadata Movement”, as Baker² has described it, has included the development of general application metadata schemas, such as the Dublin Core (DC), the Government Information Locator Service [now Global Information Locator Service] (GILS), or Digital Object Identifier (DOI), as well as domain-specific metadata schemas, such as the Text Encoding Initiative (TEI), the Encoded Archival Description (EAD), the Consortium for the Interchange of Museum Information (CIMI) metadata set, the Visual Resources Association (VRA) Core Categories, the Content Standard for Digital Geospatial Metadata (CSDGM), and the Online Information Exchange (ONIX) publishing standard – to name only a few. Such schemas are based on a common “machine-readable” syntax, such as HTML (Hypertext Markup Language), SGML (Standard Generalized Markup Language), or XML (eXtensible Markup Language). Metadata-enabled search engines can thus retrieve by precise metatags and values, those electronic resources in which a metadata record is embedded, or to which a separately housed metadata record points.

1.2 Working Group’s terms of reference

In response to a growing international interest in, and application of, metadata schemas³ the IFLA Cataloguing Section Working Group on the Use of Metadata Schemas was established at the IFLA 1998 Conference in Amsterdam, the Netherlands. The Working

¹ Levy, Pierre. 1990. *Les Technologies de l’Intelligence*. Paris: La Découverte.

² Baker, Thomas. 1999. Organizing Access with Metadata. *TIAC White Paper on Appropriate Technology for Digital Libraries at URL: http://www.tiac.or.th/tiacweb/Baker/Section2_3.html*. Accessed 31-07-02

³ For a definition of the term, “schemas”, see URL: <http://www.linktionary.com/s/schema.html>

Group drafted its terms of reference agreeing to focus on outcomes deriving from the following three objectives:

- *Objective 1:* to create an inventory of the development and implementation/application of metadata schemas in different countries
- *Objective 2:* to provide guidance (and ultimately, as appropriate, guidelines) to libraries as to when and/or how best to use metadata records and bibliographic records (catalogue records)
- *Objective 3:* to determine a metadata “core record” – i.e., a set of most commonly occurring elements in selected metadata schemas – that could be used by authors and/or publishers of electronic records to enhance resource discovery, and to provide, where appropriate, elements for incorporation into bibliographic records (catalogue records)

1.3 Working Group membership

While not all members listed below have been active throughout the full 1998-2003 period, Working Group membership has included the following:

- Fernanda Campos, Portugal
- Lois Mai Chan, USA
- Assumpcio Estivill, Spain
- Christel Hengel, Germany
- Lynne C. Howarth, Canada (Chair)
- Mona Madsen, Denmark (to 2001)
- Dorothy McGarry, USA
- Monika Muennich, Germany
- Eeva Murtomaa, Finland
- Glenn Patton, USA
- Charlotte Pederson, Denmark (2001-)
- Barbara Tillet, USA
- Mirna Willer, Croatia
- Maria Witt, France
- Maja Žumer, Slovenia

In addition, the Working Group has invited input – and, in some cases, representation – from the IFLA Sections on Information Technology, Classification and Indexing, Bibliography, and Libraries for the Blind, as well as from the Permanent UNIMARC Committee (PUC) and the DCMI Libraries Working Group.

1.4 Working Group activities and accomplishments

In fulfilling its terms of reference, the Working Group held five sets of formal meetings, beginning in Bangkok, Thailand at the IFLA 1999 Conference, and continuing at the IFLA 2000 Conference in Jerusalem, Israel, the IFLA 2001 Conference in Boston, USA, the IFLA 2002 Conference in Glasgow, Scotland, and the IFLA 2003 Conference in Berlin, Germany. Over the five year period, the Working Group concentrated on

completing its three objectives (see sections 1.4.1, 2.0, and 3.0, respectively), co-sponsored an IFLA program (see section 1.4.1) and an IFLA Discussion Group (see section 1.4.2), and engaged in discussions with the DCMI Libraries Working Group (see section 1.4.3). These various efforts have culminated in the report for worldwide review (see section 1.4.4), for which extensive international feedback is being sought.

1.4.1 Metadata Schema Development, Applications, and Implementations

Working Group Terms of Reference: Objective 1: To create an inventory of the development and implementation/application of metadata schemas in different countries.

Beginning work on the first objective within the Working Group's (WG) terms of reference, and in preparation for the IFLA 1999 Conference WG meetings, each member had volunteered to research a particular country or region, and subsequently to provide an update on how metadata schemas were being developed and/or applied in those constituencies. Such large-scale activities as the Nordic Metadata Project, the Cooperative Online Resource Catalog (CORC) – a research project undertaken by OCLC to explore the cooperative creation of a catalogue of Internet resources – the MetaWeb Project in Australia, metadata research at the U.K. Office for Library and Information Networking (UKOLN) at the University of Bath, England, and ongoing metadata projects at a number of libraries in Germany were briefly discussed. Lois Chan also provided, for information, and as an example of the kinds of metadata-related work other constituents of the bibliographic community were engaging in, a summary of a working draft of *Subject Data in the Metadata Record* – a report from the ALCTS/CCS/SAC Subcommittee on Metadata and Subject Analysis (see Final Report at URL: <http://www.ala.org/efapps/archive.cfm?path=alcts/organization/ccs/sac/metarept2.html>) (Accessed 30/10/03). Working Group members were also directed to the IFLA website where the link for electronic resources leads to a comprehensive site containing metadata schemas, examples of metadata projects, and an extensive bibliography dealing with various aspects related to metadata. The site can be accessed at URL: <http://www.ifla.org/II/metadata.htm> (Accessed 30/09/03).

