



## **ALA MIDWINTER MEETING 2016**

### **Informe de Asistencia al ALA Midwinter Meeting 2016 y a las reuniones del MARC Advisory Committee**

**Boston, 7 - 12 de Enero de 2016**

**Autor:** María Jesús Morillo Calero  
**Departamento:** Departamento de Proceso Técnico  
**Versión:** 01  
**Fecha:** 23/02/2016



## Índice

1.	INTRODUCCIÓN.....	6
2.	THING-ATHON.....	8
3.	GRUPO DE INTERÉS DE DIRECTORES DE SERVICIOS TÉCNICOS DE GRANDES BIBLIOTECAS.....	12
	3.1.El futuro del proceso físico y el almacenamiento de materiales en depósitos compartidos de materiales impresos .....	12
	3.2.Servicios técnicos externos.....	13
	3.3.¿Cómo compartimos datos en un nuevo entorno distribuido y colaborativo?.....	14
	3.4.Impactos sobre los servicios técnicos del desarrollo colaborativo de colecciones .....	14
	3.5.Estructura y moral del personal en los servicios técnicos .....	14
	3.6.Un modelo consistente y sostenible para los registros bibliográficos de empresas externas	15
	3.7.Actualización de BIBFRAME y beca para el Proyecto LD4L (Linked Data for Production)....	16
	3.8.Actualización sobre el proyecto piloto de catalogación cooperativa de CIC .....	17
4.	MARC ADVISORY COMMITTEE (1) .....	17
	4.1.DP 2016-DP07: Ampliación del uso del campo 257 para incluir Regiones autónomas en el Formato MARC 21 para registros bibliográficos .....	18
	4.2.DP 2016-DP09: Coding Named Events in the MARC 21 Authority and Bibliographic Formats	18
	4.3.DP 2016-DP13: Designación de una definición en el Formato MARC 21 de Autoridades.....	19
	4.4.DP 2016-DP14: Designación del Tipo de Entidad en el Formato MARC 21 de Autoridades ...	19
	4.5.DP 2016-DP12: Designación de información coincidente en el Formato MARC 21 de Autoridades	19
	4.6.DP 2016-DP08: Eliminación de la restricción sobre el uso de las fechas en el campo 046, subcampo \$k, del Formato MARC 21 para registros bibliográficos .....	20
	4.7.DP 2016-DP10: Definición del campo 347 (Características del Archivo digital) en el Formato MARC 21 de Holdings .....	20
5.	FORO DE ACTUALIZACIÓN DE RDA .....	21
	5.1.Progresos de RDA en relación a la gobernanza y la estrategia / Gordon Dunsire .....	21
	5.2.Actualización de RDA Toolkit / James Hennelly .....	22
6.	GRUPO DE INTERÉS SOBRE GESTIÓN DEL CATÁLOGO .....	23
	6.1.¿Qué estás haciendo, ¿qué?!? Metadatos, migración y fusión en Ex Libris / Kevin Clair, Erin Elzi	23
	6.2.GMD o No GMD: aproximación de una biblioteca a la conversión a RDA / Jim Kalwara .....	24
	6.3.Preparación de materiales impresos para su almacenamiento en un depósito externo: la experiencia de la Universidad de Notre Dame / Nastia Guimaraes .....	25
7.	GRUPO DE INTERÉS EN LA TRANSICIÓN DE LOS FORMATOS MARC.....	27
	7.1.Datos enlazados abiertos para el Descubrimiento: BIBFRAME + Schema.org en una vista experimental Bento / Jim Hahn .....	27
	7.2.De MARC a MODS XML y a los amigables datos enlazados RDF / Steven Anderson .....	27
	7.3.Datos enlazados para Archivos / Gloria Gonzalez (Zepheira), Dustin Stokes (Atlas Systems)	29



8.	GRUPO DE INTERÉS EN METADATOS.....	30
8.1.	Borrador de principios para la evaluación de estándares de metadatos / Jennifer A. Liss, Kevin Clair .....	30
8.2.	Hacer lo correcto en temas de derechos de propiedad intelectual / Emily Gore, DPLA (Digital Public Library of America) .....	31
9.	ACTUALIZACIÓN SOBRE BIBFRAME DE LA BIBLIOTECA DEL CONGRESO .....	33
9.1.	Actualización sobre el programa piloto BIBFRAME de la Library of Congress / Beacher Wiggings y Sally McCallum .....	33
9.2.	Una aproximación a BIBFRAME por parte de un distribuidor / Tiziana Possemato, Casalini Libri .....	34
9.3.	Visión de conjunto del proyecto LD4P (Linked Data for Production).....	35
9.3.1.	De-“Framing” (desmontando a) Derrida: Archivos y anotaciones / Jennifer Baxmeyer, Princeton University.....	35
9.3.2.	BIBFRAME y colecciones artísticas: Arte y Museos / Melanie Wacker y Amber Billey, Columbia University Libraries.....	36
9.3.3.	Proyectos LD4P en la Universidad de Cornell / Chew Chiat Naun .....	37
9.4.	Actualización de OCLC: BIBFRAME y el programa de datos Enlazados / John Chapmann .	37
9.5.	Actualización de Zepheira: Acelerando la Biblioteca Visible o “No esperen – Repetimos” / Eric Miller .....	38
10.	MARC ADVISORY COMMITTEE (2) .....	39
10.1.	Propuesta 2016-01: Codificación de posiciones del campo 007 para las reproducciones digitales de grabaciones sonoras en el Formato MARC 21 para registros bibliográficos .....	39
10.2.	Propuesta 2016-02: Defining Subfield \$r and Subfield \$t, and Redefining Subfield \$e in Field 382 of the MARC 21 Bibliographic and Authority Formats .....	39
10.3.	DP 2016-DP01: Definición de los subcampos \$3 y \$5 en el campo 382 (Medio de interpretación) del Formato MARC 21 para registros bibliográficos .....	40
10.4.	DP 2016-DP02: Clarificación de los valores en la posición 20 del campo 008 (Formato de Música) en el Formato MARC 21 para registros bibliográficos.....	40
10.5.	DP 2016-DP03: Números de distribuidores de materiales musicales e imágenes en movimiento en el Formato MARC 21 para registros bibliográficos.....	41
10.6.	DP 2016-DP04: Extensión del uso del subcampo \$0 (número de control del registro de autoridad o número normalizado) a campos de enlace.....	41
10.7.	DP 2016-DP05: Expansión de la definición del subcampo \$w (Número de control del registro bibliográfico) a los números normalizados.....	41
10.8.	DP 2016-DP06: Definición de los subcampos \$2 y \$0 en el campo 753 del Formato MARC21 para registros bibliográficos.....	42
10.9.	DP 2016-DP11: Puntuación en el Formato MARC21 de Autoridades.....	42
10.10.	DP 2016-DP15: Tipo de Medio y Tipo de Soporte en el Formato MARC21 Autoridades .....	42
10.11.	DP 2016-DP16: Extensión del nivel de codificación en el Formato MARC21 de Autoridades .....	43
11.	GRUPO DE INTERÉS DE DIRECTORES DE DEPARTAMENTOS DE CATALOGACIÓN ALCTS CaMMS .....	43
11.1.	Presentación.....	43
11.2.	Autoridades, Entidades, Objetos de la vida real y... ¿Gatos?: De la creación de autoridades a la gestión de identidades / Nancy Lorimer .....	43
11.3.	Contexto Archivístico Codificado: Entidades corporativas, Personas y Familias (EAC-CPF) / Katherine M. Wisser .....	46
12.	ACTUALIZACIÓN SOBRE LOS PROYECTOS DE INVESTIGACIÓN DE OCLC .....	48



