



MINISTERIO  
DE EDUCACIÓN, CULTURA  
Y DEPORTE



BIBLIOTECA  
NACIONAL  
DE ESPAÑA

## **INFORME ALA MIDWINTER MEETING 2015**

**Chicago, 30 de enero – 3 de febrero de 2015**

**María Jesús Morillo Calero**  
Directora de Departamento de Proceso Técnico  
Biblioteca Nacional de España



## Índice

1. Introducción .....	4
2. JANEATHON: Fun with Dick and Jane (and RDA): Creating Linked Data for Jane Austen and Blade Runner .....	5
3. MARC ADVISORY COMMITTEE 1 .....	7
4. CATALOGING NORMS INTEREST GROUP .....	9
4.1. Utility of the Legacy MARC Records: Analyses and Ways to Support Discovery Services / Eric Radio, University of Kansas .....	9
4.2. Ready...Set...URIs... Actionable! --Experiencing a Linked Data Mindset Now. --Preparing For the Next Generation Search and Discovery Environment / Jackie Shieh, Dan Tam Do, George Washington University Libraries .....	9
4.3. MarcMaker: A Metadata Creation Tool / Tricia Lampron, University of Illinois .....	11
5. CATALOG MANAGEMENT INTEREST GROUP .....	13
5.1. OCLC Hybrid Records / Cynthia M. Whitacre, Manager, WorldCat Quality & Partner Content Dept., WorldCat Quality Management Division, OCLC .....	13
5.2. RDA Enrichment at the University of Colorado Boulder / Paul Moeller, Director of Metadata Services; Nate Cochran , Backstage Library Works .....	14
5.3. Partly Cloudy: Collaborating to Achieve Full Catalog RDA Enrichment in a Cloud-Based LMS / Stacey Traill, University of Minnesota .....	15
5.4. Roles and Reels: RDA and Relationship Designators for Moving Images / Kelley McGrath, University of Oregon .....	15
6. MARC FORMATS TRANSITION INTEREST GROUP .....	18
6.1. Experiments in BIBFRAME: A Modular Approach / Nancy Fallgren, Metadata Specialist Librarian, National Library of Medicine .....	18
6.2. BIBFLOW: An IMLS Project / Xiaoli Li, Co-head of Content Support Services Department, University of California Davis .....	19
7. LINKED LIBRARY DATA INTEREST GROUP .....	21
7.1. Linked Data for Libraries Project / Nancy Lorimer, Interim Head of Metadata Department at Stanford University Libraries .....	21
7.2. Opening up Science with VIVO / Kristi Holmes, Director of Galter Health Sciences Library at Northwestern University .....	21
7.3. BIBFRAME: The Way Forward for Library Visibility on the Web / Victoria Mueller, Senior Information Architect and System Librarian, Zepheira .....	22
7.3.1. Linked Data for Libraries (LD4L) / Stanford University Libraries .....	24
8. LC BIBFRAME UPDATE FORUM .....	25
8.1. Introduction: BIBFRAME Developments/ Sally H. McCallum .....	25
8.2. Library of Congress BIBFRAME Pilot Plans/ Beacher Wiggins .....	25
8.3. Pilot Component: BIBFRAME Editor/ Paul Frank .....	25
8.4. Pilot component: Profiles and Search/Display / Nate Trail .....	26
8.5. Le Boeuf sur le Toit: A Surrealistic Exploration of BIBFRAME for Production / Philip Schreur .....	26
8.6. OCLC and LC Collaboration / Ted Fons, OCLC .....	27



8.7.The Libhub Initiative: Accelerating / Eric Miller, Zepheira .....	27
9. CaMMS FORUM: BIBFRAME: NOTES FROM THE FIELD .....	29
9.1.BIBFRAME Development / Angela Kroeger, Archives and Special Collections Associate, University of Nebraska at Omaha .....	29
9.2.BIBFRAME and Beyond... Linked Data Projects in Stanford University Libraries / Nancy Lorimer, Interim Head of Metadata Department at Stanford University Libraries .....	30
9.3.Thinking in BIBFRAME... Not so Far Fetched After All / Jackie Shieh, Coordinator Resource Description Group, George Washington University Libraries .....	31
10. MARC ADVISORY COMMITTEE 2 .....	33
11. HEADS OF CATALOGING INTEREST GROUP.....	35
11.1. GW'S BIBFRAME Cookbook / Jackie Shieh, Georges Washington Libraries .....	35
11.2. BIBFLOW: An IMLS Project / Xiaoli Li, Co-head of Content Support Services Department, University of California Davis .....	36
12. SAC INTEREST GROUP .....	37
12.1. What is an RDA Subject / Gordon Dunsire .....	37



## 1. Introducción

El presente informe pretende dar cuenta de los múltiples y diferentes eventos a los que he asistido en el Encuentro de Invierno de la Asociación Americana de Bibliotecas (ALA), en torno a los temas de Proceso Técnico. Dada la abundancia de conferencias, comités y grupos de interés que se producen en cada conferencia de la ALA, intenté seleccionar aquellos de más utilidad entre los múltiples celebrados en torno a los “Technical Services”. La decisión en muchos casos ha sido difícil, pues la coincidencia en el tiempo de varios actos me ha impedido asistir a algunos y he debido inclinarme por la opción aparentemente más interesante y enriquecedora, teniendo en cuenta, fundamentalmente, los intereses de la Biblioteca Nacional de España en estos temas. Asimismo recoge la información sobre las reuniones del MARC Advisory Committee que tuvieron lugar el sábado 31 de enero y el domingo 2 de febrero, en el que la BNE está representada.

A lo largo de las diferentes jornadas, desde el Janeathon a las diferentes ponencias sobre Datos Enlazados, se ha insistido en la necesidad de poner el acento sobre las relaciones entre recursos y los enlaces entre ellos antes que sobre la descripción. Debemos comenzar a pensar de forma diferente, la importancia ya no está en nuestros registros bibliográficos, sino en las relaciones entre las diferentes entidades WEMI (Work, Expression, Manifestation, Item). Como señalaron tanto Eric Miller como Victoria Mueller, de la empresa Zepheira, “un enlace vale más que mil palabras”.

El cambio hacia un entorno de datos enlazados va a implicar, no ya compartir datos, sino agregar datos procedentes de otras comunidades e instituciones como archivos y museos. Debemos hablar de metadatos y no de catalogación, señalaba Diane Hillmann, ya que tenemos que avanzar hacia ese mundo de agregación de datos y aceptar la diversidad de fuentes de donde se pueden extraer dichos datos. Debemos abandonar la idea de registro para comenzar a pensar en términos de identificación de recursos.

En diferentes foros surgió la problemática de la verdadera internacionalización de RDA como estándar de catalogación. En líneas generales, se está de acuerdo en que no se debe pensar en términos de qué nuevos países o instituciones se suman a la utilización del estándar, sino que se debe trabajar con el compromiso adquirido de alcanzar una verdadera multiculturalidad. No se puede seguir considerando el inglés como la “lingua franca”, es preciso incorporar lenguas y escrituras, entre otras muchas cosas. La aspiración de convertirse en un estándar para la catalogación de todo tipo de materiales independientemente del tipo de institución que los custodie también se puso en entredicho. Es más, se lanzó la pregunta a personal de la Library of Congress de si RDA era estándar más adecuado de cara a un contexto de datos enlazados y el propio Beacher Wiggins, director de Adquisiciones y Acceso Bibliográfico de la Library of Congress, contestó que no era la única.

Asimismo, se ha puesto de manifiesto que la experimentación con el modelo BIBFRAME está tomando diferentes vías en las distintas instituciones que están trabajando con él, incorporando otros modelos o vocabularios reconocidos, dando una nueva perspectiva, como por ejemplo en el caso de la National Library of Medicine, con un enfoque más modular. BIBFRAME es una obra en continuo proceso (“work in progress”), no un modelo cerrado. La experimentación es básica para conseguir alcanzar resultados óptimos.



**Viernes, 30 de Enero de 2015**

## **2. JANEATHON: Fun with Dick and Jane (and RDA): Creating Linked Data for Jane Austen and Blade Runner**

El Viernes 30 de enero se celebró un evento, que se prolongó desde las 8,30 h hasta las 17 h., cuyo fin era catalogar colectivamente recursos relativos a obras de Jane Austen utilizando un software diseñado específicamente para el estándar RDA: Resource Description and Access. El programa era [RIMMF3](#) (RDA in Many Metadata Formats) que permite crear registros RDA puros, así como importar registros MARC21, editarlos y exportarlos como datos enlazados. El objetivo era explorar el trabajo con declaraciones RDF (Resource Description Framework) para contribuir al mundo de los datos enlazados.

La agenda del evento incluía: una demostración del trabajo con RDA y RIMMF3 para la descripción de recursos y obras relacionadas con la película Blade Runner, la historia original de Philip K. Dick y la música creada por Vangelis para la película; trabajar en pequeños grupos sobre recursos relacionados con las obras de Jane Austen; exportar datos RDA como RDF y comprobar cómo los datos pueden agregarse a un conjunto de datos enlazados con tripletes relativas a Jane Austen; discusión sobre el trabajo con RIMMF3 y RDA, sobre los temas a los que se enfrentan las bibliotecas en el entorno de datos enlazados y catalogación compartida, así como sobre el trabajo con recursos relacionados con Jane Austen y cómo ampliarlos y extenderlos.

No se pretendía enseñar RDA, sino comprender las posibilidades que ofrece la definición de relaciones entre recursos en un entorno de datos enlazados. Con carácter previo al encuentro se había pedido a las personas interesadas en asistir que completaran toda una serie de tutoriales para aprender a trabajar con RIMMF3 y se había desarrollado un webinar online.

El “Janeathon”, como fue denominado siguiendo la estructura y designación de los hackatons tan extendidos en la actualidad, comenzó con una presentación de Gordon Dunsire sobre Blade Runner y cómo había comenzado a trabajar en la creación de su primer R-ball (RIMMF-Ball = grupo de recursos sobre un determinado tema).

Lo primero que hizo fue acudir al catálogo de la Biblioteca Nacional de Escocia, en la que trabaja, para buscar recursos relacionados con la película. Con la opción “Arrastrar y Soltar” de RIMMF3, que definió como asombrosa y que, en verdad lo es, importó registros MARC del catálogo al programa RIMMF3 que los convierte a RDA estableciendo las diferentes entidades WEM (Work, Entity, Manifestation). A continuación comenzó la misma búsqueda en el catálogo de la Biblioteca Nacional alemana que le permitió importar nuevos registros a RIMMF3. Mucha gente piensa en MARC, pero MARC no es RDA, por lo que es necesario comenzar a pensar en RDA fuera del entorno MARC.

Finalmente, expresó su convicción de que lo que estaba comenzando en la sala era algo muy serio, aunque recomendó a todos los presentes que se divirtieran. Y recordó que lo importante son las relaciones, no la descripción. Recomendó, asimismo, no pensar en si un elemento es obligatorio o no, si es “core” o no, sino más bien en si es útil o no.



A continuación los participantes, divididos en grupos por tipos de recursos y durante dos horas y media, comenzaron a trabajar partiendo de un R-ball con 75 entidades relacionadas con obras de Jane Austen, junto a algunas expresiones y manifestaciones. Los asistentes importaron registros, catalogaron recursos desde cero y establecieron relaciones entre ellos con el programa RIMMF3. Partiendo de esas 75 entidades, después de dos horas y media, el R-ball había aumentado a 126 Obras, 117 Expresiones, 134 Manifestaciones, 240 Personas y 64 entidades corporativas. En cada mesa un asistente ayudaba a los participantes a resolver dudas, ofrecer soporte y comentar los casos. El RIMMF-ball resultante tras el trabajo fue convertido a RDF para trabajar como datos enlazados.

Una vez terminada la parte práctica del evento, Deborah Fritz, desarrolladora con su marido del programa RIMMF3, comentó los resultados del trabajo e inició la discusión sobre los problemas encontrados y temas surgidos. Habló de los problemas con la creación de Expresiones de las que solo se conoce la lengua, pero no otros elementos propios de la expresión como la fecha, traductores, etc. Quizás sería necesario crear registros genéricos de expresiones con lenguas, de las que no se conozcan otros datos.