At the IFLA 2000 Conference in Jerusalem, the Working Group on the Use of Metadata Schemas sponsored a workshop to showcase a number of metadata projects currently underway around the world. Speakers described such applications as the OCLC CORC project, Dublin Core and GILS-enabled websites in Canada, the European-resident Renardus project, implementations of the Dublin Core in Germany, the design of a metadata-enabled national database in Russia, and a number of implementations – including the development of metadata registries – under the auspices of the Dublin Core Metadata Initiatives (DCMI) Working Group. With the obvious intensifying of metadata-focused activity, the Working Group determined that its initial objective of *creating an inventory* of implementations was no longer feasible. Consequently, the first term of reference was modified to *monitoring* metadata standards development and applications/implementations internationally.

By the IFLA 2001 Conference in Boston, the number and range of metadata-related

projects had grown beyond the monitoring resources of the Working Group. Moreover, it was felt that, with the emergence of metadata registries for identifying and tracking implementations and adaptations of different metadata schemas, there were formal and well-maintained sources for the international bibliographic community to access. Therefore, the Working Group suggests that, while the original first objective from the terms of reference was “to create an inventory of the development and implementation of metadata schemas in different countries”, the tremendous growth in adopting, adapting, and applying metadata schemas has rendered this objective unrealistic, and largely unattainable without substantial investment of resources.

Recommendations:

- The Working Group encourages the IFLA web content manager to continue work on the metadata site to include and maintain links to key information sources and registries for those metadata schemas most widely used within the bibliographic community.
- The Working Group further suggests that the schemas referenced in this report and used as a basis for the common core of metadata elements, constitute the minimum set of metadata standards to which the IFLA metadata site points.

1.4.2 Metadata Discussion Group

In addition to addressing its terms of reference specifically, the Working Group also discussed a proposal from the Internet Discussion Group (within the Information Technology Section) to collaborate on forming a Metadata Discussion Group. With subsequent approval from the IFLA Professional Board, the Metadata Discussion Group was established, with scheduled times for meeting at the IFLA conferences in Jerusalem (2000), Boston (2001), and Glasgow (2002). Co-sponsored by the Cataloguing and Information Technology Sections, and co-chaired by Larry Woods (IT Section) and Lynne Howarth (Cataloguing Section; Working Group on the Use of Metadata Schemas), the Metadata Discussion Group has embodied what many metadata projects, themselves, have exhibited, namely, cooperation and knowledge sharing between those who are responsible for information technology infrastructure and the systems and tools for communicating – the metadata “carriers” – and those who create and/or manage content using metadata.

1.4.3 Communicating with the Dublin Core Metadata Initiative (DCMI) Libraries Working Group

At Working Group meetings during the IFLA 2000 Conference in Jerusalem, members discussed the usefulness of engaging with the Dublin Core Metadata Initiative (DCMI) Libraries Working Group (LWG), and exchanging information concerning DCMI and IFLA Metadata Working Group activities. This contact was initiated immediately following the IFLA 2000 Conference, with both groups subsequently agreeing to share minutes of meetings and to keep one another apprised of work in progress. The chairs of the respective working groups met at the American Library Association Annual

Conference in San Francisco, USA, in June 2001, to discuss items of mutual interest that would be on the agenda of each group at the upcoming IFLA meetings. Consequently, several members of the IFLA Working Group on the Use of Metadata Schemas met jointly with the DCMI Libraries Working Group during the IFLA 2001 Conference in Boston. The proposed “Library Application Profile”, as developed by the DCMI LWG, was discussed at the meeting, with LWG chair, Rebecca Guenther encouraging IFLA Metadata WG members to sign on to the DCMI discussion list, and to participate in further deliberations concerning the Profile document.

Recommendation:

The IFLA Working Group on the Use of Metadata Schemas recommends that the IFLA Cataloguing Section continue to discuss topics, exchange documents, and/or engage in activities of mutual interest concerning metadata applications with the DCMI Libraries Working Group, and to ensure that there is some mechanism for having IFLA Cataloguing Section input to the DCMI “Library Applications Profile” as it continues to evolve.

1.4.4 Worldwide Review of the Working Group Report

Having completed a draft of its final report, the Working Group is posting the document to the IFLA website for Worldwide Review. In addition to seeking comment from libraries and other constituents within the bibliographic community, feedback is being solicited from the following with a potential interest in, or connection within, metadata schemas and applications:

- Archival and museums communities
- Publishing community
- National Libraries – specifically regarding possible application in concert with legal deposit
- Other groups with an interest in metadata interoperability – e.g., those overseeing specific metadata schemas
- Those organizations or associations currently building and/or maintaining metadata-enabled subject gateways

Responses received by 31 January, 2004, will be addressed, as applicable, in the revised final report. The Working Group is committed to having approval and final IFLA sign-off, as required, prior to the IFLA 2004 meeting in Buenos Aires, Argentina.

2.0 USING METADATA IN LIBRARIES OR OTHER INFORMATION-INTENSIVE ORGANIZATIONS

Working Group Terms of Reference: Objective 2: To provide guidance (and ultimately, as appropriate, guidelines) to libraries as to when and/or how best to use metadata records and bibliographic records (catalogue records).

Concerning the second objective, the Working Group discussed the confusion that exists within the bibliographic community as to when and how best to use metadata records versus catalogue records. What, if any, is the relationship between the two? The Working Group agreed that, at this stage, it would be more appropriate to provide context or guidance rather than formal guidelines to libraries that are incorporating electronic resources into their collections. Because metadata schemas are being implemented currently on a project by project basis, rather than being broadly applied, and because metadata schemas as well as the resources to which they are applied and the software venues in which they are being supported are very much in flux, guidelines for metadata versus bibliographic record use seem premature.