12.1.	Investigación y membresía en OCLC / Eric Childress .....	48
12.2.	La biblioteca en la vida del usuario: involucrarse con la gente allí donde viven y aprenden / Lynn Silipigni .....	48
12.3.	Transformar las descripciones bibliográficas en conocimiento procesable / Jeff Mixer	49
12.4.	Trabajar con socios e identificadores de OCLC Research / Karen Smith-Yoshimura....	51
13.	RDA TECH FORUM.....	52



## 1. INTRODUCCIÓN

Como en anteriores ocasiones, este informe tan solo tiene como objetivo exponer el contenido de los diferentes eventos a los que he asistido en el Encuentro de Invierno de la Asociación Americana de Bibliotecas (ALA), celebrado en Boston entre el 8 y el 12 de enero, en torno a los temas de Proceso Técnico y presentar los debates y decisiones adoptadas en las reuniones del MARC Advisory Committee, en el que la Biblioteca Nacional de España está representada.

Varios temas han sido recurrentes en los eventos relacionados con los procesos técnicos a los que he asistido. La deslocalización de depósitos y su gestión cooperativa, cuestiones que enlazan con las actividades bibliotecarias más tradicionales y que ponen sobre el tapete asuntos como el acceso a los documentos y su preservación, fueron tratados en diversas ocasiones y se les concedió gran importancia, por ejemplo, en la reunión de los directores de Servicios Técnicos de las grandes bibliotecas académicas. Otros, más relacionados con el futuro como la transición a un modelo de datos enlazados, estándares de descripción y codificación, o la gestión de vocabularios e identificadores, continúan siendo los principales protagonistas de las discusiones.

En relación a la evolución de RDA como estándar de descripción, se están llevando a cabo grandes esfuerzos para conseguir extender su expansión a nivel internacional, como comentó Gordon Dunsire en el Foro de actualización sobre RDA. La nueva estructura de gobernanza reduce la participación de norteamericanos y europeos para incluir representantes de cada continente que hayan implementado RDA o vayan a hacerlo en los próximos dos años. Y, desde luego, a la vista de las discusiones en el encuentro del Committee on Cataloging: Description and Access CC:DA, de la American Library Association, sobre cómo reducir su representación a un único delegado, parece que el tema va en serio. Por otra parte, también se ha procedido a crear grupos de trabajo, tanto estables como temporales, sobre temas diversos. Es de agradecer, en este sentido, la creación del grupo permanente sobre traducciones y la definición de una nueva política a aplicar a partir de ahora por la que se impone la necesidad, a quien se encargue de traducir el estándar a una lengua determinada, de traducir el RDA Reference, las etiquetas y definiciones de las entidades, elementos y vocabularios RDA para que sean añadidos al RDA Registry

Los estándares de metadatos para bibliotecas (y también los de archivos y museos) han tenido gran solidez en el pasado reciente y han permanecido prácticamente inalterables durante años. No parece que esta situación se vaya a prolongar a partir de ahora, por lo que, en el futuro, se ha repetido como un mantra, tendrán valor aquellos estándares que cuenten con amplias comunidades que participen activamente en su desarrollo y utilización. La ausencia de extensas comunidades de usuarios puede suponer el abandono y la muerte de un estándar. Para su desarrollo y mantenimiento será fundamental la transparencia en los procesos y toma de decisiones.

Un tema recurrente en este Midwinter Meeting de ALA de Boston ha sido la cuestión de la evolución del control de autoridades hacia la gestión de identidades. Gordon Dunsire, Chair del RDA Steering Committee, Diane Hillmann, de la Dublin Core Metadata Initiative, Philip Schreur, de la Universidad de Stanford, Jackie Shieh, de la George Washington University, o Nancy Lorimer, de la Universidad de Stanford, fueron algunas de las personas que hablaron de este tema. A medida que nos movemos



hacia un entorno de datos enlazados, cuestiones como el tradicional control de autoridades se convierten en básicas y aumentan su importancia. Pero ya no basta con desambiguar homonimias, resulta imprescindible identificar unívocamente autores, obras, entidades, lugares, acontecimientos, etc., movernos desde el tradicional control de autoridades a la gestión de identidades.