Tras la intervención de Deborah, Diane Hillman habló sobre RDA y Datos enlazados. Empezó hablando de que no hablamos sobre catalogación, sino sobre metadatos y lo complicado que va a ser para nosotros compartir datos con otras comunidades como museos o archivos. Es más, ni siquiera debemos hablar de compartir datos, sino de agregar datos. En el proceso de comenzar a mover RDA a otro mundo de agregación de datos, tenemos que aceptar la diversidad y vivir con ella. No pensar en términos de lo que está bien o está mal, sino en la capacidad de agregar datos provenientes de diferentes fuentes. El registro nos matará si no abandonamos la idea del registro. Necesitamos pensar en lo que significa la identificación de un recurso. Nuestros registros de autoridad no identifican personas, sino registros.

Es necesario que integremos vocabularios válidos y no solo vocabularios RDA. Es necesario extender RDA a dominios específicos. Hay muchísimos recursos maravillosos que pueden ser utilizados, como el tesoro Getty, por poner un ejemplo.

Respecto a la internacionalización de RDA es necesario dar viabilidad a otras lenguas diferentes del inglés. No se trata de que poco a poco nuevos países se vayan sumando a la adopción de RDA, se trata de que no es posible considerar la lengua inglesa como la lengua franca. RDA debe comprometerse realmente con la multiculturalidad.

Finalmente, Gordon Dunsire comentó los nuevos pasos a dar en relación a lo que había sido el Janeathon. Animó a utilizar la lista RDA-L para ampliar la discusión generada en el evento y la [web de los r-balls](#). Recomendó que si tienes más recursos que añadir al r-ball de Jane Austen, o bien quieres crear r-balls para otros temas, los exportes en ficheros de texto comprimidos y los envíes a [jschair@rdatoolkit.org](mailto:jschair@rdatoolkit.org). Animó a divertirse todavía más creando r-balls.



**Sábado, 31 de Enero de 2015**

### **3. MARC ADVISORY COMMITTEE 1**

[http://www.loc.gov/marc/mac/mw2015\\_age.html](http://www.loc.gov/marc/mac/mw2015_age.html)

Abrió la reunión del comité MAC (MARC Advisory Committee) su presidente, Matthew W. Wise, bibliotecario de la Universidad de Nueva York. Tras la presentación de sus miembros, se procedió a la aprobación de las actas del último encuentro de MARBI de junio de 2013, que fueron sancionadas por unanimidad.

A continuación se pasó a la discusión y aprobación de varias de las propuestas presentadas:

#### **Proposal No. 2015-01: Defining Values in Field 037 to Indicate a Sequence of Sources of Acquisition in the MARC 21 Bibliographic Format**

La propuesta provenía de la British Library. Tras un debate previo, la propuesta es aprobada.

#### **Discussion Paper No. 2015-DP01: Recording RDA Format of Notated Music in the MARC 21 Bibliographic Format**

A continuación se debatió el único Discussion Paper que se presentaba en los dos encuentros del MARC Advisory Committee. La propuesta partía del Canadian Committee on MARC quien defendió la necesidad de incluir el formato de música notada que establece RDA en el formato bibliográfico.

Para las preguntas 1 (¿es el formato de música notada un elemento significativo para encontrar, identificar y seleccionar recursos consistentes en música notada?) y 2 (¿es útil una codificación precisa del formato de música notada para la recuperación basada en estos términos controlados?) la respuesta fue un sí unánime.

Respecto a la pregunta 3 (¿hay algún caso de estudio para incluir este elemento aparte de en los registros bibliográficos en los registros de autoridad también?), John Myers, del Committee on Cataloging: Description and Access (CC:DA) de ALA, opinó que debería ser aplicado a ambos formatos, autoridades y bibliográficos. Se puso de relieve también el trabajo que la Musical Association había llevado a cabo en este tema y la necesidad de armonizar las definiciones.

Respecto a la pregunta final que se refería a las opciones preferidas para la codificación, varios miembros del grupo destacaron su preferencia por el campo 348. Reinhold Lieuvemann, de la Biblioteca Nacional alemana, explicó que hay varias maneras de modelar este tipo de información, por lo que es necesario encontrar una solución flexible y trabajar en las definiciones. También se comentaron los temas concernientes a la forma en singular o plural.

Finalmente, se recomendó hacer una propuesta final para su codificación en el campo 348 con dos nuevos subcampos.



**Proposal 2015-04: Broaden Usage of Field 088 in the MARC 21 Bibliographic Format**  
(<http://www.loc.gov/marc/mac/2015/2015-dp04.html>).

Mathew W. Wise, chair del MARC Advisory Committee, introdujo la propuesta ya que al representante de la Alaska Resources Library and Information Service, que había enviado la propuesta, le había sido imposible acudir al Encuentro. International partners support the proposal.

Tras un debate previo la propuesta fue aceptada, pero con sugerencias de correcciones para los ejemplos.

**Proposal No. 2015-02: Adding Dates for Corporate Bodies in Field 046 in the MARC 21 Authority Format**

En el debate previo sobre la propuesta de la British Library de añadir fechas para entidades corporativas en el campo 046 del Formato de Autoridades se discutió sobre la necesidad de diferenciar entre las fechas de creación y fin de una entidad, de las fechas de actividad de la misma entidad. La representante de la National Library of Medicine indicó que en su institución algunas personas pensaban que esa primera fecha de actividad más temprana podría ser denominada "Non-date".

La propuesta fue aprobada.

**Proposal No. 2015-05: Definition of New Code for Leased Resources in Field 008/07 in the MARC 21 Holdings Format**

En el debate inicial sobre la propuesta presentada por la British Library se apuntó que, en la mayoría de las bibliotecas, este tipo de información se registraba en las licencias, no en la codificación MARC. Thustan Young, de la British Library, comentó que el párrafo final de la propuesta ya recoge que quizás en el futuro serán necesarios nuevos códigos, pero que la British Library había redactado la propuesta en este sentido, porque en la actualidad sí que necesitaban un código para este tipo de recursos.

La propuesta fue aprobada.

**Business meeting**

Sally McCallum comentó que veinticinco catalogadores de la Library of Congress habían estado trabajando directamente en BIBFRAME, pero, a pesar de ello, estos registros serían convertidos en registros MARC.



## 4. CATALOGING NORMS INTEREST GROUP

### 4.1. Utility of the Legacy MARC Records: Analyses and Ways to Support Discovery Services / Eric Radio, University of Kansas

En esta presentación se presentó un estudio sobre el grado de consistencia de los datos codificados en registros MARC usados en la biblioteca digital HathiTrust y la forma en la que los catalogadores usan los campos MARC a través del análisis de una muestra de registros MARC procedentes de la mencionada biblioteca digital. En un momento en el que el control bibliográfico y el papel que los metadatos juegan como apoyo al descubrimiento de materiales digitalizados, es necesario analizar la consistencia de la información.

Las primeras cuestiones que se plantearon al inicio del Proyecto fueron, fundamentalmente, de qué manera interpretan los catalogadores campos diseñados para los mismos tipos de información, qué sucede si se están utilizando de forma diferente y, por último, si se podría mejorar de forma sistemática la información consignada en dichos campos.

Entre los primeros hallazgos del proyecto encontraron que el porcentaje del uso del texto libre llegaba a un 91'46%, que había una gran consistencia en el uso de las posiciones 35-37 del campo 008 referidas a la lengua del documento o que el campo 041 en el que se incluyen códigos de lenguas cuando el código de lengua del campo 008/35-37 del registro es insuficiente para dar una información completa, como en documentos multilingües, o que incluyen una traducción, o documentos en los que el medio de comunicación es un lenguaje de signos, no se usaba con mucha frecuencia.

Los principales hallazgos a partir de análisis automáticos fueron: el solapamiento de información en múltiples campos, inconsistencias en el control de la puntuación y en el control bibliográfico, falta de granularidad y uso de vocabularios controlados. Como conclusiones, por tanto, se vio la necesidad de reducir las redundancias, incrementar la consistencia y dotar de una mayor normalización.

Se han analizado seis millones de registros de HathiTrust en el desarrollo del Proyecto.

### 4.2. Ready...Set...URIs... Actionable! --Experiencing a Linked Data Mindset Now. --Preparing For the Next Generation Search and Discovery Environment / Jackie Shieh, Dan Tam Do, George Washington University Libraries

Las bibliotecas de la Universidad George Washington han sido de las primeras en convertirse en experimentadores con BIBFRAME y han introducido el uso del \$0 para expresar el concepto de enlace con un punto de acceso autorizado. En su presentación, Jackie Shieh y Dan Tam Do comentaron los modos en que lo han utilizado para fijar la formación del personal en los conceptos de datos enlazados, en torno a identidades, vocabularios, "cosas y no cadenas de caracteres" y más concretamente los beneficios de utilizar URIs procesables suministradas por profesionales de la información. Por otra parte, hablaron sobre las técnicas y herramientas que se pueden utilizar para



incorporar URIs a los flujos de catalogación cotidianos tanto en el momento de la catalogación como en los procesos batch.

Los datos enlazados han ido ganando relevancia en la comunidad bibliotecaria. Cada vez se llevan a cabo más proyectos en diferentes áreas geográficas en este sentido y cita, junto a los proyectos de la British Library o la Biblioteca Nacional Alemana, el proyecto de datos enlazados de la Biblioteca Nacional de España, [datos.bne.es](http://datos.bne.es).

El primero de los cuatro principios definidos por Tim Berners-Lee en relación a los datos enlazados establece que se deben “utilizar URIs (Uniform Resource Identifiers) como nombres para Cosas” (Use URIs as names for Things). En este sentido, un recurso es una de esas “Cosas” que tiene una identidad. ¿Por qué debe ser uniforme? Porque la uniformidad facilita la interpretación semántica.

Pero vivimos en un mundo global en el que es indispensable tener en cuenta diferentes lenguajes y escrituras y resulta fundamental poder reutilizar un identificador en diferentes contextos. En este sentido y dado que las URIs están limitadas a un subconjunto del juego de caracteres ASCII, es más conveniente el uso de IRIs (Internationalized Resource identifiers), estándar definido por la Internet Engineering Task Force (IETF) en 2005 para ampliar las URIs con caracteres correspondientes al Juego Universal de Caracteres (Unicode/ISO 10646), y que incluye otras lenguas y escrituras como el chino, japonés, cirílico, etc.

El segundo principio de Tim Berners-Lee se refiere a la capacidad de procesado y señala que es preciso aprovechar el HTTP de las URIs para que la gente pueda localizar y consultar esos recursos. En este sentido, un identificador como, por ejemplo, (OCLC)fst00903451 no nos permite localizar y consultar el recurso, mientras que uno del tipo <http://id.worldcat.org/fast/903451> sí. El potencial es enorme, sobre todo, si contamos con herramientas de rastreo de versiones. En el entorno actual ya se han alcanzado estos dos primeros principios.

En su proyecto de datos enlazados han utilizado un subcampo, el \$0, que consideran poderoso. Le han otorgado un carácter repetible y lo han hecho presente en X00, X10, X11, X30, 240, la mayor parte de los 65X y otros campos. Para el proceso en batch han utilizado la herramienta MarcEdit, aunque señalaron que también se podrían utilizar otro tipo de herramientas para ello como Perl. La actualización de registros se ha realizado utilizando URIs proporcionados por [id.loc.gov](http://id.loc.gov), FAST y VIAF.

Los costes se han referido, sobre todo, al tiempo que es necesario invertir: para aprender, para ejecutar nuevas acciones, para llevar a cabo el control de calidad, para realizar procesos de edición adicionales. A todo ello se suma la falta de resultados visibles inmediatos. Pero los beneficios que se han obtenido han sido: un mejor conocimiento de MARC, una mayor familiaridad con [id.loc.gov](http://id.loc.gov), FAST y VIAF, una mejora en el conocimiento y utilización de la herramienta MarcEdit, así como una mayor apertura de miras y un mejor entendimiento de la diferencia entre “cosas” (recursos) por oposición a “cadenas de caracteres”.

El primer paso para ellos era extraer los datos de forma que su sistema pudiera mostrarlos de una forma adecuada.



Por ahora, OCLC no proporciona información sobre las actualizaciones de los registros. BIBFRAME está estudiando llevar a cabo cambios en las relaciones para emitir notificaciones informando de cambios y permitiendo aceptarlos o no en tus propios registros.