This piece of the Working Group's charge, then, begins with suggestions concerning how to approach a metadata project, the types of metadata that may be required to support the project, and whether to use an existing or institute a local metadata schema. The section concludes with some recommendations concerning the most appropriate use of metadata records – as opposed to, or in concert with, bibliographic records – as a potential context or guidance to libraries planning to incorporate metadata in some part of their bibliographic activities.

2.1 *Initiating a Metadata Project*

- *Defining the Scope, Inclusion Criteria, and Purpose of the Project*

Prior to initiating a metadata project, it will be important for the library or bibliographic agency to define clearly the nature, scope, coverage, inclusion/exclusion criteria for selection, and format(s) of items or objects to be included in the “digital collection” (portal, subject gateway, knowledge repository, etc.) for which metadata records will be created. For what kinds of electronic resources will identification and access (links) be required?

- *General Metadata Record Types or Structures*

Depending on what is included, and the intended purpose(s) or use of the library's specified digital collection, a number of *types* or *structures* of metadata may be considered appropriate to the configuration of the final surrogate record. As a review of a number of large-scale metadata implementations confirm, these can be broadly categorized as follows:

- *Administrative metadata*: “housekeeping” information about the record itself – its creation, modification, relationship to other records, etc. Examples of

elements pertaining to administrative metadata include, but are not restricted to, the following:

- Record number
 - Date of record creation
 - Date of last modification
 - Identification of creator/reviser of record
 - Language of record
 - Notes
 - Relationship of this record to other(s)
- ***Descriptive metadata***: information describing the physical and intellectual properties or content of a digital item or object with such elements as:
 - Title (also alternative and parallel titles; subtitles; short titles; etc.)
 - Creator (author; composer; cartographer; artist; etc.)
 - Date
 - Publisher
 - Unique identifiers (ISBN; ISSN; etc.)
 - Dynamic links (URI; URL; etc.)
 - Summary; descriptive note; review; etc.
 - Audience level
 - Physical media; format; etc.
 - Language of the item or object
 - Version
 - ***Analytical metadata***: information analysing and enhancing access to the resource's contents. Sometimes referred to as “subject metadata”, elements may include:
 - Controlled subject terms, e.g., subject headings, descriptors
 - Subject/topic keywords
 - Abstract; Table of Contents (TOC)
 - Codes derived from classification systems or categorization schemes
 - Other elements of local importance, e.g., department affiliation; links to other related e-content
 - ***Rights management metadata***: information regarding restrictions (legal; financial; etc.) on access to, or use of, digital items or objects. Such elements as the following may apply:
 - Restrictions on use
 - Permission statements
 - Subscriber/licensing/pay-per-use fees
 - Acknowledgements
 - Copyright notice
 - Retention schedules
 - Quality ratings
 - Use disclaimers

- **Technical metadata:** particular hardware or software used in converting an item/object to a digital format, or in storing, displaying, etc., may require the use of such elements as:
 - Digitizing equipment specifications
 - Camera positions
 - Shooting conditions
 - Coding parameters
 - Voice recognition and/or read-back hardware and software
 - Optical scanner specifications
 - Image rendering equipment
 - Type of file and conversion software requirements

- **Other, as determined** – e.g., particular metadata elements based on local, regional, and/or organizational requirements, or in accordance with a nationally mandated metadata standard, and not subsumed within any metadata type above.

2.2 *Selecting a Metadata Schema or Schemas*

The choice of a metadata schema or schemas to be used in creating the surrogate records for uniquely identifying and linking to digital items or objects in a collection will depend on where and how the resources will be accessed and used. For example, a local land registry site of scanned documents, photographs, and maps, accessible exclusively on the organization’s Intranet, and fully maintained in-house, might necessitate only descriptive and technical metadata. On the other hand, a digital repository created and maintained by a distributed network of national organizations with content comprised of high quality Web sites (text and images only), and limited to a subject area in a technical domain might require a mix of administrative, descriptive, and analytical metadata. Likewise, a “virtual exhibit” containing links to a variety of digital objects within an international consortium of public and private art galleries and museums would necessitate the use of technical and rights management metadata, in addition to those required for administrative, descriptive, and analytical purposes.

A final determination of metadata schema may also depend on the desired degree of *granularity*, or the amount of detail to be captured and represented in the metadata record. A “core record” – created using a metadata schema, such as the Dublin Core with its fifteen element set (any of which is optional, repeatable, and extensible) – may include, as appropriate and/or required, administrative, descriptive, analytical, and rights management metadata, and can accommodate information related to technical specifications. In some specialized domains, however, a metadata schema, such as the Dublin Core, may lack sufficient granularity (detail) to represent resources adequately, or the particular purposes to which the subject gateway is directed. The ONIX metadata standard for international publishing and publishers, or the Content Standard for Digital Geospatial Metadata (CSDGM) are two examples of rich, detailed, and highly technical metadata schemas, derived especially to deal with complex content and unique applications within the domain.