En este sentido, en los últimos años han aparecido diferentes iniciativas para la definición de identificadores que establezcan de forma unívoca dicha identidad en el caso de autores, creadores o investigadores, como [ORCID](#), [ISNI](#) o [VIVO](#), para sumarse a los tradicionales ficheros de autoridades mantenidos por las bibliotecas como el propio de la [Library of Congress](#) o [VIAF](#), el Fichero de Autoridades Virtual Internacional. Pero, hay muchos interrogantes que deben ser despejados para que esta “gestión de identidades” llegue a un buen fin: ¿cómo se crean y asignan estos identificadores?, ¿quién es el responsable de su gestión y mantenimiento?, ¿cómo se va a garantizar su preservación y disponibilidad a largo plazo?, ¿en qué casos será mejor reutilizar identificadores globales y en qué casos identificadores locales?, ¿qué hacemos si llegado el caso esos identificadores globales desaparecen?, ¿qué va a pasar con el control de autoridades tradicional? El mantenimiento de los identificadores globales se convierte en uno de los temas fundamentales, ya que para su persistencia y estabilidad se requiere compromiso, para que sean verdaderamente globales se necesita que estén permanentemente disponibles y generen confianza, lo que está indisolublemente unido a la relevancia de su procedencia.

La transición desde el uso de cadenas de caracteres como elemento de identificación a la identidad real basada sobre URIs pone sobre la mesa muchas cuestiones completamente nuevas, que nos obligan a analizar con detenimiento el proceso de crear, gestionar y utilizar identidades mediante URIs. La ventaja de las URIs es que permiten un nivel de precisión imposible de reproducir utilizando cadenas de caracteres y que las máquinas realicen la mayor parte del trabajo. Pero no todo son ventajas, también existen inconvenientes. La utilización de estos identificadores creados por instituciones ajenas nos convierte a todos en dependientes de los gestores de esos vocabularios públicos que pueden, llegado el caso, decidir que tienen otras prioridades o, incluso, que ya no están interesados en su mantenimiento, lo que provoque su abandono y desaparición.

Otra de las cuestiones que se ponen encima de la mesa en dicha transición es los modos en que vamos a compartir los metadatos que creamos, cómo vamos a dar acceso a ellos y si estamos dispuestos a ofrecerlos de forma completamente abierta. ¿Existirán miles de nodos Sparkl y nos moveremos desde un entorno local a un entorno global y mundializado? Probablemente. Y puede que, como señaló Philip Schreur, de la Universidad de Stanford, solo sea necesario almacenar los datos sobre los ejemplares a nivel local y el resto de datos en algún sitio externo.

Pero, ¿vamos a ser capaces todos de movernos rápidamente a estos nuevos entornos? Seguramente no. Las grandes bibliotecas académicas y las grandes bibliotecas nacionales lo podrán hacer relativamente pronto, costará más al resto y, probablemente, habrá muchas bibliotecas más modestas que jamás puedan dar ese salto, se comentó en varios foros.

La pregunta del millón: ¿debemos esperar a que BIBFRAME se finalice y consolide como sustituto de MARC o debemos experimentar por nuestra cuenta en la exposición de nuestros datos en la web



como datos enlazados? La Chair del Grupo de Interés en la transición de los formatos MARC mostró su preferencia por la segunda opción, señaló que tenemos serios problemas de comunicación en la comunidad y que hay ejemplos suficientes de gente trabajando en estas áreas sin esperar a que la Biblioteca del Congreso diga: ¡Ya estamos aquí! ¡Ya hemos terminado BIBFRAME!

De momento, desde la Library of Congress comentaron sus avances y sus estrategias a corto plazo: la publicación de una propuesta de vocabulario 2.0 (modelo, principios, denominación, mapeo MARC), una reingeniería de la herramienta, el inicio de un segundo proyecto piloto con el nuevo vocabulario 2.0, continuar con los proyectos LD4P (Linked Data for Production), así como explorar otros medios. Entre sus últimas realizaciones está la publicación de un nuevo informe sobre metadatos técnicos para recursos audiovisuales: <http://www.loc.gov/bibframe/docs/pdf/bf-avtechstudy-01-04-2016.pdf>. Por otra parte, la experimentación no cesa en el marco del Proyecto LD4P (Linked Data for Production), con estudios como los de Princeton en torno a la biblioteca del filósofo francés Derrida o los de la Universidad de Cornell en relación a los vocabularios RBMS o la colección de LPs del DJ Afrika Bambaataa.

Lo que se espera de todos estos desarrollos y experimentación es la definición de nuevas herramientas de descubrimiento que permitan extraer conocimiento relevante a partir de todo tipo de datos e informaciones. Google ya está trabajando desde hace unos años en la creación de un gráfico del conocimiento, que ha evolucionado hacia lo que denominan la Cúpula del conocimiento, basándose en modelos probabilísticos y de aprendizaje automático. OCLC, a través de su proyecto piloto EntityJS, pretende adaptar un modelo similar para datos bibliográficos, añadiendo contenidos suministrados por los usuarios y fuentes externas que le permita la construcción de una nueva plataforma de descubrimiento. ¿Será capaz OCLC de construir esa Cúpula del Conocimiento Bibliotecario? Tendremos que estar al tanto.

## 2. THING-ATHON

Organizado por ALA Publishing y el RSC (RDA Steering Committee) Development Team, se celebró el 7 de enero en la Biblioteca Lamont de la Universidad de Harvard un nuevo Hackaton relacionado con RDA, denominado Thing-athon. El objetivo del evento era crear metadatos utilizando [RDA: Resource Description & Access](#) y el editor [RIMMF](#), con el fin de explorar las posibilidades de RDA fuera y más allá de un entorno MARC. Siguió la estela de los anteriores [Jane-athons](#) celebrados en torno a la obra de Jane Austen en 2015 en el ALA Midwinter Meeting de Chicago<sup>1</sup> y en la conferencia anual de ALA de San Francisco.