### **4.3. MarcMaker: A Metadata Creation Tool / Tricia Lampron, University of Illinois**

Tricia Lampron comenzó su presentación con un resumen de los recursos con los que cuenta la Biblioteca de la Universidad de Illinois en Urbana-Champaign (en adelante, UIUC): veintisiete bibliotecas más una especializada en estudios internacionales y más de trece millones de volúmenes. Cada vez son más frecuentes las donaciones de materiales a la biblioteca y el número de recursos a gestionar crece de forma acelerada. La necesidad de hacer disponibles al público dichos recursos en el menor tiempo posible, mejorar el flujo de trabajo en la catalogación, incluir más áreas geográficas y lenguas, y la gestión de metadatos y visibilizar las colecciones que permanecen ocultas a los usuarios, junto a la falta de personal formado, llevaron a la UIUC a poner en marcha un proyecto denominado “MarcMaker” para la creación de una aplicación web que sirviera para facilitar y acelerar el mencionado flujo de trabajo, que pudiera proporcionar registros de calidad para materiales que necesitan ser catalogados desde cero.

Existían diferentes casos. En primer lugar, había materiales que necesitaban registros de catalogación de nivel mínimo en MARC; otros proyectos requerirían conversión a otros esquemas de metadatos partiendo de MARCXML; para los materiales nacidos digitales como, por ejemplo, tesis doctorales, habría que tener en cuenta el esquema MODS; finalmente, en el entorno de datos enlazados, sería necesario contar con HTML en relación a Schema.org.

La aplicación está basada en HTML, CSS y Javascript. Se puede elegir entre varios formatos de exportación: MARC/MARCXML, MODS, Schema.org, para lo que se ha trabajado en el mapeo de estos diferentes estándares.

Una vez comprobado que el sistema era estable, seis personas con diferente experiencia en catalogación trabajaron en los ensayos y crearon registros de monografías en varias lenguas y en todos los formatos utilizando MetadataMaker para crear los registros.

El flujo de trabajo con los registros MARC fue el siguiente: se subían registros a OCLC Connexion que eran revisados, se añadían encabezamientos de materia y clasificación basados en las palabras clave asignadas por el usuario, se añadían a OCLC, se exportaban al catálogo local y seguían el proceso normal.

Los problemas que se encontraron se referían a errores de codificación, el nombre siempre estaba mapeado a “autor”, no había apoyo para las entradas con diacríticos, era necesaria información sobre la publicación en los casos en que no estaba identificada, había problemas con la complejidad de los datos en los campos 264 o la necesidad de recontar aquellos materiales no paginados.

Respecto a la formación, no hubo una formación previa. Los problemas más frecuentes se refirieron a la fuente preferida de información en el caso de los títulos y la información sobre la



publicación; las fechas, si se tomaban del copyright o de la publicación; los números de páginas y dimensiones; las palabras claves y los tipos de notas, como de series, otros títulos y otras lenguas.

Se prevén mejoras relativas al desarrollo de menús desplegables para seleccionar roles, cajas de comprobación para bibliografías e índices, apoyo para diferentes alfabetos y para su romanización, así como servicios API en relación a encabezamientos de materia FAST y enlazado a VIAF. Más información sobre el proyecto se puede encontrar en:

MetadataMaker GitHub: <https://github.com/dkudeki/metadata-maker>.

MetadataMaker Website: <http://iisdev1.library.illinois.edu/marcmaker>.



## 5. CATALOG MANAGEMENT INTEREST GROUP

### 5.1. OCLC Hybrid Records / Cynthia M. Whitacre, Manager, WorldCat Quality & Partner Content Dept., WorldCat Quality Management Division, OCLC

Cynthia M. Whitacre, de OCLC, comenzó hablando de porqué OCLC cuenta con registros híbridos y en qué consisten, ya que la cantidad de estos registros crece a diario en OCLC.

Si buscas en Google “hybrid record”, lo primero que te aparece es este enlace: <http://piaptk.limitedrun.com/products/535023-pre-order-cd-record-series-digital-analog-hybrid-discs>. Pero no es de esto de lo que quería hablar sino de los registros bibliográficos que tienen elementos RDA en ellos, no son registros RDA en sentido estricto y no están codificados como registros RDA.

¿Por qué existen registros híbridos en OCLC? Tiene que ver con la reducción de la necesidad de que los catalogadores dominen varios tipos de reglas, con la reducción de costos y con hacer que dichos registros sean funcionales en un entorno de catalogación con el estándar RDA. Si nos encontramos con un registro anterior a la época de AACR2, ¿conocemos las reglas con las que fue creado o nos tomamos libertades para adaptarlos a las nuevas necesidades de nuestro catálogo? Los registros híbridos simplifican los procesos de catalogación y reducen costes al hacer cambios de forma automática sin el recurso en la mano. Necesitamos estar seguros de que todos los registros antiguos son tan útiles como se pueda en términos de indización, recuperación y presentación de los datos.

OCLC utiliza las [PCC Guidelines on Hybrid Bibliographic Records](#), de octubre de 2012) para todos los registros bibliográficos presentes en WorldCat, así como la [Declaración de OCLC sobre RDA](#) que se hizo efectiva a partir del 31 de marzo de 2013. En este sentido, OCLC quiere eliminar todas las designaciones generales de material, puesto que no forman parte de RDA, pero puesto que la política PCC señala que no se deben eliminar al hacer registros híbridos, están actuando de esa manera y no lo harán hasta abril de 2016, tres años después de la implementación de RDA, tal y como señala PCC.

Expuso algunos ejemplos de cómo se está trabajando con estos registros en temas como el desarrollo de abreviaturas, el añadido de los campos 33X mostrando de una forma apropiada los términos no ingleses al usuario final, la sustitución del símbolo del copyright cuando aparezca tan solo una c delante de fechas de copyright, cambios en los puntos de acceso, etc.

Explicó que el personal de calidad de Worldcat está trabajando en la hibridización de todos los registros. Este personal normalmente tiene tres sesiones abiertas con Connexion a la vez que procesan un total de 10.000 registros cada una, incluyendo mejoras como el añadido de los campos RDA. También se corrigen aquellos registros con errores de validación.



Los registros bibliográficos híbridos son poderosos, señaló, porque conducen al futuro. Si contamos con datos consistentes será más fácil convertirlos a datos enlazados y sistemas futuros que no usen ya MARC para la codificación.

Para cualquier duda o comentario ofreció las siguientes direcciones de correo electrónico: [whitacrc@oclc.org](mailto:whitacrc@oclc.org), [rdapolicy@oclc.org](mailto:rdapolicy@oclc.org), [askqc@oclc.org](mailto:askqc@oclc.org).

## **5.2. RDA Enrichment at the University of Colorado Boulder / Paul Moeller, Director of Metadata Services; Nate Cochran , Backstage Library Works**

Paul Moeller comenzó dando algunos datos sobre las bibliotecas de la Universidad de Colorado Boulder: cuatro millones de registros bibliográficos, un millón de registros de autoridad, aplican las políticas PCC (NACO, BIBCO, CONSER) y utilizan Innovative Interfaces Sierra (1.2.2\_6).

Desde 2008 trabajan con Backstage Library Works en el control de autoridades y la limpieza de la base de datos. Con ellos han trabajado en el enriquecimiento de unos tres millones de registros bibliográficos. Pero, ¿qué razones llevan al enriquecimiento de registros? Con la implementación de RDA como estándar para la descripción, las bibliotecas comienzan a plantearse qué hacer con sus registros previos que ya tienen en sus catálogos locales y a plantearse la necesidad de hibridarlos o enriquecerlos. Los avances en el proyecto BIBFRAME, que tiene como objetivo reemplazar a MARC, se considera que es necesario tener datos consistentes para desarrollar sistemas de datos enlazados que sustituyan a los actuales sistemas integrados de bibliotecas.

En mayo de 2014 se formó un grupo para revisar la planificación del proyecto. Los miembros del grupo fueron dos personas de Backstage Library Works, dos bibliotecarios catalogadores y un bibliotecario experto en recursos musicales. Entre los pasos dados se han añadido los campos 33X a los registros y se han realizado procesos de conversión de 260 a 264 que han tenido algunos problemas ya que no siempre era fácil distinguir entre un editor y un distribuidor en los campos 260.

En agosto se procedió a la fase de comprobación con un archivo de 625 registros de libros, publicaciones periódicas, recursos musicales, mapas, grabaciones sonoras, materiales visuales y archivos de ordenador. Se detectaron algunos pequeños problemas, como que si la información en 006/007 es incorrecta es imposible aplicar de forma adecuada los campos 3XX. Es imposible convertir malos datos en buenos datos.

Se notificó a los catalogadores que no debían borrar o editar registros bibliográficos catalogados con anterioridad da agosto de 2014 hasta nueva orden, así como a las bibliotecas. En siete días se procedió a recargar todos los registros bibliográficos de nuevo actualizados y enriquecidos.

Nate Cochran, de la empresa Backstage Library Work, comentó que ofrecen enriquecimiento de registros RDA, que es gratuito para sus clientes. También comentó que de 2013 a 2014 el número de clientes interesados en enriquecimiento de registros con RDA ha pasado de un 48% a un 96%.



### **5.3. Partly Cloudy: Collaborating to Achieve Full Catalog RDA Enrichment in a Cloud-Based LMS / Stacey Trill, University of Minnesota**

La Universidad de Minnesota cuenta con 21 bibliotecas distribuidas en cinco campus. El sistema de gestión de bibliotecas utilizado por ellas entre 2002 y 2013 fue ExLibris de Aleph. A partir de finales de 2013 emigraron a Alma, sistema de gestión bibliotecaria en la nube. Previo a la migración, se habían planteado un proyecto de enriquecimiento de registros a RDA con la ayuda de la empresa Backstage Library Works, cuya planificación comenzó en abril del año 2013, para terminar totalmente en el otoño del mismo año.

Tanto lo relativo al control de autoridades como las elecciones de mejora descriptiva fueron realizados por un pequeño equipo de tres catalogadores y un bibliotecario de sistemas. En abril de 2013 comenzaron las primeras pruebas con unos diez mil registros correspondientes a todos los formatos y lenguas. La empresa Backstage Library Works sugirió algunas mejoras basadas en los resultados de estos tests iniciales.

En julio ya se disponía de un perfil para mejora y enriquecimiento de autoridades a RDA y en agosto se comenzó a trabajar. En septiembre se sugirieron mejoras a la empresa Backstage Library Works y el test final se realizó en octubre.

Mientras tanto se había producido un retraso en la migración a ALMA. Se trabajó en colaboración con personal de Ex Libris para desarrollar todo el proceso, pero seguía habiendo incertidumbre respecto a la escalabilidad de los procesos de exportación e importación, la indexación en Alma, etc. Se comenzó a trabajar con pequeños conjuntos de registros.

En relación al control de autoridades, no se hicieron cambios en los registros ya que los ficheros de autoridades se mantenían de forma centralizada por la empresa Ex Libris. Las opciones que se manejaron para el enriquecimiento de los registros se refirieron fundamentalmente a la descripción: se eliminaron las descripciones generales del material del \$h de los campos 245 (aunque fueron retenidas en un campo local), se actualizaron los campos 25X y 260, se convirtieron los campos 260 en 264 (este proceso lo han realizado aproximadamente un 85% de las bibliotecas académicas que han trabajado con la empresa Backstage Library Works), se actualizaron los campos 300, los nombres de lugares de congresos en X11, las abreviaturas se desarrollaron en las fechas, así como los términos de relación.

El Proyecto finalizó con el enriquecimiento de más de cinco millones y medio de registros. Los desafíos y resultados se pueden resumir en la dificultad para predecir el tiempo de ejecución de los procesos en ALMA, en la necesidad de limpiar y/o borrar una gran cantidad de registros que no se habían detectado en el proceso de la migración, la identificación también de registros con errores críticos de codificación MARC para su posterior depuración, etc. Los cambios que se han operado en los registros han otorgado mayor consistencia y claridad para los usuarios.