In addition to deciding on the level of detail to be captured in metadata-enabled records, the choice of schema can be narrowed in response to questions, such as the following:

- Is there a structured, rich format metadata standard that is appropriate to the items/objects selected for, and intended purpose of, the digital collection?
- Which metadata elements or fields would be most useful to the community of users the digital collection is intended to support? How much detail should those fields support?
- Which metadata elements or fields would be most useful to those who are creating and/or maintaining the digital collection? How much detail should those fields support?
- Will the use of, or access to, this digital collection be restricted in any way? How will (should) this be made explicit in the metadata record?
- Are there any requirements related to language, or format of material, or type of media for which particular (or additional) metadata elements or fields must be provided?
- Are there requirements to create or share resources among a network of collaborators with responsibility for the digital collection(s) Are (additional) metadata fields required to support network cooperation?
- If the use of more than one metadata schema is envisioned or required (sharing resources across networks), are authoritative cross-schema mappings (crosswalks) readily and immediately available to facilitate and maintain interoperability? Can resources represented in one metadata schema (or standard) be exchanged with collaborators who are using a different schema (or standard)?
- How widely used is a particular schema, and in what applications or environments comparable to the one currently proposed? How robust and/or flexible is the schema within different contexts?
- How readily can one migrate from this particular schema to another should data conversion be required at some time?
- How or how well does a particular schema comply with mandated organizational local, national, or international standards, if any?
- What human (numbers; education; training), technical, financial, or other resources are required to support the application of the metadata schema, and does the organization or operation have those resources readily and sufficiently available? Are there other practical constraints to implementing and maintaining a particular schema or schemas?

Having answered any or several of the preceding questions, the choice to use one or more *standardized* metadata schemas may be confirmed. Alternatively, an individual, organization, or network of libraries may determine that a local or “home grown” solution – a set of locally-determined and supported metadata elements – is the preferred option. Similarly, some choose to combine elements of an established standard, such as the Dublin Core, with elements appropriate to the local situation of resources and objectives. There is no single recipe or “one-size-fits-all solution to which metadata schema or standard to use in initiating a project.

Recommendation:

Where an in-house approach is selected, *the Working Group recommends that the core record framework (see Section 3.2) be used as both a starting point and a minimum standard for metadata content.*

2.3 Choosing between Metadata and Bibliographic Records

2.3.1 Suggestions regarding where to use “non-traditional” metadata schemas to create records describing electronic resources and/or digital objects

- *On their own – uniquely*

A library may choose to use a metadata schema to markup a collection of electronic resources that may not previously have been accessible to end-users. Such collections may include materials that only exist in electronic format (e.g., subject-focused websites), or that have been digitized (using some kind of optical scanning or digital image capture technology) and are being added to a library’s Intranet, Internet, portal, or knowledge repository for public or private use. For example, a collection of paper maps could be scanned and metadata records describing and providing dynamic links to those digital images could be created using an appropriate existing metadata schema (e.g., DC; CSDGM; VRA), or a locally devised metadata schema. The latter could be fully independent of any existing schema, or be derived based on an existing schema that is extensible.

Depending on the metadata schema selected, the library should be aware of the purpose(s) for which the surrogate descriptive metadata record(s) is/are being created (internal and/or external discovery of the resource; legal deposit compliance; e-business application; inventory control; etc.), in order to ensure that all metadata elements required are embedded in the record template. Are additional metadata for administrative, technical, legal use, archival, or other purposes, required? Moreover, how will metadata be derived, modified, and maintained across time? A procedures manual may be required to ensure consistency of application and use across time and across an institutional environment. Likewise, how will links from the record to the item be created and maintained in a dynamic environment? Finally, to what extent will this digital collection be integrated with other databases or repositories of electronic materials (“legacy knowledge”) within the institution’s collections?

- *Instead of/in the place of a “traditional” bibliographic record*

In some cases libraries are choosing to continue using “traditional” cataloguing standards (e.g., MARC 21 or UNIMARC with content standards ISBD or AACR) for physical collections of print and/or audio-visual items, and “emerging” metadata standards (DC; TEI; CIMI) for electronic/digital resources accessible via the Web. There has been movement towards standardizing on one metadata schema to facilitate end-user understanding of, and access to, materials regardless of their type of physical format. Such initiatives include the Resource Description Framework (RDF), the Open Archives

Initiative (OAI), and the Semantic Web project. The development of XML (itself based on SGML) has provided a common syntax for facilitating “interoperability” among metadata schemas. As individual elements within each schema are mapped to, and expressed in the language (or syntax) of, XML, exchange of data within the XML framework is greatly facilitated and transparent to the end-users. For consistency across multiple knowledge stores some libraries are choosing to move from a dual or multiple standards approach (e.g., a MARC format with ISBD for printed text and for physical objects and DC for electronic resources regardless of type), to a single metadata schema application (e.g., DC for all materials; TEI regardless of whether the text is printed in a paper-based or digital environment). This is more appropriately done, or more readily accomplished, when the volume of “legacy collections” (i.e., those for which records were created using long-established codes, standards, or protocols) is low or non-existent. Where this is not the case, planning for retrospective conversion of records may be required.

- *In addition to a “traditional” bibliographic record (complementary)*

With the growth in cooperative or collaborative projects, the number of tools that will automatically convert a record from one metadata schema format to another (e.g., expressing a GILS record within the Dublin Core metadata set) has been growing. So-called legacy records can readily co-exist with emerging metadata standards. A metadata-enabled record describing an electronic resource to which it is linked can be captured and converted into a MARC format for inclusion in a library’s online public access catalogue (OPAC) or WebPAC. When a Web-based search engine or Web crawler “discovers” a digital item, the metadata embedded within the HTML <Head> can be used as a foundation for the surrogate record that is added to the library’s internal knowledge repository. In short, metadata expressed in one environment can be harvested and re-used in another, as appropriate or required. This approach may be especially beneficial to linking across different subject domains, disciplines, fields, or applications, including those associated with archives, museums, art galleries, education, publishing, or government – to name only a few.