Los asistentes se distribuyeron en mesas especializadas en diferentes temas y materiales, mesas que debían haber elegido previamente: **digitalización**, centrada en la creación de metadatos con RIMMF para recursos digitales y digitalizados; **manuales universitarios**, enfocada hacia la catalogación de diversas ediciones de la misma obra con o sin cambios en el contenido, o de ediciones con cambio en

<sup>1</sup> Para más información sobre el Jane-athon celebrado en Chicago, se puede consultar la información relativa al mismo en la antigua página web del Joint Steering Committee para el Desarrollo de RDA, <http://www.rda-jsc.org/archivedsite/Jane-athon.html>, así como mi informe sobre el ALA Midwinter Meeting de 2015 celebrado en Chicago, [http://www.bne.es/webdocs/Prensa/Noticias/2015/ALA\\_Midwinter\\_2015.pdf](http://www.bne.es/webdocs/Prensa/Noticias/2015/ALA_Midwinter_2015.pdf)



el soporte, pero no en el contenido; **catalogación de partituras; fondo antiguo y materiales raros**, que no solo se dedicó a la producción de metadatos para este tipo de recursos, sino también a la discusión sobre el alineamiento entre RDA y DCRM2 [Descriptive Cataloguing of Rare Materials (Books)], el código de catalogación utilizado para el fondo antiguo en Estados Unidos, tema que ya se había tratado en la reunión del Joint Steering Committee de noviembre de 2015 y en el seminario internacional sobre el tema que se celebró paralelamente; **Strings vs Things** (cadenas de caracteres versus cosas), mesa dedicada a discutir sobre la necesidad de reemplazar cadenas de caracteres legibles por humanos, por “cosas” o identificadores legibles por máquina que enlazan a datos de otras entidades; y, finalmente, **repositorios institucionales, archivos y comunicación académica**, que se dedicó a reflexionar sobre cómo relacionar y enlazar datos de repositorios institucionales y archivos con catálogos de bibliotecas.

Tras la pausa para la comida, Gordon Dunsire, chair del RDA Steering Committee, y Diane Hillmann, de la Dublin Core Metadata Initiative, hablaron sobre las posibilidades que ofrece RDA en un entorno de datos enlazados, con una presentación denominada “Beyond Thing-athon: RDA in the field”. Los temas sobre los que se centró la intervención fueron el impacto de la digitalización, el vocabulario referente al género en RDA y la gestión de identidades.

En la estructura RDA para las manifestaciones de recursos digitalizados, la manifestación de la obra impresa se relaciona con aquellas de la obra digitalizada (puede haber varias dependiendo del formato) mediante una relación “has electronic reproduction” (tiene reproducción electrónica), mientras que la expresión de ambas es la misma. La relación inversa de la manifestación del recurso digitalizado a su original impreso sería “is electronic reproduction of” (es reproducción electrónica de). En los proyectos de digitalización es preciso pensar antes de actuar, tener muy en cuenta todos los aspectos relacionados con la descripción de los recursos, porque se puede dar el caso de que dicha descripción de los recursos esté siendo completamente ineficaz.

Pusieron como ejemplo de caso de estudio para cuestiones relacionadas con la gestión de vocabularios la forma de registrar la información sobre el género de una persona. La norma 9.7.1.3 de RDA define la instrucción sobre el tema y establece tres posibilidades: mujer, hombre y no conocido. La American Library Association envió una solicitud al RDA Steering Committee para añadir el término “transgénero” como una opción más. En RDA se define el género como “aquel con el que una persona se identifica” y existen muchos términos que utiliza la gente para identificar su propio género. Por su parte, la instrucción 8.12.1.3 en RDA establece que el género no forma parte del punto de acceso y que se debe indicar la fuente de la que se extrae la información. Y la regla 0.12 sobre la codificación de datos RDA señala que, para determinados elementos, las instrucciones RDA incluyen un vocabulario definido específicamente para su uso en RDA o se establece el uso de un vocabulario determinado, para lo que es necesario identificar correctamente dicho vocabulario o esquema (por ejemplo, el estándar ISO 3166 para nombres de países y sus subdivisiones).

Un problema añadido a la consignación de datos sobre el género de una persona gira en torno a que las personas transgénero pueden cambiar de género a lo largo de su vida, lo que dificulta el mantenimiento de este tipo de información que, además, es de carácter privado y puede ser también influenciado culturalmente. Además, el gobierno británico y el canadiense han anunciado



recientemente que van a revisar los vocabularios relacionados con el género a nivel nacional en cada uno de estos países. Por todo ello se hace necesario un sistema de gestión de vocabularios que los incluya a todos, por ejemplo, en el Open Metadata Registry.

Estamos, comentaron, en un momento de cambio y caos persistente, pero, para minimizar el caos es preciso auditar el cambio de forma eficiente. Si algo se puede decir que es absolutamente cierto en torno al tema de los datos enlazados es que no hay ninguna verdad ahí fuera. Nada se olvida, nada se borra, pero las declaraciones pueden ser menospreciadas. Cada declaración se copia y se enlaza con otra.

Pero, ¿quién mantiene los identificadores (URIs)? La British Library ha publicado toda su Bibliografía Nacional como datos enlazados y otras instituciones van a hacer lo mismo próximamente. Sin embargo, la pregunta sobre quién se va a encargar del mantenimiento de los identificadores sigue siendo relevante ya que para su persistencia y estabilidad se requiere compromiso, para que sean verdaderamente globales se necesita que estén permanentemente disponibles y para que sean confiables es preciso que su origen, su procedencia ofrezca la mayor confianza.

Debemos pasar de datos cerrados a datos abiertos. Las aplicaciones cerradas, por ejemplo, una base de datos local, no necesitan utilizar URIs y en ellas lo que no se ha grabado no existe. En aplicaciones abiertas, por ejemplo, semánticas, todas las cosas deben tener una URI asociada y lo que no está grabado simplemente es que todavía no se ha hecho. Piensa globalmente, actúa localmente debe ser también la máxima aquí. Es preciso publicar los datasets locales con valores y elementos locales en una estructura global con la debida diligencia, utilizando herramientas abiertas, así como mapeos de los propios elementos o valores locales al más cercano a nivel global.

En relación a la gestión de vocabularios surgen toda una serie de cuestiones fundamentales que hay que tener muy presentes. Los identificadores legibles por máquina resultan imprescindibles para trabajar en un entorno de datos enlazados, pero, ¿cómo se crean y asignan?, ¿quién es responsable de su gestión y mantenimiento?, ¿cómo se garantiza su preservación y disponibilidad a largo plazo?, ¿cuándo es mejor asignar identificadores locales (actuar localmente)?, ¿cuándo es más conveniente reutilizar identificadores globales (actuar globalmente) y pensar en las posibilidades en caso de que dichos identificadores globales desaparezcan (pensar localmente)? ¿qué pasa con los identificadores legibles por humanos, es decir, con el control de autoridades?

