



World Library and Information Congress: 70th IFLA General Conference and Council

22-27 August 2004
Buenos Aires, Argentina

Programme: <http://www.ifla.org/IV/ifla70/prog04.htm>

Code Number: 024-S
Meeting: 148. IFLA-CDNL Alliance for Bibliographic Standards (ICABS)
Simultaneous Interpretation: -

Metadatos, Protocolo y Actividades de Identificación: Alianza Biblioteca del Congreso IFLA/CDNL para el Reporte de Estándares Bibliográficos, 2004, Buenos Aires

Sally H. McCallum
Biblioteca del Congreso de EE.UU.

Sumario

La contribución de la Biblioteca del Congreso de EE.UU. a la Alianza para Estándares IFLA/CDNL incluye formatos MARC 21, metadatos, protocolos para recuperación de información, e identificadores. Este informe cubre algunas de las actividades relevantes recientes en cada una de estas áreas. Para MARC 21, es mencionado el desarrollo de MARCXML y el conjunto de herramientas MARC para transformaciones de formato. En el área de metadatos, se delinea el Estándar de Descripción de Objetos de Metadatos (MODS) junto con MADS, su nuevo acompañante para datos de autoridades. Destacado en el área de los protocolos de recuperación de información, se encuentra el nuevo servicio Web derivativo Z39.50 denominado SRW/SRU; y bajo identificadores, establecido especialmente para la comunidad de información, está el nuevo “info” URI. Completa este informe la situación del modelo de Archivo Virtual Internacional de Autoridades.

Como han oído hoy, ICABS es un concepto nuevo para apoyar el trabajo de IFLA y su participación en una miríada de importantes actividades de información. Algunos de los programas ya estaban bien establecidos en IFLA, en otros casos aún estamos entablando relaciones. Una premisa básica ha sido que las actividades sean extensiones de aquellas que ya se encuentren operando dentro de alguna de las instituciones participantes en la Alianza. Mi institución, la Biblioteca del Congreso de EE.UU., es responsable de cuatro áreas referidas a:

MARC 21 y sus derivados en XML; el protocolo de recuperación de información Z30.50 y su versión de próxima generación; metadatos y esquemas de metadatos basados en XML; identificadores persistentes. Además estamos asociados con Die Deutsche Bibliothek en un modelo experimental de Archivo Virtual Internacional de Autoridades (VIAF por sus siglas en inglés.)

Hoy quisiera referirme en breve a uno o dos temas en cada una de estas áreas que considero especialmente importantes o que han desarrollado algún tipo de actividad significativa en el transcurso de este último año.

MARC 21

Como ustedes probablemente saben, MARC 21 está basado en una estructura de formato ISO 2709, que es un estándar muy bien establecido mundialmente. La estructura, que también es usada para UNIMARC y muchos otros formatos similares a MARC, ha funcionado bien y seguirá siendo importante por muchos más años. Pero en la actualidad, muchos protocolos nuevos están más “cómodos” con una estructura para el récord bibliográfico basada en XML. En respuesta a esa necesidad, en estos últimos años han sido desarrolladas estructuras alternativas para los elementos de datos en MARC 21. La más importante es MARCXML, que provee un camino de ida y vuelta ininterrumpido de MARC 21 a MARC en XML.

Estos desarrollos han sido importantes para permitir que nuestros vastos depósitos de datos en MARC 21 sean más fácilmente intercambiables con herramientas y protocolos en XML. Ejemplos de estos protocolos son: el protocolo de la Iniciativa para Archivos Abiertos (OAI), el nuevo protocolo de búsqueda y recuperación (SRW/SRU) basado en Z39.50, y el Estándar de Codificación y Recuperación de Metadatos (METS). Usando estas herramientas de transformación en la Biblioteca del Congreso, las respuestas a búsquedas Z39.50 y SRW ahora pueden ser enviadas en formatos varios, incluyendo MARC 21, MARCXML, MODS, y Dublin Core. Estamos trabajando con colegas daneses para establecer un estándar ISO para la representación, en forma coherente con MARCXML, de cualquier dato ISO 2709 en XML.