3.0 METADATA CORE RECORD FRAMEWORK

Working Group Terms of Reference: Objective 3: To determine a metadata “core record” – i.e., a set of most commonly occurring elements in selected metadata schemas – that could be used by authors and/or publishers of electronic records to enhance resource discovery, and to provide, where appropriate, elements for incorporation into bibliographic records (catalogue records).

3.1 Deriving the Framework

Much of the Working Group’s initial work was focused on its third objective. In Bangkok, Thailand, members discussed the need for defining a core level metadata record as a reference point for authors and publishers of electronic resources. We agreed that it would be useful to create a core record based on the framework of the Functional Requirements for Bibliographic Records (FRBR). Thus, a set of metadata element types would be required to fulfill each of the four functions defined by FRBR, as follows:

- Find (data corresponds to search criteria)
- Identify (confirm entity described in record corresponds to entity sought; distinguish between 2 entities with same title)
- Select (language; version)
- Obtain (place order, request; remote access)

The actual metadata element(s) assigned (i.e., entity names and values) would depend on which metadata schema was being used (e.g., Dublin Core; TEI; EAD; GILS). The Working Group concurred that it would be useful to make recommendations as to which elements would be mandatory versus optional for both electronic serial and integrating resources and monographic resources.

Members also agreed that the eight areas of the International Standard Bibliographic Description (ISBD), and the fifteen elements of the Dublin Core metadata schema might provide a useful starting point for deriving a baseline set of constituent named metadata elements for describing any electronic resource in any domain regardless of the metadata schema used (i.e., schema-independent). With Olivia Madison’s crosswalk mapping the Dublin Core to the FRBR, and Eeva Murtomaa’s mapping of ISBD(ER) to the FRBR as examples, the Working Group compiled a list of metadata elements that could be included in an FRBR-compliant record.

In essence, the proposed core record framework stands apart from, or is independent of, all metadata schemas, and is intended to address the following question. “Regardless of which metadata schema is being used⁴ to encode or markup an electronic resource, which metadata elements within a schema should be included in a surrogate record describing this resource so as to facilitate the processes of finding, identifying, selecting, and/or

⁴ Including existing metadata schemas (e.g., DC; TEI; EAD; VRA; CIMI; MARC 21), or those that are developed “in-house” or for local use only.

obtaining the item or object? While the proposed core record *contains* metadata elements, it is not in any way assumed to be, or to represent, or to replace any other metadata schema *per se*. It should be considered, rather, as a model, or framework, or structure for metadata elements descriptive of any type or format of electronic resource in any intellectual discipline or knowledge domain.

With the proposed list of elements having been drafted, Working Group members agreed to take responsibility for one (or more) of the elements, to compare how each was represented (entity names and values) across a sample of metadata schemas, and, based on that analysis, to define and/or refine the terminology of the assigned draft core record element further. Members also agreed to identify where there were gaps between and among the elements in the original crosswalk of metadata schemas, and to try to define what functions those elements that exist in one domain, but not in another, served. Further, how should that function be defined and expressed in the draft core record list of elements, if not already articulated in the draft list?

A crosswalk mapping eight metadata schemas (source metadata standards) to the MARC 21 format (baseline or target standard)⁵ was developed to facilitate the comparison process. This “master crosswalk” (located at URL: <http://www.fis.utoronto.ca/special/metadata/shortwalk.asp>) contains those schemas that are among the more widely used in the bibliographic control community, including:

- Encoded Archival Description (EAD)
- Dublin Core (DC)
- Government Information Locator Service (GILS) metadata schema – now Global Information Locator Service
- Text Encoding Initiative (TEI) Header
- Visual Resources Association (VRA) Core Categories
- Consortium for the Interchange of Museum Information (CIMI) metadata set
- Content Standard for Digital Geospatial Metadata (CSDGM)
- Online Information Exchange (ONIX) publishing standard

In addition to analyzing what elements were being used within each metadata schema, how they were represented and defined, and what function(s) they served, the “master crosswalk” also provided a ready visual summary of elements that match across all schemas (i.e., is it an exact match or are there variations in tag names or meanings, e.g., date of origin *versus* creation?), elements that correspond between two systems or among three or more, elements that are clearly unique to a domain and missing from all others, and elements where some inherent ambiguity exists (i.e., the same entity name being used in different metadata schemas to represent different concepts or contexts (e.g., “organization” as institution or as process).

⁵ St. Pierre, M., and LaPlant, W.P. (1998, October 15). Issues in crosswalking: Content metadata standards. NISO. <http://www.niso.org/press/whitepapers/crswalk.html> (Accessed 20/10/03)

During the IFLA 2001 Conference in Boston, USA, Working Group members presented an analysis of metadata elements and their treatment across a set of metadata schemas (as per the crosswalk - see URL above) as assigned at the IFLA 2000 meetings. As described previously, these elements, themselves, were derived from the FRBR framework, and with reference to the Dublin Core metadata schema. The analysis of the draft “core record” elements revealed, in brief, that while some elements (e.g., name assigned to the resource) were well and consistently represented across most or all of the selected metadata schemas, others appeared with highly variant names or meanings (e.g., coverage date; date of creation; date of origin), while others were missing from one or more schemas. There was considerable ambiguity associated with such elements as “resource identifiers” (i.e., the resource itself; the location of the resource), and “resource type” (i.e., data type, format, “container”, arrangement, etc.).