### **5.4. Roles and Reels: RDA and Relationship Designators for Moving Images / Kelley McGrath, University of Oregon**



Kelley McGrath explicó que su presentación estaba basada en observaciones realizadas después de analizar una larga lista de roles en registros MARC de imágenes en movimiento y los trabajos de OLAC en un proyecto para tratar de extraer datos más estructurados de registros bibliográficos ya existentes relativos a películas y videograbaciones (<http://olac-annotator.org/>). Uno de los objetivos era analizar campos con texto libre (245\$c, 260\$b, 264\$b, 508, 511) acerca de quién hizo qué en registros MARC para tratar de convertir esa información en datos que pudieran ser procesables de manera automática, crear un fichero de respuestas correctas y utilizarlo con fines de formación y evaluación.

Para ello decidieron contar con la colaboración del público a través de un formulario web. En una primera línea se añadía la información sobre la declaración de responsabilidad que se quería analizar. Por ejemplo, "written and directed by Charles Chaplin". Para desgranar esta información se contaba con diferentes cuadros de introducción de datos: uno para identificar la lengua del rol o de la función, otro para consignar el nombre tal y como aparecía en los créditos, otro para intentar enlazar dicho nombre con un 1XX o 7XX dentro del registro, otro para la función tal y como aparece en los créditos y otro para la traducción al inglés de esa función. En otro ejemplo, "kyakuhon, Tsuda Fumio, los datos quedarían expresados como:

Lengua de la función: Japonés  
Nombre en los créditos: Tsuda Fumio  
Rol o función: kyakuhon  
Traducción al inglés del rol/función: guión

Gracias a esta forma de trabajo están consiguiendo identificar roles según lenguas. Según sus datos, la distribución de roles sigue una proporción de un 80/20, ya que la mayor parte de los roles se incluyen dentro de un número pequeño de categorías que han sido recogidas de forma tradicional en los registros de imágenes en movimiento. A pesar de ello, existe un gran número de subcategorías dentro de cada una de dichas categorías. Por ejemplo, dentro del rol "productor" se pueden encontrar variantes o subcategorías como "productor ejecutivo", "co-productor", "productor asociado", "productor senior", "productor de video", etc.

Por otra parte, se han detectado muchos términos no ingleses que no se corresponden exactamente con un término inglés o utilizado en la industria cinematográfica anglo-americana. "Kikaku", por ejemplo, significa literalmente en japonés "planificación", pero no existe una palabra que se use de forma semejante en los créditos de películas norteamericanas. ¿Qué significa el término español "realización"? ¿Se refiere a la producción, a la dirección?

También existen dificultades en relación a los roles que se han transcrito en su base de datos de roles/funciones que pueden resumirse en "vagos", "ambiguos" e "incoherentes". Los primeros, los "vagos" son los más comunes. Por ejemplo, "responsable, Lê Mỳ Phu'o'ng", "concepto, Clive Sugars" o "supervisor, Dr. Nurdin Perdana". ¿Qué función exacta desempeñan?

La ambigüedad también es un problema. "By" puede referirse al autor de una obra de teatro (By Shakespeare), al compositor de una ópera (By W. A. Mozart), por ejemplo. "Music" puede referirse tanto al compositor como al ejecutante. "Songs", puede designar al autor de la letra, de la música o al intérprete.



Por último, existen muchas declaraciones de autoría completamente incoherentes como “An Old Photo film presents” o “un film von Xavier Koller”. Muchas de ellas probablemente son errores de transcripción, pero no todas.

¿Qué se nos escapa en RDA y MARC? La forma de relacionar nombres de actores con los personajes que representan, o el modo de relacionar roles con partes ya que hay muchos registros que describen más de una obra. Dado que no existe una forma normalizada de solucionar este tipo de problemas, los catalogadores utilizan todo tipo de artimañas para explicitar este tipo de relaciones, lo que dificulta que un ordenador pueda interpretar dicha información de forma automática. La discusión sobre estos temas no tiene fin en listas como RDA-L y el Joint Steering Committee for the Development of RDA continua ampliando la lista de designadores de relación, sin que se vea el final a dicho trabajo.

¿Cómo se podría solucionar el tema de las declaraciones de responsabilidad difíciles de interpretar? Las bibliotecas no son las únicas que se encuentran con este tipo de problemas. ¿Podría ser la solución la combinación utilizada por Internet Movie Database (IMDb) de categorías amplias con roles/funciones muy específicas en texto libre?

Ejemplo: **Romeo y Julieta (1968)**

Vocabulario controlado para términos genéricos:

William Shakespeare (**Teatro**)

Franco Zeffirelli (**Director**)

Franco Brusati (**Guión**)

Otros miembros del equipo:

Leonard Whiting ... **Romeo**

Olivia Hussey ... **Juliet**

John McEney ... **Mercutio**

Las bibliotecas suelen hacer también este tipo de cosas, pero lo que no hacen y sí IMDb es conectar este tipo de información de forma que sea procesable de forma automática.

Finalmente, Kelley McGrath animó a participar en el proyecto acudiendo a su web: <http://olac-annotator.org>. Y también participando en su lista de discusión: <https://lists.uoregon.edu/mailman/listinfo/olac-credits>.



## 6. MARC FORMATS TRANSITION INTEREST GROUP

### 6.1. Experiments in BIBFRAME: A Modular Approach / Nancy Fallgren, Metadata Specialist Librarian, National Library of Medicine

La National Library of Medicine ha colaborado con la Library of Congress en el desarrollo de BIBFRAME desde sus inicios. Sin embargo, el pasado 21 de noviembre de 2014 la propia Nancy Fallgren anunció a través de la [lista de discusión dedicada a BIBFRAME](#) que la National Library of Medicine daba un paso atrás en su proyecto con BIBFRAME para comenzar a experimentar con un enfoque más modular. En este sentido, se palpaba la expectación, por la gran asistencia de público al evento del MARC Format Transition Interest Group.

Nancy Fallgren comenzó su presentación hablando de los objetivos que guiaron el desarrollo de BIBFRAME como sustituto de MARC y comentó que tanto el Grupo de Trabajo de la Library of Congress sobre el Futuro del Control Bibliográfico, como el Comité Coordinador del Test americano de RDA como el informe BIBFRAME han coincidido en que debía ser un modelo basado en la web, agnóstico respecto a las reglas de catalogación empleadas y, por tanto, flexible y extensible, como útil más allá de la comunidad catalogadora a nivel bibliográfico, ampliamente comprensible y utilizable.

Pero tal y como ha demostrado con su visualización Jenn Riley, <http://www.dlib.indiana.edu/~jenrile/metadatamap/>, la cantidad de estándares de metadatos existentes es apabullante, por lo que es preciso que, para alcanzar los objetivos que han guiado desde sus inicios el modelo BIBFRAME, se haga desde una perspectiva modular, no basado en un estándar existente, como pueda ser MARC al que pretende reemplazar, sino utilizando los estándares descriptivos que ya han sido desarrollados por expertos, conectando datos equivalentes entre diferentes estándares descriptivos y siendo flexible con los cambios en dichos estándares descriptivos.

Un modelo BIBFRAME con este carácter modular, que es el que persigue en estos momentos la National Library of Medicine, debe ser capaz de integrar diferentes modelos como Dublin Core, Schemas locales, Press00, RDA, VRA, MODS, EAD, etc.

¿Qué se ha hecho hasta ahora? Utilizar el vocabulario BF esbozado por la empresa Zepheira y añadir el perfil de aplicación de metadatos a registros estándar BIBCO y CONSER, siguiendo las políticas PCC para RDA, con un mapeo al vocabulario BF, otro RDA/RDF y revisar los trabajos llevados a cabo.

Los próximos pasos a dar serán revisar y refinar el vocabulario con Zepheira, compartir BF Core v.1 y observaciones a él con la comunidad BIBFRAME, con PCC y con el Joint Steering Committee para el Desarrollo de RDA, testear el mapeo/conversión con metadatos preexistentes, testar la creación de nuevos datos bibliográficos y colaborar con Jackie Shieh de la George Washington University.



Se realizará una prueba de concepto con la figura del investigador Marshall W. Nirenberg, Premio Nobel de Fisiología en 1968 y todos los recursos relacionados con él en la National Library of Medicine (perfiles en Science, fotos, textos digitalizados, colección de manuscritos, citas de artículos). Se trabajará en relación al mapeo y conversión de datos ya existentes en NLM en MARC, Medline XML, perfiles XML, Dublin Core, su conversión a RDF y el mapeo de esquemas descriptivos al vocabulario BF. Para finalizar se compartirán todos los metadatos existentes con OCLC.

Se está trabajando con Zepheira para crear un perfil RDA para su herramienta Scribe y se harán pruebas paralelas de testeo con la Library of Congress en la primavera de 2015.

## **6.2. BIBFLOW: An IMLS Project / Xiaoli Li, Co-head of Content Support Services Department, University of California Davis**

BIBFLOW es un Proyecto de investigación de las bibliotecas de la Universidad de California Davis de dos años de duración (mayo 2014 – abril 2016), con la colaboración de la empresa Zepheira. Oficialmente se conoce como “Reinventando la catalogación: Modelos para las futuras operaciones bibliotecarias” y de lo que trata es de dirimir el impacto de la adopción de BIBFRAME sobre los flujos de trabajo de los Departamentos de procesos técnicos en las bibliotecas académicas. El nombre BIBFLOW deriva de BIBframe + workFLOW.

Uno de los resultados previstos es una hoja de ruta que ayude a las bibliotecas a pasar de un modelo de trabajo basado en la catalogación a un sistema nativo de datos enlazados/BIBFRAME. Pero, ¿cómo desarrollar esa hoja de ruta? Datos enlazados representa un salto evolutivo muy importante para las bibliotecas, no una simple migración a un nuevo sistema. El trabajo bibliotecario implica un compromiso diario con una gran colección de sistemas de software, instituciones y empresas suministradoras. En este sentido, cambiar a un ecosistema basado en datos enlazados no significa conocer simplemente su impacto en los datos bibliotecarios, sino en toda esa compleja red de flujos de trabajo.

Hasta ahora se han centrado en ocho flujos de trabajo y en cada uno de ellos han trabajado en casos de estudio como “catalogación de monografías originales utilizando el cliente de OCLC Connexion”. En este caso, se han definido los actores (bibliotecarios con gran experiencia, un especialista en catalogación con experiencia media, un catalogador de copia con un nivel bajo de experiencia y sistemas y herramientas de apoyo), se ha descrito la historia de usuario, así como la métrica de éxito y los pasos a dar. Invitó al público asistente a analizar y realizar comentarios sobre estos flujos de trabajo en el blog del proyecto: <http://www.lib.ucdavis.edu/bibflow/workflow-analysis>, fundamentalmente, en comparación con los de sus propias instituciones.

Para hacer posible la transición a ese nuevo modelo de datos enlazados/BIBFRAME, las bibliotecas necesitan sistemas orientados a datos enlazados. Han evaluado las posibilidades de Kuali-OLE, pero la versión analizada (OLE.1.0) no soporta las operaciones en relación a datos enlazados/BIBFRAME. Sin embargo, consideran que la base de datos Bibliográfica de OLE (DocumentStore), es robusta, extensible y capaz de trabajar con datos enlazados, lo que desde el punto de vista de desarrollo es muy ventajoso.



A partir de los flujos de trabajo identificados y de la evaluación de Kuali-OLE van a desarrollar un modelo de implementación preliminar de datos enlazados/BIBFRAME, que podrá funcionar y comunicarse de forma sincronizada sin importar si los datos son originalmente MARC o datos enlazados/BIBFRAME.

Los próximos pasos serán mejorar BIBFRAME Scribe, la herramienta desarrollada por Zepheira para catalogar directamente en el entorno BIBFRAME, para añadir servicios externos y desarrollar perfiles BIBFRAME; programar Kuali-OLE de forma que los catalogadores puedan utilizar Scribe para crear descripciones BIBFRAME de diversos materiales y almacenar los datos en RDF; desarrollar herramientas y servicios de transformación de datos; así como identificar y conectar un OPAC de código abierto.

La existencia de diferentes sistemas para distintos materiales como repositorios, catálogos bibliográficos, catálogos de materiales de archivo, frustra a los usuarios. El sueño de Xiaoli es poder contar con una única interfaz para todos los tipos de materiales, independientemente del estándar utilizado para su descripción y del origen de las descripciones de los materiales, y vinculada a todo el contenido de la web.