No puedo abandonar el área de acción de MARC 21 sin antes mencionar la herramienta de MARC 21 a FRBR que hemos desarrollado y ofrecemos en la página Web de MARC 21 a usuarios interesados en experimentar con FRBR. FRBR es un concepto importante salido de IFLA y mantenido por otro participante de la alianza ICABS, la Biblioteca Británica. Esta es simplemente una herramienta – no proclama ser una herramienta FRBR completa – pero puede ser muy útil para evaluar las posibilidades de empleo de conceptos FRBR y la coherencia y potencial de los datos que uno posee. Está diseñada para operar sobre un juego de registros recuperados de una búsqueda, siendo por lo tanto, una herramienta para convertir a FRBR posterior a la búsqueda.

Metadatos

Quisiera examinar dos actividades de este año último. La primera está conectada con la iniciativa MARC 21 ya que el desarrollo del Estándar de Descripción de Objetos de Metadatos (MODS) surgió de una necesidad frecuentemente expresada de un formato XML para recursos

electrónicos que pudiera ser usado para facilitar el procesamiento de registros para dichos recursos. MODS tiene etiquetas basadas en lenguaje (en lugar de numéricas), algunas características especiales para recursos electrónicos, y buena compatibilidad con MARC 21. A MODS le falta la minuciosidad de MARC pero se aproxima en las áreas fundamentales. No reemplaza a MARC para aplicaciones que requieren el grado de minuciosidad para descripción y recuperación que brinda MARC 21.

Puede ser utilizado como una visión alternativa de datos en MARC para ítems de conversión digital que ya han sido catalogados para el catálogo de la biblioteca usando MARC 21. Un subconjunto de esos metadatos puede ser necesario para un repositorio que archive y disemine la versión del ítem como recurso electrónico. Es también un formato de fácil uso para el ingreso original de recursos electrónicos que han “nacidos digitales,” y particularmente útil para recursos de gran volumen como las páginas Web. El formato tiene una capacidad “recursiva” importante que permite la inclusión jerárquicamente relacionada de información sobre los recursos electrónicos. El lenguaje, tipo de letra, y transcripción pueden ser indicados para cada elemento de datos en MODS.

Otra cualidad valiosa de MODS es su acompañante más reciente: el Estándar de Metadatos para Descripción de Autoridades (MADS). MADS se relaciona con el formato de Autoridades MARC 21, de la misma forma que MODS se relaciona con el formato Bibliográfico MARC 21, pero simplifica datos y se asocia a MODS en formas novedosas y útiles – influenciado por el desarrollo de FRBR. MADS facilita el registro de formas autorizadas y formas de consulta de nombres para entidades básicas como nombres, títulos, materias geográficas, de género, materias de tópico y materias temporales.

Vale la pena tener en cuenta que el desarrollo de ambos formatos ocurre en el Internet, por medio de un listserv abierto. En ambos casos la Biblioteca del Congreso lanzó el formato original y luego se dedicó a editar y a diseminar las funciones por medio de una página Web.

El segundo progreso realmente importante en el ámbito de los metadatos en el transcurso de estos últimos años es el desarrollo de METS. METS puede ser denominada como una técnica para la grabación, en forma altamente flexible, de metadatos relacionados con todos los aspectos de un objeto digital. Esto incluye los metadatos descriptivos en que usamos MARC y MODS para codificar y también técnicos, de preservación, administrativos, estructurales, etc. Los paquetes de METS para un recurso electrónico incluyen o apuntan a estos metadatos y especifican el formato de los metadatos, aumentando la flexibilidad. Por supuesto flexibilidad trae aparejado un costo y los usuarios originales de METS están definiendo también perfiles para distintos tipos de recursos que aumentará el potencial para su interoperabilidad.

Por último, quiero destacar un recurso que tal vez les sea útil cuando exploren entre los distintos lenguajes de expresión de derechos en busca de derechos de metadatos. Encargamos un informe para examinar los más prominentes en la actualidad (principios del año 2004), principalmente ODRE, METSrights, Creative Commons y MPEG-21/5. Sus metas subyacentes y premisas fueron examinadas y fue establecida una clasificación sistemática de evaluación para su uso.

Recuperación de Información: Z39.50 y SRW

En la actualidad, el protocolo para recuperación de información Z39.50 se encuentra ampliamente divulgado en los sistemas de bibliotecas. Sin embargo, el progreso reciente más interesante en esta área, el “Z39.50 de próxima generación” de esta tecnología adaptada al ámbito de la Web, es SRW y SRU. SRU (Servicio URL de Búsqueda y Recuperación por sus siglas en inglés) permite a los usuarios hacer una búsqueda usando título, nombre, identificador y otros parámetros vía un URL (http GET) y obtener registros como respuesta. SRW (Servicio Web de Búsqueda y Recuperación por sus siglas en inglés) mantiene un servicio Web que opera en la Web basado en protocolo, el SOAP, en lugar de hacerlo directamente a través de HTTP. Es más robusto que SRU y puede ir directamente sobre HTTP.