The Working Group determined that, with some minor changes as suggested from the analysis, the core of elements that could be applied regardless of metadata schema (i.e., schema-independent), consisted of the following ten (10) elements:

- Subject
- Date
- Conditions of use
- Publisher
- Name assigned to the resource
- Language/mode of expression
- Resource identifier
- Resource type (i.e., what the resource *is*, rather than what it is *about*)
- Author/creator
- Version

Lois Chan dubbed this ten element set, the “core of cores”. It was decided that, as follow-up to the Working Group meeting at IFLA 2001, members would write up a summary of their analysis to include the following:

- A brief definition of what the element includes (see section 3.2)
- Treatment (naming/definition/application) of the element across each schema represented on the crosswalk (see URL, above) (see also Appendix II)

These summaries were subsequently reviewed at IFLA 2002, and further refined in preparation for the IFLA 2003 meeting in Berlin, Germany. They form the basis of a draft set of guidelines concerning core metadata elements to include in any metadata record (“core of cores”). In terms of application, any metadata record, regardless of schema used to render that record, can contain some or all of the core elements in order to enhance the indexing of, and/or access to, the resource being described.

3.2 Core record elements and definitions

SUBJECT: Provides a term, keyword, or phrase that describes, identifies, or interprets the intellectual content of a work and what it depicts or expresses. These may include

proper names (e.g., people or events), geographic locations (places), time period covered, or topics (e.g., iconography, concepts, themes, or issues). Depending on the metadata schema being used, the descriptive terms used to communicate the subject of the work may be derived from either controlled vocabularies or natural language. In different metadata schemas, the element “subject” may pertain to different facets. In some cases, the term “subject” is used to mean topic or theme; in others, it may be used as a collective term implying various facets such as topic, time, place, etc. In the broad sense, it is defined as “what a resource is *about*.”

DATE: Indicates the particular year and may, as appropriate, include a month and/or day associated with the work.. Dates and times may be used for a number of different reasons and in a number of different contexts. The date(s) could describe when the work was created, published, modified, accessed, etc.

CONDITIONS OF USE: Indicates the limitations and legal rules that may restrict or deny access to a work, or that affect how the work (or, in CSDGM, the metadata describing the work) is to be used *after* access has been granted. Generally, these constraints are applied to ensure the protection of privacy or intellectual property. Restrictions may include regulations, special procedures imposed by a repository, donor, legal statute, or other agency regarding reproduction, publication, or quotation of the described materials. May also indicate the absence of restrictions, such as when copyright or literary rights have been dedicated to the public.

PUBLISHER: Provides the name, location, and other identifying and/or contact information concerning an entity responsible for making a resource available, whether by production, manufacture, maintenance, distribution, etc.

NAME ASSIGNED TO THE RESOURCE: The name or phrase given to a work (or code set in some cases in CSDGM), often referred to as title. It may consist of a word, phrase, character, or group of words and/or characters. The title may or may not adhere to bibliographic standards. Schemas describing artistic works and images (VRA and CIMI) also include the names of complex works or series and the discrete units within these larger entities (e.g., a print from a series, a panel from a fresco, a building within a temple complex), or may identify only the larger entity itself.

LANGUAGE/MODE OF EXPRESSION: Identifies the language and/or script, sublanguages, dialects, etc., of the intellectual content of the work. Language information may be indicated through the use of complete words or predetermined alphabetic, numeric, or alpha-numeric codes.

RESOURCE IDENTIFIERS: Unique names, alphabetic codes, or numbers associated with a work that are used consistently to distinguish one resource from another.

RESOURCE TYPE: Can be divided into two facets: type of content; and type of carrier. Carrier deals with the 'package' of the resource and content is how the resource is

presented (e.g., genre, data type). This category contains tags that describe the physical format, rendering, appearance, or construction of the work.

AUTHOR/CREATOR: Name(s) of organization(s) or individual(s) responsible for creating or compiling the intellectual or artistic content of the work. May include a brief statement indicating the nature of the responsibility.

VERSION: Provides information on the version, edition, or adaptation of a particular work, or relationships to other works.

3.3 *Potential applications for the common core record*

In the process of distilling a list of common elements that would provide the framework for a core metadata record independent of any particular metadata schema, the Working Group determined that the ten elements would offer national libraries a uniform data structure to recommend to those creating metadata-enabled resources for legal deposit. Thus, regardless of which metadata schema was used, a requirement to include those schema-specific elements that complied with the minimal record set of ten elements included within the “core of cores” framework, would apply a common standard for digital resource submission. This uniform metadata record structure would be conducive to subsequent record exchange – should that materialize at the local, national, or international level – and would facilitate mapping activities and crosswalking in support of interoperability. With this scenario in mind, the Working Group **recommends that national libraries and other bibliographic or cataloguing agencies consider adopting the ten element common core framework as a standard metadata record structure for organizing digital resources, including those submitted to national library collections under legal deposit.**

In addition to the recommendation for consideration by national libraries and other bibliographic or cataloguing agencies, the Working Group also identified some potential opportunities and challenges for extending the “common core”. While briefly outlined below, the Working Group is aware that the following issues remain to be addressed or resolved:

- Relating the “common core” to content standards, such as ISBD, AACR, and/or to metadata standards, such as DC
- Relating the “common core” more closely to the FRBR – within a theoretical framework, as well as within the context of application
- Relating the “common core” to other work ongoing in the metadata community
- Relating the “common core” to other metadata domains – not reinventing the wheel
- Determining possible applications, implementations, and strategies for disseminating the “common core record”
- Expanding the English language “common core” to the broader multilingual and multiscrypt environments
- Addressing national and international requirements for description, access, and dissemination of metadata-enabled resources

- Remaining open and flexible to a changing information technology (IT) environment, to new syntaxes, to possible new/emerging resource types, etc.
- Liaising with other communities
 - IT/IS
 - W3C
 - DCMI
 - ISO and other member bodies
- Liaising with other domains
 - Museums and galleries
 - Archives
 - Publishing
 - Other – geospatial; rights management; e-commerce; etc.
- Creating a “common language bibliographic record” with a “common core record”
- Exploring opportunities for common ground – for “interoperability” – syntactically and semantically
- Facilitating universal access to a variety of resources for a broad constituency of domains – expanding the world of knowledge
- Bringing important “traditional” skill sets and expertise to “emerging” digital arena

While currently beyond the scope of the Working Group, each of the preceding opportunities and challenges warrants consideration, and may perhaps be more appropriately addressed in the final report deriving from Worldwide review of the proposed “common core” as described in the consultation draft.