Proporcionó una dirección de correo electrónico, [bibflow@lib.ucdavis.edu](mailto:bibflow@lib.ucdavis.edu), y la página del blog del proyecto, <http://lib.ucdavis.edu/bibflow>, para enviar comentarios.



**Domingo, 1 de Febrero de 2014**

## **7. LINKED LIBRARY DATA INTEREST GROUP**

### **7.1. Linked Data for Libraries Project / Nancy Lorimer, Interim Head of Metadata Department at Stanford University Libraries**

[LD4L \(Linked Data for Libraries\)](#) es un proyecto a dos años financiado por la Andrew W. Mellon Foundation en el que participan las Bibliotecas universitarias de Cornell, Harvard y Stanford con el objetivo de crear un modelo semántico de almacenamiento de información en relación a recursos de investigación. Nancy Lorimer en su presentación ofreció una visión de los trabajos desarrollados en el primer año del proyecto como el desarrollo de la ontología LD4L, la migración de datos desde MARC a BIBFRAME y la integración de modelos de metadatos dispares, así como de datos no provenientes del mundo bibliotecario. Realizó una presentación posteriormente el mismo día en el [CAMMS Forum](#) en la que abundó en el tema.

### **7.2. Opening up Science with VIVO / Kristi Holmes, Director of Galter Health Sciences Library at Northwestern University**

Kristi Holmes habló del papel de VIVO en la apertura de datos científicos. Comenzó comentando que las bibliotecas están muy bien posicionadas para ocupar el papel de facilitar la representación abierta de información en relación a la investigación para estimular la colaboración, el descubrimiento y las actividades de evaluación. Por otra parte, sistemas de información sobre investigación como VIVO juegan un papel principal al permitir este trabajo tanto a nivel nacional como internacional.

La Web semántica no trata de poner datos en la web, sino de establecer enlaces de forma que una persona o una máquina puedan explorar esa red de datos. Con datos enlazados puedes encontrar otros datos relacionados con aquellos.

¿Por qué es esto importante a nivel institucional? A nivel de facultad, la investigación es cada vez más interdisciplinar por lo que es importante encontrar colaboradores, ver cómo se comporta la competencia, estar al tanto de la investigación llevada a cabo por otras instituciones tanto a nivel nacional como internacional, encontrar colegas y comunidades con los mismos intereses. A nivel de instalaciones y personal, los directores de bibliotecas necesitan alinear sus planes estratégicos con las necesidades en continua evolución de sus usuarios, así como identificar áreas de crecimiento de la investigación a través del incremento de las publicaciones, por ejemplo. A nivel de administradores, los centros de investigación son cada vez más amplios y diferentes, y es necesario mejorar la visibilidad de la institución.

La información sobre los investigadores se ha optimizado a través de una infraestructura de estándares y tecnologías que permite relaciones entre universidades, agencias y sociedades tanto dentro de Estados Unidos como fuera.



Respecto a los aspectos relacionados con la red científica, la información sobre los investigadores se ha optimizado utilizando una infraestructura de normas y tecnologías basadas en la Web que permite una descripción distribuida de los datos, legible por máquina y una aplicación web que genera enlaces entre universidades, agencias y sociedades. Esta aplicación web semántica abierta es VIVO que, además, es un modelo de información y una comunidad abierta, además de una ontología que permite el descubrimiento de la investigación que se está llevando a cabo a través de las diferentes disciplinas en una institución. Recopila datos de fuentes solventes y ofrece perfiles detallados de investigadores y facultades.

VIVO es una aplicación para web semántica de código abierto que proporciona una visión integrada del trabajo académico de una organización que puede facilitar el descubrimiento en muchos niveles tanto en el interior de una organización como más allá. Lo que permite VIVO es establecer una red nacional de científicos. Lo que buscan es construir una amplia red de datos más grande que cualquier otro esfuerzo anterior, siendo a la vez agregadores, creadores y consumidores. Un perfil en VIVO permite descubrir colegas o recursos por área de trabajo, autoría, colaboraciones, mostrar habilidades, logros académicos, visualizar áreas de especialización temática y redes de colaboración, entre otras cosas.

En la actualidad, 42 Universidades ya están en VIVO: varios campus de una misma universidad, conexiones regionales, consorcios, trece universidades holandesas y universidades alemanas. La Universidad de Brown utiliza por ejemplo la plataforma VIVO para incluir los CVs de su personal investigador. El Weill Cornell Medical College ofrece datos de varias fuentes y mezcla sus propios datos con datos que proceden de PubMed. La Duke University también incluye datos de fuentes muy variadas.

Se ha creado también una herramienta basada en Drupal para visualizar datos semánticos. El eje de trabajo es unir datos más herramientas más la participación de los propios investigadores.

La búsqueda es el tema principal en VIVO. La tecnología en VIVO es asombrosa, pero lo más alucinante es la participación de la gente. Existen en la actualidad un montón de proyectos VIVO alrededor del mundo. Las bibliotecas en VIVO están avanzando con mucho éxito hacia la web semántica ejerciendo una variedad de roles relacionados con las plataformas de investigación en red.

### **7.3. BIBFRAME: The Way Forward for Library Visibility on the Web / Victoria Mueller, Senior Information Architect and System Librarian, Zepheira**

Victoria Mueller, de la empresa Zepheira, que trabaja con la Library of Congress en el desarrollo de BIBFRAME, dividió su charla en tres partes fundamentales: una introducción al contexto de datos enlazados en la web, una explicación de lo que es BIBFRAME y sus objetivos y una tercera parte sobre cómo actuar para que los datos de las bibliotecas sean visibles en la web.

Comenzó hablando del tema que considera fundamental: la visibilidad en la web de las bibliotecas. No se trata ya de crear un nuevo producto, sino de construir una comunidad, un proceso social diseñado para integrar a diferentes partes implicadas que puedan hacer avanzar los estándares. Es necesario entrelazar lo mejor de las personas con lo mejor de los datos. Se habla



mucho de los “científicos de datos”, está muy de moda, pero, ¿de qué estamos hablando? Los bibliotecarios están adelantados a su tiempo y las nuevas normas pueden llegar de una forma extremadamente rápida.

En la actualidad estamos en un nuevo paradigma en el que los datos de las bibliotecas deben poder ser encontrados utilizando motores de búsqueda semánticos, los datos de las bibliotecas deben poder ser encontrados fuera de las bibliotecas. La web se está convirtiendo cada vez más en la plataforma de integración y gestión de datos más generalizada. En este sentido, si conseguimos exponer los datos de las bibliotecas de un modo que la web los comprenda estaremos beneficiándonos de las ventajas de esta plataforma. Para ello es necesario utilizar RDF ya que es el modelo común para crear datos en la web. RDF se parece a las piezas de LEGO que deben ser utilizadas para construir algo diferente.

Z39 y MARC21 ya no son las herramientas más adecuadas en la actualidad. Esa fue la razón por la que la Library of Congress contactó con la empresa Zepheira, para construir un nuevo entorno apropiado y adaptado a la web, en un entorno de datos enlazados. El objetivo de BIBFRAME es volver a concebir y a largo plazo, implementar un nuevo entorno bibliográfico para las bibliotecas que convierta esta red en central y convierta la interconexión en el lugar común.

El foco inicial se centró en los datos bibliográficos, pero pronto se vio los beneficios de atender también a datos no bibliográficos como recursos visuales, musicales. Las comunidades pueden definir perfiles BIBFRAME de acuerdo a sus propias necesidades. Cualquier cosa puede ser modelada en BIBFRAME. Lo bello es lo pequeño. En este sentido, el modelo BIBFRAME es flexible y diseñado para acomodarse a las necesidades de la comunidad bibliotecaria. En él se pueden combinar diferentes perfiles como series, monografías modernas y antiguas, música, VRA, ISAD(G).

Finalmente, habló de cómo entrar en acción. En la actualidad no se puede localizar en Google la información de los recursos que tu biblioteca dispone de una determinada obra porque la información está en tu catálogo, pero no es reconocible por la web. En este sentido, un enlace vale más que mil palabras. Es preciso incrementar la visibilidad de todas las bibliotecas en la web. Zepheira ha desarrollado la iniciativa Libhub, a la que invitan a participar a todas las bibliotecas que quieran tanto por la vía de convertirse en un pionero en la adopción de BIBFRAME, como asistiendo a formación en torno a BIBFRAME y Datos Enlazados. Asimismo, en el plano individual animan al compromiso.

Las bibliotecas, explicó Mueller, han gastado una gran cantidad de tiempo, energía y esfuerzo en describir sus recursos. De lo que se trata con BIBFRAME es de exponer ese esfuerzo de una forma extremadamente fácil.

¿Qué es lo que puede hacer una biblioteca hoy en día? Repensar su papel para hacer que las instituciones de la memoria se conviertan en los motores de credibilidad que necesita la web, aprender sobre estándares de datos enlazados como schema.org y BIBFRAME, así como participar en Libhub para ayudar a entrelazar lo mejor de la web con lo mejor de las bibliotecas.



### 7.3.1. Linked Data for Libraries (LD4L) / Stanford University Libraries

El Proyecto LD4L (Linked Data for Libraries) de las universidades de Cornell, Harvard y Stanford cuenta con una beca de la Andrew W. Mellon Foundation para su desarrollo durante dos años a partir de enero de 2014. Los objetivos del proyecto son construir una ontología y establecer fuentes de datos enlazados que proporcionen relaciones, metadatos, etc.

La estructura técnica se basa en Fedora, Solr, Blacklight, Hydra-Head. Se trabaja con diferentes estándares de datos bibliográficos (MARC, MODS, EAD), de datos personales (VIVO/CAP, ORCID, ISNI, VIAF). A ellos se unen datos de uso en relación a la circulación, cita, trabajo con contenidos como exposiciones, guías de investigación, temarios, etiquetas, etc.

Respecto a la ontología, la Library of Congress ha desarrollado la ontología BIBFRAME que pretende convertirse en el sustituto final de MARC. Los componentes probables de la ontología LD4L serán BIBFRAME junto a otros tipos bibliográficos adicionales como [FaBiO](#) (FRBR-aligned Bibliographic Ontology), la [Music Ontology](#), Schema.

Los desafíos en relación a la ontología se centran fundamentalmente en el carácter inestable en la actualidad de BIBFRAME, ya que es un proceso todavía no finalizado y en continuo movimiento. En este sentido se sigue trabajando en la interpretación de bf:Person, el modelado de bf:Event); la reconciliación de entidades y sus procesos, integrando identificadores ya existentes.

Casos de uso:

1) Annotations (datos bibliográficos + curation): Se trata de construir una colección virtual y etiquetar recursos académicos y de investigación para apoyar la reutilización. (Standáres: Digital Publishing Interest Group; EPUB Standards agency)

2) Autoridades: Descubrir obras a través de información sobre personas y sus relaciones, descubrir obras a través de acontecimientos y sus relaciones.

3) Linked Open Data.

Entre enero y junio de 2015 se harán públicos ejemplos de Harvard y Stanford. En febrero de 2015 habrá un taller de dos días para 25 personas pertenecientes a diez o doce archivos, bibliotecas e instituciones de la memoria.

El resultado del Proyecto será una ontología LD4L extensible y abierta, compatible con VIVO y BIBFRAME. Las bibliotecas universitarias de Cornell, Harvard y Stanford se unieron en este Proyecto debido a su interés común en los temas de Datos enlazados.

Al finalizar la presentación comenzó el turno de preguntas. Una persona lanzó la cuestión de donde estaba OCLC en todo este tema, a lo que se respondió que acababa de publicar un documento sobre el tema. Otra persona del Smithsonian habló del Proyecto piloto que estaban llevando a cabo con VIVO y preguntó si sería posible compartir experiencias, éxitos, fracasos, etc.