Las bibliotecas siempre han lidiado con la identidad de las 'cosas' utilizando cadenas de caracteres, utilizando puntos de acceso como herramientas de gestión y cadenas de caracteres coincidentes, para identificar obras, personas, series, entidades corporativas, etc. La transición desde el uso de cadenas de caracteres como elemento identificatorio a la identidad real basada sobre URIs pone sobre la mesa muchas cuestiones completamente nuevas, que nos obligan a observar con detenimiento el proceso de crear, gestionar y utilizar identidades mediante URIs. La ventaja de esto es que las URIs permiten un nivel de precisión imposible de reproducir utilizando cadenas de caracteres y permite que las máquinas lleven a cabo todo el trabajo.



RDA, en este momento, además de añadir capacidades multilingües, quiere resolver las necesidades de otras comunidades relacionadas con el patrimonio cultural mediante estrategias de extensión. A nivel local, desarrollando nuevos elementos que se puedan utilizar junto a RDA, pero también extendiendo elementos ya existentes de forma que resulten útiles a otras comunidades. Existen cuestiones todavía no resueltas que tienen que ver con la publicación, los mapeos, el mantenimiento y el desarrollo de buenas prácticas, pero una estrategia de extensión puede ayudar a rebajar el nivel para conseguir una cooperación realmente significativa entre comunidades que describen recursos patrimoniales.

Se asume que todo el mundo utilizará vocabularios públicos y sus identificadores, pero los inconvenientes de esa estrategia se centran en que se convierte a sus usuarios en dependientes de los gestores de esos vocabularios públicos que pueden tener otras prioridades o, incluso, pueden llegar a desaparecer completamente, y también hace más difícil la extensión local desde el mismo momento en que los propietarios de dichos vocabularios públicos puedan elegir no reconocer extensiones, por lo que los vocabularios globales deberían ser el mínimo común denominador. Los vocabularios públicos de carácter global deberían reservarse para los metadatos públicos distribuidos globalmente y los vocabularios locales deberían utilizarse para metadatos locales y mapearse a los vocabularios globales públicos, una vez publicados.

Las URIs opacas (números) abordan la volátil realidad de los vocabularios. Las URIs canónicas, basadas en etiquetas se fijan para siempre, incluso aunque las etiquetas sobre las que se basaron inicialmente cambien. Pero los vocabularios cambian, al igual que lo hacen sus etiquetas. Se pueden crear pseudónimos léxicos cuando cambian las etiquetas de manera que los términos originales continúen redirigiéndose a las URIs opacas originales, para mantener la máxima estabilidad.

Tanto si se usan cadenas de caracteres como si se usan 'cosas' (URIs) es necesario establecer buenas estrategias de gestión que se centren en temas como transparencia (versión apropiada para uso de los gestores de datos), conciencia del coste (la toma de decisiones debe estar basada en el conocimiento de las necesidades de los implicados), integración (suministro de servicios a usuarios no angloparlantes, pasando a un primer plano sus preocupaciones) y receptividad (utilizando la tecnología, Github, para mantener el contacto con los usuarios).

Existe también una necesidad de clarificar buenas prácticas. En este sentido, el proyecto de NISO "Bibliographic Roadmap", supone un nuevo esfuerzo para observar con detenimiento todo este entorno y desarrollar buenas prácticas en torno al uso y la reutilización de datos (políticas de gestión, licencias, selección y descubrimiento), su documentación (información sobre vocabularios, su gestión y actualización) y preservación (desarrollo de estrategias en relación a vocabularios abandonados, creación de prácticas de sostenibilidad y recogida de información transversal). ¡Únicamente los catalogadores pueden leer RDA sin ayuda!

Prepárense, señalaron, para comentar el trabajo de NISO durante el período de revisión que se abrirá el próximo año, y consideren si a sus comunidades les gustaría saber más de estos temas. Ambos, Gordon Dunsire y Diane Hillmann, señalaron, están trabajando con NISO en este proyecto y les encantaría escuchar las preocupaciones de todo el mundo.



Para terminar la jornada, Deborah Fritz abrió un turno de comentarios en el que se señalaron las dificultades de importar registros desde MARC a RIMMF, en relación a los repositorios institucionales las debilidades en temas como propietarios o digitalización. Otras cuestiones como los niveles de catalogación en relación a los recursos digitalizados por ejemplo en Internet Archive y si se puede considerar la página web como compilación., interoperabilidad, mapeos, BIBFRAME, también fueron debatidos en las diferentes mesas.

### **3. GRUPO DE INTERÉS DE DIRECTORES DE SERVICIOS TÉCNICOS DE GRANDES BIBLIOTECAS**

La reunión habitual de los directores de departamentos de Servicios Técnicos de las grandes bibliotecas académicas americanas, pertenecientes a la división de ALA [Association for Library Collections & Technical Services](#) fue bastante más heterogénea, si comparamos con reuniones anteriores, en cuanto a la cantidad y diversidad de temas tratados, con un cambio también evidente a nivel organizativo. Depósitos compartidos entre varias instituciones, servicios técnicos deslocalizados, la moral o el desaliento del personal, desarrollo colaborativo de colecciones, modos de compartir datos en un nuevo entorno distribuido o el proyecto LD4P (Linked Data for Production) en torno a BIBFRAME, fueron, entre otros, los asuntos debatidos.

#### **3.1. El futuro del proceso físico y el almacenamiento de materiales en depósitos compartidos de materiales impresos**

Paula Sullenger, de la Universidad de Texas, habló del almacenamiento de impresos en depósitos compartidos, de cómo habían conseguido simplificar el proceso en su Universidad, con una reducción por ejemplo, en los procesos de clasificación de los materiales. Otros trabajos como la asignación de códigos de barras y de tiras de seguridad a los materiales se han seguido realizando como antes. Paula Sullenger invitó a compartir sus experiencias al resto de asistentes.

Beth Camden, de la Penn University, comentó que, en su caso, sí que clasificaban los materiales, aunque solo se hacía constar en los registros, no en el ejemplar físico. Cris Cronin, de la Universidad de Chicago, señaló que ellos sí que estaban formando al personal para clasificar y añadir notaciones a los ejemplares, incluso en el caso de los ebooks. Esta clasificación se hacía con todos los materiales. Beacher Wiggins, de la Library of Congress, también comentó que realizaban todo el proceso.