Estos protocolos de búsqueda y recuperación basados en XML son semánticamente compatibles con Z39.50, adaptando las partes más útiles de este último protocolo al ámbito Web. Esto permite a los usuarios proveer portales SRW/SRU a ya existentes aplicaciones Z39.50. Desarrollado por un grupo editorial internacional, examinado a través de un listserv abierto, la versión 1.1 fue lanzada a principios de este año (reemplazando a 1.0, una versión experimental lanzada en el año 2002). Ya ha sido incorporada en algunos productos, e inclusive, se puede acceder a la institución a la que pertenezco, la Biblioteca del Congreso, por medio de un portal SRW/SRU.

Este emprendimiento resultó también en el desarrollo de un Lenguaje Común de Consulta (CQL por sus siglas en inglés). CQL intenta zanzar la brecha existente entre los distintos métodos de lenguajes de búsqueda: aquellos que son potentes pero complejos, crípticos, y difíciles de usar, como el XQuery o SQL; y aquellos como Google que son simples y fáciles de usar pero de limitada funcionalidad.

SRW puede ser combinado con OAI y OpenURL generando herramientas potentes y flexibles para acceder a recursos en la Internet.

Identificadores Persistentes

La Biblioteca del Congreso se encarga de observar la evolución de identificadores persistentes. Tenemos planeado formar un equipo con la Biblioteca Nacional de Australia, que tiene un buen sitio Web con información sobre Identificadores y ofrece un sitio Web complementario en LC. En el sitio Web LC/ICABS, pondremos información actualizada y descripciones de actividades recientes en W3C y IETF. Tendremos un hipervínculo apuntando al sitio Web australiano PADI ya que allí se encuentra mucha de la información disponible sobre este tema.

Cuando comenzamos a definir nuestro rol en esta área inmediatamente vimos que era necesaria la colaboración con otro sitio Web de ICABS.

Un avance importante este último año en este campo fue el del “info” URL. Identificadores son un tema complejo, pero hay tres características claves que la comunidad busca en ellos – persistencia, practicidad, y unicidad – y por ahora los intentos por lograr conjugar todas estas en un identificador han fracasado. El “info” URL está diseñado para facilitar el desarrollo de identificadores que satisfagan el criterio de unicidad y también para facilitar su asignación

autorizada para las comunidades editorial y de bibliotecas. Al focalizarse en el criterio de unicidad, el “info” URL puede ser utilizado para consultar recursos de tipo digital o no, esquemas, elementos de datos, códigos, etc. Surgió como un recurso para OpenURL pero también fue reconocida su funcionalidad que permite a los sistemas de identificación de legado formar parte de la arquitectura global de la red. Muchos de estos identificadores han sido registrados como “namespaces” bajo el “info” URL, incluyendo el DOI, LCCN (Library of Congress Control Number), entre otros conocidos identificadores.

VIAF

Fui encomendada por Die Deutsche Bibliothek (DDB) para reportar sobre la actividad del Archivo Virtual Internacional de Autoridades (VIAF), que es una responsabilidad que comparte con la Biblioteca del Congreso en ICABS. Esta actividad explora un modelo potencial para un “VIAF.” En este último año, se firmó un acuerdo entre los asociados en el proyecto, que incluye OCLC, y se intercambiaron archivos para registros coincidentes. Varias rondas de refinamiento de los algoritmos coincidentes serán necesarias, seguidas de la integración de las correspondientes registros de autoridades de nombres personales correspondientes y verificación de las coincidencias por medio del control de los registros bibliográficos. Estos registros pertenecientes a DDB, LC, y Bayerischer Bibliotheksverbund (BVB) serán utilizados para construir un VIAF inicial que concuerde con este modelo. Se espera completar esta primera fase en el 2004. De tener éxito, se procederá a facilitar el acceso al “server” del archivo VIAF experimental vía protocolos OAI y a través de un conector (interface) multilingüe.

Traducción de Fernando C. Alvarez
Library of Congress