## 8. LC BIBFRAME UPDATE FORUM

El Foro de actualización sobre BIBFRAME de la Library of Congress se presentaba como una oportunidad para debatir sobre los progresos y preocupaciones en torno al Proyecto BIBFRAME, considerado como la iniciativa liderada por la Library of Congress para la creación de una nueva estructura bibliográfica. Participaron en el Foro:

- Por parte de la Library of Congress: Sally H. McCallum, Jefa de la Network Development/MARC Standards Office; Beacher Wiggins, Director del Departamento de Adquisiciones y Acceso Bibliográfico; Paul Frank, Coordinador de los programas NACO y SACO; Nate Trail, Coordinador de Proyectos Digitales.
- Por la Universidad de Stanford: Philip Schreur, Director del Departamento de Metadatos de las Bibliotecas de la Universidad de Stanford.
- Por la OCLC: Ted Fons, Director Ejecutivo de Servicios de Datos y Gestión de Calidad de WorldCat.
- Por la empresa Zepheira: Eric Miller, presidente de Zepheira.

### 8.1. Introduction: BIBFRAME Developments/ Sally H. McCallum

Sally McCallum habló del desarrollo del grupo que ha realizado el banco de pruebas durante el último año y animó a todo el mundo a participar en el experimento. Comentó que los holdings y las notas se van a incluir en el modelo próximamente y que un experto en RDF está trabajando en la actualidad en la revisión y ajuste del modelo para conseguir una mayor consistencia.

### 8.2. Library of Congress BIBFRAME Pilot Plans/ Beacher Wiggins

Beacher Wiggins comentó acerca del Proyecto piloto desarrollado en la Library of Congress para catalogar directamente en BIBFRAME, en el que han participado entre 25 y 30 trabajadores que fueron formados previamente en temas como RDA, ontologías y Web semántica.

El foco de atención técnica se centra ahora en la revisión de vocabularios, así como el dar cabida a todo tipo de lenguas y escrituras, por una parte, y con especial atención sobre los materiales especiales. En este sentido están trabajando en el desarrollo de materiales para la formación de catalogadores.

Comentó como este Proyecto piloto les había servido y así lo habían utilizado para saber cómo mirar hacia el futuro.

### 8.3. Pilot Component: BIBFRAME Editor/ Paul Frank

Paul Frank comenzó hablando de la gran cantidad de problemas que habían surgido en torno a la formación. Comentó cómo es posible en la actualidad trabajar con el editor BIBFRAME directamente y crear “algo”, que no podemos considerar un registro bibliográfico ya, sino algo completamente diferente.

El conocimiento previo en torno a la catalogación es lo que va a permitir avanzar. Mencionó que todo aquel que experimente con el Editor BIBFRAME no necesita conocer toda la infraestructura que hay detrás, sino que lo único que precisa es comenzar a usar el editor. Señaló que el uso del Editor BIBFRAME permite adquirir un mejor conocimiento de vocabularios, ontologías y todo lo que hay por



detrás. Además, su uso puede calmar el malestar o el desasosiego que se percibe en torno a estos temas tan nuevos.

#### **8.4. Pilot component: Profiles and Search/Display / Nate Trail**

Se está trabajando en toda una serie de herramientas como: el Editor BIBFRAME, el Editor de Perfiles, una herramienta de búsqueda y visualización, SRU en MARCXML.

En relación al Editor BIBFRAME, los perfiles están documentados en la web de la Library of Congress: <http://www.loc.gov/bibframe/docs/bibframe-profiles.html>. Ha sido la empresa Smart Software Solutions la que ha desarrollado el Editor de Perfiles.

La herramienta de búsqueda SRU de BIBFRAME ha sido desarrollada por IndexData: [www.indexdata.com](http://www.indexdata.com).

La herramienta BIBFRAME Search/Display está basada en Fedora 4 con Elastic Search y cuenta con dos componentes: bibframe datastore, bibframe catalog.

También se trabajado en una herramienta de comparación y transformación de MARC a BIBFRAME.

Un cambio significativo que se va a producir próximamente es la actualización para permitir trabajar con registros de holdings.

#### **8.5. Le Boeuf sur le Toit: A Surrealistic Exploration of BIBFRAME for Production / Philip Schreur**

Lo primero que comenté es que había cambiado el título de la presentación en el último momento para ofrecer lo que denominé una exploración surrealista a BIBFRAME, aludiendo en el título al conocido café-cabaret parisino en el que se reunían los artistas de vanguardia del período de entreguerras. Hablé de las instituciones que están experimentando con BIBFRAME para explorar las posibilidades que ofrecen los Datos Enlazados: las bibliotecas de la Universidad de Stanford, la Universidad George Washington, la Universidad de Cornell, la Universidad de Columbia, entre otras.

Los objetivos que comparten son compartir espacio en la nube, compartir archivos de investigación, material relativo a autoridades y como se debería trabajar con ellas en este nuevo contexto.

Hablé de un Nuevo Proyecto de la UC Davis (Universidad de California Davis) conocido como BIBFLOW, centrado en una única institución y en las herramientas desarrolladas por la empresa Zepheira, con atención especial en la integración de dichas herramientas con OLE.

Comenté la vía seguida por las bibliotecas de la Universidad de Stanford en este proyecto que se centra en temas como la catalogación por copia a partir de los registros proporcionados por empresas suministradoras, la conversión de MARC a BIBFRAME, la catalogación original directa en BIBFRAME. El último proyecto que están considerando en la actualidad es el trabajo sobre un dominio particular, en este caso, autoridades relativas al ámbito musical, para lo que están trabajando con la Music Cataloging Association. Esto les permite volver a pensar con más detenimiento acerca de la búsqueda y descubrimiento de estos materiales.



## 8.6. OCLC and LC Collaboration / Ted Fons, OCLC

Ted Fons se refirió a la estrategia seguida que se centra en el modelado de cosas de interés para la web (centrándose especialmente en los elementos de datos para el descubrimiento web), en facilitar que este tipo de cosas esté disponible en estructuras que la web reconozca y le sean familiares (comenzando con Schema.org, la gestión de entidades y no de registros...) y mejorar los flujos de trabajo en bibliotecas (mejorar el descubrimiento, reinventar la catalogación vendría después).

Se refirió al nuevo libro blanco definido como “puntos en común”, redactado por Carol Jean Godby, de la OCLC, y Ray Denenberg, de la Library of Congress y titulado [\*Common Ground: Exploring Compatibilities Between the Linked Data Model of the Library of Congress and OCLC\*](#). Este resumen ejecutivo compara y contrasta las iniciativas de datos enlazados compatibles de ambas instituciones. Durante el año 2015 se publicará un análisis técnico más detallado. La principal conclusión del informe se centra en señalar que el alineamiento de ambos modelos es acertado y quizás es incluso más defendible que en 2013 porque los principales conceptos BIBFRAME son ahora más consistentes con los conceptos correspondientes definidos en el modelo de OCLC basado en Schema.

La OCLC está trabajando en estos momentos en el grafo del conocimiento bibliotecario (the Library Knowledge Graph). Si trabajamos correctamente, muchas cosas mejorarán.

A nivel de la entidad Obra, ya hay más de 197 millones de descripciones de obras y URIs correspondientes a dichas obras, con Schema.org y Bibliograph.net, recolectadas desde WorldCat y enriquecidas con datos en RDF de otros núcleos con el objetivo de mejorar el descubrimiento.

Progress: Work Entity (plus than 197 million work descriptions and URIs; with Schema.org + Bibliograph.net; harvested from WorldCat data and enriched from other hubs RDF data formats); improving discovery.

¿Se puede medir el impacto de estos trabajos? El tráfico comenzó a incrementarse en mayo de 2014 cuando empezamos. ¿Qué podemos hacer? Seguir catalogando. Añadir nuestros metadatos bibliográficos y holdings a WorldCat, todo lo que viene siendo tradicional en nuestro trabajo.

## 8.7. The Libhub Initiative: Accelerating / Eric Miller, Zepheira

Eric Miller, de Zepheira, comenzó comentando que han estado trabajando en la perspectiva del largo plazo con el Proyecto Bibflow, que pretende cambiar los flujos de trabajo en los procesos de catalogación. Durante el pasado año han compartido observaciones. Entre ellas, las más relevantes son: en primer lugar, una biblioteca es más que su colección, son sus activos, sus instalaciones, los servicios que ofrece a la comunidad, etc.; en segundo lugar, cuando hablamos de datos enlazados, lo más importante se refiere a los enlaces, los datos vienen después; en tercer lugar, la Web de datos ya está aquí; por último, las bibliotecas no están participando en todo su potencial.

Un aspecto clave de esta falta de participación se puede resumir en las palabras de Chuck Gibson, Director and CEO (Chief Executive Officer) de la Worthington Public Library (Worthington Libraries), cuando dice que “...cuando la comunidad busca en la web algo que tenemos, deberíamos



mostrarnos como una opción”. Pero no estamos hablando el mismo lenguaje que la web entiende. Hemos creado barreras y la gente no puede encontrar nuestros materiales en la web.

Durante el último año señala que han creado algunas herramientas como el visualizador de recursos BIBFRAME (BIBFRAME Resource Viewer), que ha sido una de las herramientas más provechosas ya que nos ofrece la posibilidad de comprender nuestros propios datos. Se ha incrementado el número de recursos y, además, estos recursos están altamente conectados. Se ha incrementado el número de autoridades. Existe un vocabulario para expresar dichos recursos, tipos y relaciones por la forma en que están conectados, en un lenguaje que la web entiende. **Un enlace vale más que mil palabras.**

Existe un continuum desde RDA, MARC, VRA, EAD y BIBFRAME o Schema.org. Además, **el panorama está en continua evolución. El Schema.org de hoy va a ser completamente diferente del Schema.org de mañana.**

La noción de visibilidad enlaza con comenzar a jugar a llegar a la web, ser capaces de hablar el lenguaje de la web, el lenguaje que la web entiende, de manera que entre los resultados de una búsqueda pueda aparecer junto a los resultados de Amazon los de nuestra propia biblioteca, por ejemplo.

Durante el verano de 2015 Zepheira introducirá su propio juego de datos optimizado para el enlazado a través de recursos bibliotecarios en Libhub.

Entrelazar lo mejor de las Bibliotecas con lo mejor de la Web, no trata sobre jugar con un sistema ni con actualizar a un servicio, sino que trata de hablar el lenguaje que la Web entiende.

En el **turno abierto de preguntas**, se señaló desde el público la necesidad de que exista un serio compromiso de toda la comunidad bibliotecaria, incluidos las empresas que comercializan software para bibliotecas. También se habló de la extrema importancia de crear perfiles para trabajar con diferentes tipos de materias. Philip Schreur, de la Universidad de Standford, señaló que sería necesario crear estos perfiles para ellos.

Por su parte, la recomendación de Ted Fons de OCLC fue “No sientan pánico”.

Eric Miller comentó que el cambio siempre asusta, por lo que es muy importante la formación, cómo podemos empezar y ser más eficientes trabajando de esta manera.

Por último, alguien del público, dirigió una pregunta a las personas de la Library of Congress que participaban en el acto: ¿Es RDA la mejor opción? Beacher Wiggins, director de Adquisiciones y Acceso Bibliográfico de la Library of Congress, contestó: No, no es la única.



## 9. CaMMS FORUM: BIBFRAME: NOTES FROM THE FIELD

El foro de la Sección de Catalogación y Gestión de Metadatos de la ALA se centró en los avances de BIBFRAME, lo que sabemos y lo que no sabemos y las experiencias de dos bibliotecas que están ya trabajando con el modelo. El hashtag utilizado para tuitear el evento fue #alctscammsfm15mw.

### 9.1. BIBFRAME Development / Angela Kroeger, Archives and Special Collections Associate, University of Nebraska at Omaha

Angela Kroeger ([akroeger@unomaha.edu](mailto:akroeger@unomaha.edu)) se encargó de dar una visión de conjunto sobre el estado actual del desarrollo de BIBFRAME, las herramientas que están ya disponibles y las que están en desarrollo, las cuatro clases principales dentro del modelo, para terminar con una pizca de filosofía.

Lo primero que señaló es que BIBFRAME es una obra en progreso. Es un modelo de datos enlazados expresados en RDF para la descripción de recursos, que no está basado en RDA o en FRBR, pero que es compatible con ambos. Rompe el esquema del registro bibliográfico tal y como lo conocemos, ya que los recursos se describen como un conjunto de piezas de metadatos enlazados. Las relaciones entre los diferentes recursos es lo que sirve para reunirlos.