Scott Wicks, de la Universidad de Harvard explicó que no clasificaban los ejemplares físicos desde hace dos décadas. Cuando implementaron Primo les asaltaron las dudas sobre este tema. Cuando se ofrece la posibilidad de poder ojear a nivel físico, se da también dicha posibilidad a nivel virtual, ya que se les presionó para añadir este tipo de posibilidades de búsqueda (browsing).

Bob Wolven, de la Universidad de Columbia, dijo que, en su caso, habían cambiado el flujo de trabajo, al igual que manifestó la representante de la Penn University, en la que se ha conseguido reducir el número de horas dedicadas al proceso físico de los ejemplares. Esta última también cuestionó el retorno de la inversión en relación a la colocación de tiras de seguridad en determinados



materiales. Sin embargo, esto último no fue compartido por otros directores que, en algún caso, llegaron a plantear la pregunta sobre qué cantidad de materiales especiales se está dispuesto a perder reduciendo la seguridad para conseguir un supuesto ahorro.

Philip Schreur, de la Universidad de Stanford, señaló que ellos están cambiando la dirección de su planteamiento hacia una reducción de la colección principal.

La representante de la Penn University preguntó al grupo sobre los costes del proceso si se encarga a empresas externas. Desde la Universidad de Cornell comentaron que ellos cuentan con una política, que todavía no ha alcanzado el máximo nivel de estabilidad, por la que no invierten en el proceso de materiales no europeos. También comentó que cuentan con un gran número de estudiantes contratados.

Betsy Friesen, de la Universidad de Minnesota, comentó que es necesario contemplar, en aras de la eficiencia, el equilibrio entre el coste y el tiempo del suministro de materiales cuando se ubican en depósitos externos compartidos.

Finalmente, se habló del que iba a ser el tema de discusión futuro en este ámbito, como es el cambio de los flujos de trabajo en un escenario en el que los materiales impresos continúan decreciendo en las colecciones en favor de los electrónicos.

### **3.2. Servicios técnicos externos**

Karla Strieb, de las Bibliotecas de la Universidad de Ohio, centró su intervención en el asunto de los servicios técnicos externos y comenzó realizando una pequeña encuesta a los asistentes en torno a la utilización de instalaciones de servicios externos. Entre otros, John Riemer, de UCLA, señaló que ellos sí que utilizaban instalaciones externas.

Philip Schreur, por su parte, explicó que la situación en Stanford era complicada, con movimientos continuos de ida y vuelta, incluso en el Departamento de Conservación y Preservación. Es fundamental para la toma de decisiones analizar los temas de transporte y, según Schreur, seguir pensando en términos de los flujos de trabajo. Es preciso preguntarse por aquellos cambios más “dramáticos” que se están llevando a cabo.

Jim LeBlanc, de la Universidad de Cornell, comentó que no le gustaría tener que mover los servicios técnicos fuera del campus debido a todos los temas relacionados con datos enlazados, gestión de metadatos, ebooks, etc. Podría comprometerse el futuro de los servicios técnicos, afirmó.

Por otra parte, se señaló que hay funciones que no necesitan presencia física en el campus. Quizás no todos los servicios técnicos puedan trasladarse fuera, pero sí determinados procesos.



### **3.3. ¿Cómo compartimos datos en un nuevo entorno distribuido y colaborativo?**

Philip Schreur, de la Universidad de Stanford, habló de que es preciso reconocer que, en relación a la manera de compartir datos académicos y bibliográficos, nos estamos moviendo a un escenario completamente diferente. Estamos creando muchísimos más metadatos y, no solo debemos pensar en cómo compartirlos como datos enlazados, sino también en cómo dar acceso a todos los datos que la universidad crea. Tal nuevo escenario va a ser muy diferente del modelo actual de OCLC, por ejemplo, para compartir metadatos. Existirán miles de nodos Sparkl. En este sentido, Zepheira está trabajando en el modo de tomar todos esos metadatos e introducirlos en la web.

Si diez bibliotecas diferentes comparten sus datos, va a ser preciso trasladarse desde un entorno local a un entorno global y mundializado. Explicó que tan solo será necesario almacenar los datos sobre los ejemplares a nivel local, sin embargo, el resto de datos serán almacenados en algún sitio externo. Philip Schreur señaló que las grandes bibliotecas académicas deben desplazarse lo más rápidamente posible a este escenario. Aunque también es probable que las más pequeñas tarden algo más y algunas no lleguen jamás a efectuar ese cambio.

### **3.4. Impactos sobre los servicios técnicos del desarrollo colaborativo de colecciones**

Beth Camden partió de una experiencia propia en la de la Penn University de colaboración con universidades del norte de Brasil para hablar del desarrollo de colecciones en un entorno colaborativo. Una de las principales cuestiones que se planteó fue la necesidad de compartir acuerdos sobre comunicación, flujos de trabajo, tareas, etc.

Chris Cronin sugirió que el desafío no es solo MARC, sino la utilización de pequeños consorcios locales. Existe una tensión entre lo local y las necesidades de datos a nivel más global. Martin Kurth, de la Universidad de Yale, comentó que no solo existe esa tensión que señalaba Chris Cronin, sino también otra entre una visión más realista y ligada al presente y otra más futurista hacia lo que está por venir.

Beth Camden planteó otra cuestión en relación a cómo deberían ser y qué características debería tener un acuerdo de cooperación para el desarrollo colaborativo de colecciones para ser realmente bueno. Chris Cronin comentó que no podemos permitirnos quedarnos como los únicos que procesan y desarrollan colecciones, sino que debemos pensar en cómo desarrollar el trabajo de nuestros recursos humanos en entornos colaborativos.

### **3.5. Estructura y moral del personal en los servicios técnicos**

Philip Schreur contó que las Bibliotecas de la Universidad de Stanford cuentan con una nueva iniciativa para ser más abiertos y transparentes en cuanto a la información sobre sus recursos humanos, rango de salarios, etc. Señaló que estaban trabajando y pensando en cómo mantener la



moral del personal, aunque comentó que es muy difícil, fundamentalmente porque Estados Unidos es un país muy caro y es complicado mejorar la moral y motivar al personal, sin mejorar los salarios.