SELECTED BIBLIOGRAPHY OF METADATA RESOURCES

The literature dealing with metadata schemas, applications, and project implementations is vast and growing. The following represents a selection of articles and websites that address metadata in general, and/or that cover topics, such as application profiles, crosswalks and mapping, interoperability, and registries. Where a URL link is provided, and unless noted otherwise, the site was last accessed October 15, 2003.

The Working Group particularly recommends the International Federation of Library Associations and Institutions (IFLA) site, *Digital Libraries: Metadata Resources*, for its comprehensive listing of documents and tools. Periodically updated, this valuable resource is available on IFLANET at URL: <http://www.ifla.org/II/metadata.htm> .

Baca, M., ed. (2000). *Introduction to Metadata: Pathways to Digital Information*. Los Angeles, CA: Getty Research Institute. Version 2, 2000, available at URL: http://www.getty.edu/research/conducting_research/standards/intrometadata/index.html

Baker, T., Dekkers, M., Heery, R., Patel, M., and Salokhe, G. (2001). What terms does your metadata use? Application profiles as machine-understandable narratives. *Journal of Digital Information*, 2(2). Available at URL: <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Baker/>

Buckland, M. K., Chen, A., Chen, H.-M., Kim, Y., Lam, B., Larson, R. R., et al. (1999). Mapping entry vocabulary to unfamiliar metadata vocabularies. *D-Lib Magazine*, 5(1). Available at URL: <http://www.dlib.org/dlib/january99/buckland/01buckland.html>

Caplan, P. (2003). *Metadata fundamentals for all librarians*. Chicago: American Library Association.

Chan, L. M., and Zeng, M. L. (2002). Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis. *IFLA Journal*, 28 (5/6), 323-327. Available at URL: <http://www.ifla.org/V/iflaj/art280506.pdf>

Cromwell-Kessler, W. (1998). Crosswalks, metadata mapping and interoperability: What does it all mean? In Murtha Baca, ed. *Introduction to Metadata: Pathways to Digital Information*. Los Angeles, CA: Getty Information Institute. Pp. 19-22.

Day, M. (2001). *Metadata: Mapping between metadata formats*. UKOLN. Available at URL: <http://www.ukoln.ac.uk/metadata/interoperability/>

Day, M. (1999). *The metadata challenge for libraries: A view from Europe*. Available at URL: <http://www.ukoln.ac.uk/metadata/presentations/metadiversity/challenge.html>

Day, M. (2000). Resource discovery, interoperability and digital preservation: Some aspects of current metadata research and development. *Vine*, 117: 35-48.

Day, M., Duval, E., and Causton, L. (2001). *Schemas metadata watch report #5*. Available at URL: <http://www.schemas-forum.org/metadata-watch/5.html>

Dekkers, M. (2001). Application profiles; or, how to mix and match metadata schemas. *Cultivate interactive*, 3 (29 January 2001). Available at URL: <http://www.cultivate-int.org/issue3/schemas/>

Dovey, M. J. (2000). "Stuff" about "stuff": The differing meanings of "metadata". *Vine*, 116: 6-13.

Dublin Core Metadata Initiative. (2002). *Dublin Core Metadata Registry*. Available at URL: <http://fiord.ulis.ac.jp/dcregistry/index.html>

Duval, E., Hodgins, W., Stuart, S., and Weibel, S. (2002). Metadata principles and practicalities. *D-Lib Magazine*, 8(4), Available at URL: <http://www.dlib.org/dlib/april02/weibel/04weibel.html>

Gill, T. (2000). Metadata and the World Wide Web. In Baca, M., ed. *Introduction to Metadata: Pathways to Digital Information, Version 2*. Los Angeles, CA: Getty Research Institute. Available at URL: http://www.getty.edu/research/conducting_research/standards/intrometadata/2_articles/gill/index.html

Gilliland-Swetland, A. J. (2000). Introduction to metadata: Setting the stage. In Baca, M., ed. *Introduction to Metadata: Pathways to Digital Information, Version 2*. Los Angeles, CA: Getty Research Institute. Available at URL: http://www.getty.edu/research/conducting_research/standards/intrometadata/2_articles/index.html

Heery, R., Gardner, T., Day, M., and Patel, M. (2000). *DESIRE [Development of a European Service for Information on Research and Education] II metadata registry framework*. Available at URL: <http://www.desire.org/html/research/deliverables/D3.5/d35.html>

Heery, R., and Wagner, H. (2002). A metadata registry for the Semantic Web. *D-Lib Magazine*, 8(5). Available at URL: <http://www.dlib.org/dlib/may02/wagner/05wagner.html>

Hillman, D. (2003). *Using Dublin Core*. Dublin Core Metadata Initiative. Available at URL: <http://dublincore.org/documents/2003/08/26/usageguide/>

Howarth, L.C., Cronin, C., and Hannaford, J. (2002). Designing a Metadata-Enabled Namespace for Accessing Resources Across Domains. In Howarth, L.C., Slawek, A., and

Arsenault, C., eds. *Advancing Knowledge: Expanding Horizons for Information Science: Proceedings of the 30th Annual Conference of the Canadian Association for Information Science/Association canadienne des sciences de l'information – held with the Congress for the Social Sciences and Humanities of Canada, University of Toronto, Toronto, Canada, 30 May-1 June, 2002*. Pp. 223-232.