BIBFRAME no pretende ser una mejora de MARC, sino la herramienta que permita reemplazar MARC. Tampoco es algo así como un OPAC de nueva generación o un nuevo sistema integrado de gestión de bibliotecas, sino una colección de herramientas y plantillas para uso de catalogadores y diseñadores de sistemas.

Las herramientas que están ahora mismo disponibles en el sitio web de la Library of Congress son: una herramienta de comparación de MARC a BIBFRAME, una herramienta de transformación de MARC a BIBFRAME, así como un Editor BIBFRAME. Las herramientas que se está desarrollando en estos momentos son la herramienta de búsqueda y descubrimiento, y un perfil para creación y edición.

La razón para el desarrollo de BIBFRAME se centra fundamentalmente en que MARC es difícil de interpretar por las máquinas. Además, RDA y FRBR están diseñadas para un nuevo entorno de datos enlazados. Es necesario que cuando una persona haga una búsqueda de una obra, un autor, una edición de una obra, etc., en Google encuentre junto a los resultados de Wikipedia, de Amazon o de la librería Barnes & Noble, los de nuestras propias bibliotecas.

El entorno de datos enlazados se basa en tripletas. Así, en RDF: sujeto – predicado – objeto. En FRBR: entidad – relación – atributo. En BIBFRAME: recurso – relación – propiedad. Por ejemplo, expresado en RDF: Edgar Allan Poe – es el creador – de “El cuervo”.

Como ejemplo de un modelo basado en datos enlazados, habló del proyecto de la Biblioteca Pública de Kansas “[Civil War on the Western Border](#)” (Premio Roy Rosenzweig 2014 a la Innovación en Historia Digital), proyecto que ofrece acceso libre a materiales seleccionados combinado con una capa temática de ensayos y artículos enciclopédicos sobre el período de 1854 a 1865 en la zona. El sitio está diseñado utilizando FRBR, RDA, BIBFRAME y DRUPAL.



Las cuatro clases principales en BIBFRAME son: Work, Instance, Authority y Annotation. La primera, Work, es la suma de la clases Obra y Expresión de FRBR y sus propiedades se aplican a todas las ediciones y todos los formatos. Es un concepto abstracto. La segunda, Instance, se corresponde con la Manifestación de FRBR. Ya es algo concreto y sus propiedades son específicas de una edición y un formato. La tercera, Authority, se refiere a los agentes, lugares, temas, períodos de tiempo y se puede enlazar a vocabularios controlados de muy diferentes fuentes, o a nivel local crear nuevas URIs para conceptos de autoridad que no estén representados en otras fuentes. Por último, la clase Annotation corresponde al Ítem en FRBR, entre otras cosas, ya que incluiría, además de la información sobre holdings, sumarios, críticas, etc.

Existen ya dos editores diferentes con los que se puede trastear y experimentar: el Editor BIBFRAME de la Library of Congress, muy granular y detallado, dirigido a catalogadores formados; y BIBFRAME SCRIBE de Zepheira, más reducido y simplificado, dirigido a personal no catalogador. Pronto se desarrollarán otro tipo de aproximaciones.

En relación a lo que separa a BIBFRAME de RDA y FRBR, comentó que “Obra” significa cosas diferentes para diferentes comunidades, FRBR no es un modelo óptimo para recursos audiovisuales ni para publicaciones periódicas. En BIBFRAME las relaciones son más flexibles y extensibles.

OCLC y W3C están trabajando en una extensión de Schema Bib hacia un nuevo vocabulario: en [BiblioGraph.net](http://BiblioGraph.net). BIBFRAME se usaría para descripción, intercambio de datos y Schema.org para facilitar el descubrimiento a través del uso de motores de búsqueda comunes.

También comentó el nuevo proyecto de la empresa Zepheira: Libhub Initiative, que servirá para publicar recursos BIBFRAME en la web abierta, para monitorear la visibilidad en motores de búsqueda. Está previsto que este nuevo proyecto se dé a conocer en el verano de 2015.

## **9.2. BIBFRAME and Beyond... Linked Data Projects in Stanford University Libraries / Nancy Lorimer, Interim Head of Metadata Department at Stanford University Libraries**

Nancy Lorimer situó el origen de los trabajos en torno a los temas de datos enlazados en la Universidad de Stanford en un taller celebrado en el año 2011. En el año 2014 las universidades de Cornell, Harvard y Stanford comenzaron un proyecto conjunto de datos enlazados conectando recursos de información académicos: el [proyecto LD4L](#).

Las fuentes de los datos en el proyecto son tanto datos bibliográficos (MARC, MODS, VRA, EAD); datos de personas (VIVO, ORCID, ISNI, VIAF); como datos de uso (Circulación, cita, exposiciones, guías de investigación, etiquetas, programas académicos).

Los resultados del proyecto LD4L han sido una ontología LD4L extensible de código abierto compatible con VIVO, BIBFRAME y otros esfuerzos de datos enlazados de otras bibliotecas, así como un mapeo transparente desde MARC a Solr via BIBFRAME.



Desde Zepheira se ha trabajado en la formación de 24 miembros de su personal, de los departamentos de Adquisiciones, Metadatos, Servicios Técnicos Musicales, Servicios Técnicos de la Bibliotecas especializadas en Asia y Próximo Oriente, Colecciones Especiales, Cartografía y la Biblioteca de Tecnología.

Entre los planes de la Universidad de Stanford no está mover todo su trabajo en producción a Bibframe, al menos no de momento. Se está trabajando en tres flujos de trabajo orientados a la producción: catalogación por copia, catalogación original y grabaciones sonoras musicales, ya que hay mucha gente interesada en la música en la Universidad de Stanford. Una decisión que está por tomar es la de si catalogar en MARC o en BIBFRAME. Finalmente se va a trabajar de forma colaborativa en la producción de datos enlazados en torno a [“Les Six”](#)

### 9.3. Thinking in BIBFRAME... Not so Far Fetched After All / Jackie Shieh, Coordinator Resource Description Group, George Washington University Libraries

Se anunciaba que Jackie Shieh se iba a centrar en hablar sobre el estado actual del trabajo de las bibliotecas de la Universidad George Washington, tras dos años explorando la transición desde MARC to Bibframe y en explicar lo que se había convertido en práctica diaria de lo que todavía quedaba por hacer y seguía siendo un desafío.

Comenzó hablando de lo que significa “Datos enlazados” y lo definió como un método de publicación de datos estructurados de forma que puedan estar interconectados y ser más útiles. Recordó, en este sentido, la frase de Tim Berners-Lee, “... la web semántica no trata solo de poner datos en la web, sino de hacer enlaces, de forma que una persona o una máquina pueda explorar la web de datos. Cuando tienes datos enlazados, puedes encontrar otros relacionados”. Es, por tanto, fundamental en el entorno bibliotecario enlazar entidades.

Los conceptos claves que deben ser estudiados en profundidad en este entorno son: identificadores, ‘cosas’ y representación en contexto. Puso ejemplos relativos a clases y propiedades que son básicos para el aprendizaje de BIBFRAME:

Clase: **en mayúsculas, siempre en singular** → bf:Work

Propiedad: **en minúsculas** → bf:creador

Propiedad del objeto: **tiene verbo** → bf:tieneExpresión

→ bf:tieneAnotación

Tipo de dato: **tiene nombre** → bf:isbn

Camel case (mayúsculas/minúsculas) **para más de una palabra** → bf:otroFormatoFísico

Subclase y subpropiedad

El fin y la promesa de BIBFRAME es reemplazar a MARC, o incluso a algún otro formato, maximizar la plataforma Web mediante la interconexión de las colecciones de las bibliotecas, aumentar la visibilidad de la biblioteca en la Web y permitir al usuario descubrir y encontrar datos relacionados.



En las bibliotecas de la Universidad George Washington se ha realizado la aproximación a BIBFRAME partiendo de tres perspectivas diferentes: los datos, el equipo y el proceso. En relación a los datos, las tareas que se han debido realizar son la selección de conjuntos de datos, el análisis de dichos datos, su mapeo, su transformación y el depurado de campos obsoletos (011, 029, 039, 262, 350, 440, 503, 870-873). Los conjuntos de datos que se seleccionaron fueron monografías, obras multiparte y otros formatos variantes. Respecto a la transformación de los datos, en relación a MARC es necesario tener datos más que con cadenas de datos; en relación a RDA es preciso ver cómo funciona en BIBFRAME y en relación a EAD, cómo puede funcionar en BIBFRAME.

Hay que tener muy claro que pasar tus datos desde MARC a BIBFRAME no va a transformar tus datos, van a seguir siendo los mismos datos, sean buenos o malos.

Se contó con un equipo de ocho personas: tres pertenecientes al Grupo de Descripción de Recursos, una al Centro Global de Recursos, un Archivero digital, y durante la primera mitad del año 2014, otras tres más de los Servicios de Catalogación y Metadatos de la Burns Law Library. Para la constitución de un equipo dedicado a estas tareas es preciso identificar talentos y fortalezas, promover y animar al aprendizaje independiente, alimentar las relaciones, guiar a aquellos liderazgos potenciales y desarrollar habilidades en la gestión de proyectos.

En BIBFRAME, se utilizan URIs (Uniform Resource Identifiers=Identificadores Uniformes de Recursos) para identificar de forma unívoca un recurso. Vía un servicio de enlaces como ORCID se puede conectar a otras autoridades generadas en otra parte, de forma que se puedan desambiguar a usuarios con nombres similares y/o conectar diferentes fuentes de información para crear una imagen más completa de un mismo autor.

Durante el año 2014 se han llevado a cabo análisis, selección y transformación de datos; se ha creado una Wiki y una lista de distribución para facilitar la comunicación; se ha implementado el Editor de la Library of Congress; se ha trabajado rastreando temas en Github en torno a la herramienta de transformación a BIBFRAME de la Library of Congress; se ha recibido formación adicional por parte de la empresa Zepheira y se ha comenzado a trabajar con BIBFRAME como parte importante de la vida diaria. Lo que se ha conseguido es tener una conciencia más global de la naturaleza de los datos, se ha mejorado respecto a la capacidad de discernir y diferenciar entre una entidad del tipo Work y otra del tipo Instance (Manifestación en FRBR), se ha trabajado también desde la perspectiva del estándar archivístico EAC-CPF (Encoded Archival Context: Corporate Bodies, Persons, and Families) y, lo más fundamental, se ha incorporado la colaboración como nueva norma de trabajo.

¿Qué queda todavía pendiente?: una aproximación de los perfiles más modular al modelo de datos, una mejor identificación de obras y autoridades, un modelo para los holdings, la indización y recuperación de datos creados en BIBFRAME, la incorporación o extensión a otros esquemas, proporcionar un orden a la descripción o emparejar campos de escrituras latinas y vernáculas.

Respecto al presente año 2015, van a trabajar en el testeo de vocabularios considerados principales, en una interfaz web para la entrada de datos, en la búsqueda y recuperación de datos MARC y BIBFRAME nativos para generar BF, así como en la búsqueda y recuperación vía protocolo SRU.



## 10.MARC ADVISORY COMMITTEE 2

[http://www.loc.gov/marc/mac/mw2015\\_age.html](http://www.loc.gov/marc/mac/mw2015_age.html)

### **Proposal No. 2015-03: Description Conversion Information in the MARC 21 Bibliographic Format**

Para esta propuesta no había habido un Discussion Paper previo. La representante de la National Library of Medicine preguntó la razón a lo que Sally McCallum, de la Library of Congress, respondió que no era una propuesta de la Library of Congress, aunque así se reflejaba en la propuesta, sino que la había formulado la Universidad de Stanford y se había considerado interesante incorporarla. En la Library of Congress, comentó, trabajan con registros que provienen de suministradores privados y cuando pasan el proceso de revisión, la información sobre su origen se pierde.

Desde la National Library of Medicine se expresó la preocupación en relación a la utilidad de este tipo de información a largo plazo. Se preguntó si consideraban tan solo interesante contar con este tipo de información a nivel local. Desde la National Library of Medicine se expresó que no consideraban útil ese tipo de información y su preocupación por quién va a mantenerla. Además, se expresó que fuera del entorno MARC no va a ser información útil.

Otros participantes como John Myers consideraron la propuesta como razonable.