Han trabajado sobre dos puntos básicos. En primer lugar, las bibliotecas siempre han sido lugares de trabajo atractivos para la gente y, en segundo lugar, están intentando cambiar puestos bibliotecarios tradicionales por otros más atractivos. Cada vez son más necesarias tareas relacionadas con la programación, bibliotecarios especializados en el ámbito digital, por lo que se potencian estas áreas y ello resulta injusto para los puestos de trabajo tradicionales de servicios técnicos. Además, estas personas cuentan con un conocimiento especializado en servicios técnicos de un gran valor que no quieren perder, pero también es verdad que su formación se restringe a un área muy específica.

Desde la Universidad de Minnesota comentaron que al personal especializado en catalogación original se le había ofrecido la posibilidad de trabajar con el personal de la biblioteca digital y, gracias a ello, estaban adquiriendo nuevas habilidades. Beacher Wiggins expuso que en la Library of Congress estaban trabajando en la renovación y reestructuración de las actividades relacionadas con los servicios técnicos y, por ejemplo, los bibliotecarios ahora catalogaban incluso aunque trabajaran en el departamento de Adquisiciones. Todo el personal es capaz de asumir muchas más tareas nuevas, dijo, por lo que estaban procediendo a esa renovación de las descripciones de los puestos de trabajo. En la Universidad de Columbia, explicaron también, en la actualidad el Departamento de servicios técnicos cuenta con los puestos de trabajo más cualificados de todo el Servicio de Bibliotecas.

Otros beneficios que se apuntaron en relación al trabajo en los servicios técnicos es la flexibilidad, ya que es posible plantear temas como el teletrabajo. En este sentido, Jim LeBlanc, de la Universidad de Cornell, apuntó que se habían visto obligados a recortar su presupuesto de forma brutal, por lo que, para reducir costes, habían intentado mejorar aspectos como la flexibilidad, potenciando el teletrabajo, entre otras cuestiones.

Finalmente, desde la Universidad de Columbia se lanzó la pregunta de si no deberíamos bajar nuestras expectativas para alcanzar un equilibrio. Por ejemplo, en su caso están tratando los registros realizados por empresas externas como si fueran sus propios registros.

### **3.6. Un modelo consistente y sostenible para los registros bibliográficos de empresas externas**

Bob Wolven, de la Universidad de Columbia, planteó el tema de cómo mejorar los registros bibliográficos en los nuevos sistemas de descubrimiento. Philip Schreur, de la Universidad de Stanford, comentó que parece que se debe optar por diferentes direcciones en función del tipo de registros. Por ejemplo, ellos han creado registros en su propio sistema para unas 100.000 fotografías, con escasos metadatos que, en algunos casos, se han limitado al título.

A nivel de consorcio este tema también es importante, se comentó, ya que no se puede forzar a las Bibliotecas consorciadas a mejorar los registros realizados por empresas externas.



Philip Schreur explicó que, en el caso de la Universidad de Stanford, no consiguen tener lo que necesitan. Les gustaría tener un identificador para cada creador, por ejemplo, por lo que está tratando de concentrarse en este tipo de cuestiones que son las que más les benefician e interesan, y de ser flexibles con el resto.

En la Universidad de Minnesota, en el caso de los registros bibliográficos relativos a ebooks han seguido las recomendaciones que OCLC distribuyó hace algunos años para la revisión de registros bibliográficos realizados por empresas externas de ebooks.

Otro problema que se planteó fue el impacto de la catalogación en el presupuesto global. También cómo ejecutar el gasto a un nivel colaborativo y cómo gastar el dinero de los demás. Finalmente, también se mencionó que sería fundamental tener en cuenta el impacto de los procesos de digitalización.

### **3.7. Actualización de BIBFRAME y beca para el Proyecto LD4L (Linked Data for Production)**

De nuevo Philip Schreur, de Stanford, tomó la palabra para hablar de los avances de BIBFRAME del que señaló que es un objetivo, un blanco en movimiento. Necesitamos movernos, dijo, en dirección a un escenario completamente diferente en relación a los datos enlazados, pero la transición es compleja.

Con una beca de la Mellon Foundation, seis grandes Bibliotecas académicas americanas - Columbia, Cornell, Harvard, Library of Congress, Princeton y Stanford - han estado trabajando desde 2014 en el proyecto LD4L (Linked Data for Production). En 2015 presentaron un informe de actividad a la Mellon Foundation. El objetivo era trabajar en la transición a ese nuevo escenario de datos enlazados, al que no solo las bibliotecas se dirigen.

Se ha estado trabajando en cómo producir metadatos para las colecciones de todas ellas hasta que BIBFRAME esté completamente consolidado. En este sentido, en Stanford han desarrollado una extensión de BIBFRAME para recursos musicales, en colaboración con la Music Library Association.

Es preciso desarrollar un flujo de trabajo completo, de principio a fin, que cubra todas las etapas del proceso incluidas la catalogación por copia, la catalogación original, el trabajo con registros no MARC y las colecciones especiales en sus repositorios digitales. Cuando se habla de todo el proceso se entiende desde la adquisición de los materiales a su descubrimiento. Cornell está trabajando en desarrollar herramientas que se usen en este tipo de entornos de descubrimiento.

El personal de la Library of Congress, por su parte, ha estado trabajando desde septiembre en la creación de descripciones directamente en BIBFRAME. En Cornell están también trabajando en la creación de flujos de trabajo para todas las tareas.



En Canadá también se están desarrollando proyectos de datos Enlazados, en los que se está analizando fundamentalmente BIBFRAME y temas como los identificadores, entre otros.

Philip Schreur, de Stanford, comentó que en esta transición a datos Enlazados se han dado cuenta de que dependen mucho de los registros de empresas externas, por lo que es muy importante trabajar con ellos en este tema.

La Biblioteca Nacional de Medicina por su parte indicaron ha estado trabajando activamente con la Library of Congress y han publicado la [segunda versión beta de sus Encabezamientos de Materia en RDF](#).