Hudgins, J., Grace, A., and Brown, E. (1999). *Getting mileage out of metadata: applications for the library*. Chicago: American Library Association.

Hunter, J., and Lagoze, C. (2001). *Combining RDF and XML schemas to enhance interoperability between metadata application profiles*. Available at URL: <http://archive.dstc.edu.au/RDU/staff/jane-hunter/www10/paper.html>

Lagoze, C., and Hunter, J. (2001). The ABC ontology and model. *Journal of Digital Information* 2(2). Available at URL: <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/>

Lagoze, C. (2001). Keeping Dublin Core simple: Cross-domain discovery or resource description? *D-Lib magazine*, 7(1). Available at URL: <http://dlib.org/dlib/january01/lagoze/01lagoze.html>

Leviston, T. (2001). *Metadata registry requirements and realities*. Available at URL: <http://www.dstc.edu.au/Research/Projects/Infoeco/publications/registry-discussion.pdf>

Manola, F., and Miller, E., eds. (2003). *World Wide Web Consortium RDF primer*. Available at URL: <http://www.w3.org/TR/rdf-primer/>

Miller, A. P. (1998). An introduction to Resource Description Framework. *D-Lib Magazine*. Available at URL: <http://www.dlib.org/dlib/may98/miller/05miller.html>

Miller, A. P. (1996). Metadata for the masses. *Ariadne*, 5. Available at URL: <http://www.ariadne.ac.uk/issue5/metadata-masses/>

Milstead, J., and Feldman, S. (1999). Metadata: Cataloguing by any other name. *Online* 23 (1). Available at URL: <http://www.cbuc.es/5digital/1.pdf>

Neuroth, H., and Koch, T. (2001). *Metadata mapping and application profiles: Approaches to providing the cross-searching of heterogenous resources in EU project Renardus*. Paper presented at the DC-2001, International Conference on Dublin Core and Metadata Applications, Tokyo. Available at URL: <http://www.nii.ac.jp/dc2001/proceedings/product/paper-21.pdf>

Schwartz, C. (2001). Metadata. In *Sorting out the web: approaches to subject access*. Westport, Conn.: Ablex. Pp. 9-45.

St. Pierre, M., and LaPlant Jr., W. P. (1998). *Issues in crosswalking content metadata standards*. National Information Standards Organization (NISO). Available at URL: <http://www.niso.org/press/whitepapers/crsswalk.html>

State and University Library at Gottingen, Germany (SUB). (2003). *Metadata Server*. Available at URL: <http://www2.sub.uni-goettingen.de/index.html>

Thornely, J. (2001). *Metadata - What is it?* State Library of Queensland. Available at URL: <http://www.slq.qld.gov.au/pub/staff/metadata.htm>

UKOLN Metadata Group. (1999). *A review of metadata: A survey of current resource description formats*. Work Package 3 of Telematics for Research project DESIRE. Available a URL: <http://www.ukoln.ac.uk/metadata/desire/overview/>

UK Office of Library and Information Networking. (2002). *UKOLN Interoperability Focus*. Available at URL: <http://www.ukoln.ac.uk/interop-focus/>

Woodley, M. (2000, May 07). Crosswalks: The path to universal access? In Baca, M., ed. *Introduction to Metadata: Pathways to Digital Information, Version 2*. Los Angeles, CA: Getty Research Institute. Available at URL: http://www.getty.edu/research/institute/standards/intrometadata/2_articles/woodley/index.html

World Wide Web Consortium. (2003). *Resource Description Framework*. Available at URL: <http://www.w3.org/RDF/>

APPENDIX I

**PROPOSED CORE RECORD METADATA ELEMENTS
AND FUNCTIONAL REQUIREMENTS (FRBR)**

<i>FUNCTIONAL REQUIREMENTS</i> ----- ELEMENTS:	<i>IDENTIFY</i>	<i>SELECT</i>	<i>FIND</i>	<i>OBTAIN</i>
Subject	X	X	X	
Date	X	X		X
Conditions of use				X
Publisher	X	X		X
Name assigned to the resource	X	X	X	X
Language/mode of expression		X		
Resource identifier	X		X	X
Resource type	X	X		X
Author/creator	X	X	X	X
Version	X	X		X

APPENDIX II

**COMMON CORE RECORD ELEMENTS AND THEIR
PRESENCE IN OTHER METADATA SCHEMAS**

SCHEMAS ----- ELEMENTS	MARC 21	UNI MARC	DC	TEI	EAD	VRA	CSDGM/ FGDC	CIMI	GILS	ONIX
Subject	X	X	X	X	X	X	X	X	X	X
Date	X	X	X	X	X	X	X	X	X	X
Conditions of use	X	X	X		X		X	X		X
Publisher	X	X	X	X	X *		X	X	X	X
Name assigned to the resource	X	X	X	X	X	X	X	X	X	X
Language/ Mode of expression	X	X	X	X	X				X	X
Resource identifier	X	X	X	X		X	X	X	X	X
Resource type	X	X	X	X	X	X	X	X	X	X
Author/ Creator	X	X	X	X	X	X	X	X	X	X
Version	X	X	X	X			X			X

* at the collection level