El representante de la Biblioteca Nacional alemana comentó que ellos tienen una política sobre consignar información de los procesos de conversión en sus registros y consideró que para la web podía ser interesante contar con ese tipo de información acerca del origen de los registros para, por ejemplo, valorar e identificar el grado de confianza.

La principal preocupación de la British Library, comentó su representante, al respecto se centró en lo contradictorio de la definición al incluir los términos "origen" y "conversión", ya que plantearía si se refiere solo a registros originales o a registros corregidos también. En este sentido, propuso cambiar la frase para establecer: "...utilizado para proporcionar información sobre el origen del registro MARC que ha sido convertido automáticamente desde otra estructura de metadatos". También expresó la necesidad de que se redefine lo relativo al subcampo \$k y que se añada un subcampo específico para el formato de conversión. En este sentido, John Myers comentó que podría ser un subcampo \$b, referido a la versión de la transformación.

El representante de la Biblioteca Nacional Alemana sugirió la necesidad de dos nuevos subcampos adicionales en relación a la importación de datos desde MARC a PICA.

La propuesta se aprobó, con la abstención de la National Library of Medicine y los cambios propuestos en torno a la definición, leyenda y subcampos.

### **Proposal No. 2015-06: Defining New Subfield in Field 382 for Coding Number of Ensembles**

La propuesta de la Music Library Association sobre la definición de un nuevo subcampo en el campo 383 para la codificación del número de conjuntos fue aprobada tras un debate inicial en el que se plantearon fundamentalmente preguntas sobre su necesidad por parte de los miembros del Comité que no cuentan en sus colecciones con este tipo de materiales musicales.





**Lunes, 2 de Febrero de 2014**

## **11. HEADS OF CATALOGING INTEREST GROUP**

Las dos presentaciones del “Heads of cataloging interest Group” se centraron en el establecimiento de objetivos estratégicos en los proyectos que se presentaban y cómo conseguir el mayor impacto con los recursos limitados con los que se cuenta. Por un lado, los ensayos con BIBFRAME de las bibliotecas de la Universidad George Washington y, por otro, el proyecto BIBFLOW de la Universidad de California Davis, desarrollado en colaboración con la empresa Zepheira.

### **11.1. GW’S BIBFRAME Cookbook / Jackie Shieh, Georges Washington Libraries**

Jackie Shieh planteó en primer lugar que debemos ser sinceros y honestos con nosotros mismos a la hora de plantearnos proyectos de este tipo: ¿qué queremos? ¿ser considerados en nuestras instituciones como líderes capaces de llevar a la práctica proyectos? ¿creadores de metadatos? ¿catalogadores? ¿cuáles son nuestras metas y objetivos?

Los temas en los que se centró fueron: las inversiones necesarias en formación e infraestructuras, las habilidades que nuestro personal necesita desarrollar para un futuro de catálogos y flujos de trabajo en datos enlazados, cómo conseguir el mayor impacto en nuestros objetivos estratégicos contando con escasos recursos y qué conversaciones debemos tener con todas las personas implicadas en nuestra biblioteca y en nuestra institución.

En este sentido, los objetivos que se planteó el equipo de la George Washington University para 2014 en relación a BIBFRAME fueron: formación del equipo de experimentación; selección, transformación y evaluación de datos; realización de pruebas con el Editor y con el vocabulario BIBFRAME; creación de mecanismos de comunicación (lista de correo interna, Wiki para la documentación, recursos de aprendizaje, etc); diseño de materiales para la formación y creación de un servidor de pruebas. Existían unos condicionantes como eran el entorno de trabajo y la gestión del cambio asociada a él, la limitación temporal del proyecto y la urgencia en llevarlo a cabo, la flexibilidad presupuestaria, etc., que motivaban el establecimiento de unos objetivos específicos, medibles y realistas.

Para conseguir llevar a buen puerto un proyecto como éste es necesario vestirse con varios trajes: el traje de educador, el de intérprete, el de animador... Y es necesario pensar en cómo convertir tu frustración en algo positivo para tu biblioteca.

¿Cómo llevar a cabo la gestión del cambio? Es necesario seleccionar de forma cuidadosa un equipo pequeño de trabajadores, comenzar de una forma simple y a pequeña escala, permanecer centrado en lo específico, buscar la retroalimentación e innovar, fortalecer las habilidades tecnológicas y de trabajo, e intentar evitar los peros, considerando alternativas.

La formación del personal se llevó a cabo mediante ejercicios individuales y en grupo, con encuentros cada dos semanas y webinars.



En todas las fases del proyecto se han utilizado herramientas de código abierto que son muy útiles cuando dispones de pocos recursos, como las Google Tools, MediaWiki, Flow.io, Sharepoints, IdealScale, Draw.io). De especial valor han resultado los casos de uso y requerimientos que se recogen en la página web dedicada a BIBFRAME: <http://bibframe.org/documentation/bibframe-usecases/>.

Es preciso fomentar la colaboración para generar mayor conocimiento. En este sentido, el equipo de la Biblioteca Pública de Denver les propuso trabajar en torno a la figura de una residente histórica de Denver, Molly Brown, que sobrevivió al hundimiento del Titanic, perteneció al movimiento sufragista y fue la primera mujer que se presentó a unas elecciones en Estados Unidos: <http://www.mollybrown.org>. Molly Brown es una persona, pero también es un edificio, la casa en la que vivió, y tuvo una biografía tan rica que suponía un buen ejemplo para trabajar en un modelo de datos que permitiera extraer toda la información relacionada con su persona. El equipo de la George Washington University trabajó en un esquema sobre el que el equipo de Zepheira trabajó a posteriori.

Para este año 2015 se proponen realizar experimentos con vocabularios importantes, desarrollar una interfaz web para la introducción de datos, una herramienta de búsqueda y recuperación y un OPAC para datos BIBFRAME.

### **11.2. BIBFLOW: An IMLS Project / Xiaoli Li, Co-head of Content Support Services Department, University of California Davis**

Xiaoli Li realizó la misma presentación para la reunión del Heads of Cataloging Interest Group que había expuesto el sábado en la reunión del [MARC Format Transition Interest Group](#), por lo que remitimos a páginas anteriores.

Tras la presentación, desde el público se plantearon algunas cuestiones. En primer lugar, el problema de cómo plantear cuestiones como la colaboración cuando tienes que intentar convencer a tus jefes de la necesidad de tus proyectos y cómo actuar. Xiaoli Li comentó que si no tienes financiación a través de una gran beca, por ejemplo, debes seguir al tanto de lo que está ocurriendo en otras instituciones y sus proyectos en primer lugar. En segundo lugar, debes analizar tus propios datos y ver la manera de limpiarlos. También es bueno plantearse cuestiones acerca de cómo usar datos ajenos a los tuyos, por ejemplo identificadores ISNI. Evaluar este tipo de registros puede ser útil para cambiar tu forma de pensar. La cuestión fundamental es que debemos operar un cambio en nuestra mentalidad. Otra cuestión fundamental debe ser no reinventemos la rueda, Schema ya está ahí, utilicémoslo.

Por último, insistió en la importancia del conocimiento acumulado por el personal con años de servicio para las tareas de limpieza de datos. Debemos hacerles sentir especiales y animarles a que pongan sobre el papel todo lo que saben de su trabajo en la institución.



## 12. SAC INTEREST GROUP

La segunda reunión del Subject Analysis Committee de ALA comenzó con una ponencia de Gordon Dunsire, Chair del Joint Steering Committee for the Development of RDA, sobre lo que es una material en RDA.

### 12.1. What is an RDA Subject / Gordon Dunsire

Gordon Dunsire comenzó comentando que en el encuentro del Joint Steering Committee for the Development of RDA, celebrado en noviembre en Washington, el JSC estuvo de acuerdo en avanzar en la definición de las materias en RDA, pero será solo aplicable a la entidad Obra. Además se incorporará a RDA Toolkit un nuevo apéndice M que explicitará designadores de relación de materias. También se harán cambios en el juego de elementos de RDA, los designadores de relación y el Registro RDA. Este trabajo en relación a las materias está alineado con los modelos FRBR y de datos enlazados (Dublin Core Abstract Model).

El modelo FR ha evolucionado a lo largo del tiempo respecto a las materias y se ha invertido una gran cantidad de tiempo en la reflexión acerca de ellas, tanto a nivel de FRBR, con la definición de las entidades del grupo 3 (concepto, objeto, acontecimiento y lugar), como a nivel de FRAD (las entidades no se limitaban al contexto de las materias) y FRASAD, que ampliaron los datos de autoridad y de autoridad de materia de FRBR. El pasar de hablar de “registros” en FRBR a “datos” en FRAD/FRASAD revela los cambios que se han ido produciendo en el campo bibliográfico a lo largo de los últimos años. Los tres modelos en la actualidad están consolidados en un único modelo. El modelo FRASAD solo se refiere a datos de materias, mientras que FRBR hablaba de las entidades del grupo 3 (Concepto, Objeto, Acontecimiento y Lugar) como materias. En este sentido el Joint Steering Committee para el Desarrollo de RDA trabaja sobre el modelo FRASAD, y, aunque quizás mantenga las entidades del grupo 3 de FRBR, no se limitarán al contexto de las materias. Por ejemplo, el JSC está considerando tratar la entidad “Lugar” como una entidad en sí misma.

En FRASAD, la entidad Obra es la misma que en FRBR. La entidad “Thema” representa la materia como un concepto general, mientras que Nomen representa encabezamientos de materia, clasificaciones y notaciones que etiquetan un Thema. Una Obra puede ser la materia de otra Obra, por lo que puede ser también un tipo de Thema. Este mismo análisis se aplica a todas las entidades de los grupos 1 y 2 de FRBR, que, además, son también entidades RDA.

Obra → tiene como materia → Thema → tiene como nombre → Nomen

La relación “tiene como nombre” puede ser también definida como “tiene título” en el contexto de la Obra. Lo mismo se puede aplicar a la Expresión, la Manifestación y, en algunas circunstancias, al Ítem.

De hecho, un Thema puede ser cualquier cosa. En términos de entidades, Thema es lo mismo que “Cosa” (Thing). Thema puede ser incluido en un tratamiento más general en RDA, dejando solo la relación de materia entre la Obra y cualquier cosa que sea su materia (una Obra, una Expresión, Una Manifestación, un Ítem, una Persona, una Familia, una Entidad corporativa). Todas las entidades RDA están incluidas de forma implícita como objeto de la relación de materias. Las relaciones



recíprocas son explícitas para las entidades primarias Obra, Expresión, Manifestación e Ítem, e implícitas para las entidades secundarias Persona, Familia y Entidad Corporativa.

Las relaciones descriptivas de la Obra (Apéndice J.2.3 de RDA) podrían ser para la descripción de una Obra: análisis de una obra, comentario sobre una obra, crítica de una obra, evaluación de una obra, revisión de una obra.

Las relaciones descriptivas de la Expresión (Apéndice J.3.3 de RDA) utilizan el mismo análisis que en los designadores de relación para la Obra, pero esto contradice el modelo básico FRISAD en el que tan solo la Obra puede tener una materia, mientras que una Expresión no puede tener una materia.

Las relaciones descriptivas de la expresión serán redefinidas en RDA.

Se va a añadir un Nuevo apéndice M para designadores de relación entre entidades diferentes (excepto para Obra que tiene materia Obra):

Obra [tiene] materia [expresión] Expresión

[es] descripción de (expresión)

[es] análisis de (expresión)

[es] comentario sobre (expresión) y así sucesivamente...

Habría que preguntarse si necesitamos en realidad una relación al nivel del comentario de la expresión o si podría ser un género o una forma.

Los designadores de “materias” del Apéndice J se trasladarán al Apéndice M. En la actualidad se está discutiendo si son necesarios estos designadores más refinados o si pueden ser sustituidos por otros subtipos de obras basadas en la forma o el género, por ejemplo, una Obra que es un comentario, o una Obra que es un análisis.

ALA va a preparar una propuesta en relación al restablecimiento de este tipo de información.

Conclusiones:

- “Materia” (Subject) no es una entidad
- “Materia” es una **relación** between una Obra y otra cosa:
  - Obra “(tiene como) **materia**” Algo
    - Relación recíproca: Algo es **materia** (de)” una Obra

En la actualización de RDA Toolkit de abril de 2015 se incluirán actualizaciones en torno a las materias.