### **3.8. Actualización sobre el proyecto piloto de catalogación cooperativa de CIC**

Cris Cronin (Universidad de Chicago), y Mary Laskowski, (Universidad de Illinois Urbana-Champaign), comentaron los avances del Proyecto piloto de CIC (Committee on Institutional Cooperation ), sobre catalogación cooperativa de recursos en otras lenguas, para las que se requiere un conocimiento experto.

Comentaron un nuevo proyecto que acababa de terminar sobre catalogación por copia, del que van a publicar un informe próximamente. Uno de los resultados del proyecto ha sido la determinación de costes de catalogación de materiales en otras lenguas que difieren mucho de unas a otras. El coste medio se sitúa en unos veinticinco dólares, que ascienden a treinta dólares en el caso de materiales en lenguas no latinas.

Animaron a la participación en el Proyecto ya que necesitan que se impliquen más instituciones. Y recordaron que uno de los grandes beneficios de CIC (Committee on Institutional Cooperation ), es el ahorro en costes y esfuerzos.

## **4. MARC ADVISORY COMMITTEE (1)**

La primera reunión del MARC Advisory Committee se produjo a primera hora de la mañana del sábado. Tras la rueda de presentaciones habitual, Mathew E. Wise, Chair del Comité dio paso a la aprobación de las actas de la reunión anterior de junio, que fueron aprobadas por unanimidad. A continuación, comenzó el debate sobre siete documentos presentados por diferentes instituciones. Toda la información sobre la agenda y los documentos y propuestas a debatir se puede encontrar en: [https://www.loc.gov/marc/mac/mw2016\\_age.html](https://www.loc.gov/marc/mac/mw2016_age.html).



#### **4.1. DP 2016-DP07: Ampliación del uso del campo 257 para incluir Regiones autónomas en el Formato MARC 21 para registros bibliográficos**

El documento DP 2016-DP07 fue presentado por Catherine Gerhart, representante de OLAC (Online Audiovisual Catalogers), institución responsable de su envío al Comité, quien explicó que el campo 257 era muy utilizado por catalogadores de videgrabaciones para hacer explícita la información sobre el país en el que se localizan las empresas productoras de las películas. Existen otras regiones, comentó, que habitualmente aparecen también en las listas de películas por países, a pesar de no serlo, como Hong Kong y Palestina, por lo que sacar a la luz dicha información para el usuario resulta importante. Muy pocos cambios en el formato son necesarios, dijo, para poder permitirlo.

Se preguntó si se suponía que todas las regiones autónomas del mundo, que son muchas, estarían incluidas y la representante de OLAC respondió que siempre que fueran productoras de películas deberían estarlo. También explicó que este problema tenía más que ver con temas de financiación que con las características especiales de las películas realizadas en una determinada región autónoma.

La mayor parte de los asistentes no identificaban problemas relativos a la ampliación del campo 257 para incluir regiones autónomas y se aprobó su regreso ya como propuesta para el siguiente encuentro del Comité.

#### **4.2. DP 2016-DP09: Coding Named Events in the MARC 21 Authority and Bibliographic Formats**

Jay Weitz, de OCLC, presentó este documento sobre la codificación de acontecimientos con nombres formales, como batallas, terremotos, erupciones volcánicas, guerras, hambrunas, huelgas, etc., en los Formatos MARC21 tanto de Autoridades como de Registros Bibliográficos que, hasta ahora, se habían venido codificando en el campo 611. Esto ha tenido como consecuencia una gran confusión para los catalogadores ya que la definición del campo 611 limita su uso a congresos y conferencias, así como a otro tipo de acontecimientos que se relacionan en el [Apéndice E](#) del Formato MARC21 para registros bibliográficos, sin embargo, este apéndice establece que unos se deben codificar en 650 y otros en X11. Dado que la implementación de FAST (Faceted Application of Subject Terminology), que clasifica la información relativa a materias en diferentes facetas, necesita una codificación clara de aquellas se presentan dos posibles soluciones para solventar la ambigüedad actual en la codificación:

1. Redefinir las etiquetas X11 para incluir los acontecimientos con un nombre formal.
2. Definir una nueva serie de etiquetas X47 para codificar los acontecimientos con un nombre formal.

Mathew H. Wise, chair del Comité, expuso que de los comentarios enviados parecía deducirse una preferencia por la segunda opción, por lo que propuso sondear si era así entre los asistentes. Excepto



el representante de la British Library, todos los asistentes mostraron su preferencia por la segunda opción. El representante de la Deutsche National Bibliothek explicó que ellos preferirían optar por una tercera opción ya que es necesario diferenciar también los agentes.

Dada la preferencia mayoritaria por la segunda opción se aceptó que se presentara como propuesta en el siguiente encuentro el uso del campo X47 con definición de subcampos.

### **4.3. DP 2016-DP13: Designación de una definición en el Formato MARC 21 de Autoridades**

Reinhold Heuvelmann, representante de la Deutsche National Bibliothek, presentó el documento para la definición de un nuevo campo, preferiblemente 668 (o entre los campos 667 y 68X), para la introducción de información relativa a la definición de entidades.

Desde la Biblioteca del Congreso comentaron que no veían la necesidad de crear un nuevo campo ya que el campo 680 podría utilizarse para incluir ese tipo de información. También explicaron que con los nuevos vocabularios de género/forma de la Biblioteca del Congreso estaban dando este tipo de definiciones en ese campo 680.

John Myers, representante del Committee on Cataloging: Description and Access (CC:DA), preguntó si sería posible introducir nuevos subcampos para explicitar este tipo de información. La representante de las Bibliotecas y Archivos de Canadá sugirió la posibilidad de añadir un nuevo indicador. Finalmente, se indicó que se explorara la posibilidad de utilizar los campos 680 o 678, introduciendo algunos ajustes en la presentación final como propuesta en la siguiente reunión del Comité.

### **4.4. DP 2016-DP14: Designación del Tipo de Entidad en el Formato MARC 21 de Autoridades**

Este documento también fue presentado por la Deutsche National Bibliothek y proponía cómo codificar qué tipo de entidad se está describiendo en un determinado registro de autoridad. Se sugirió en la discusión que sería más apropiado expandir la estructura de subcampos actual en lugar de crear un nuevo campo, o utilizar el campo 072 para este propósito a la hora de plantearlo como propuesta para la próxima reunión del Comité.

### **4.5. DP 2016-DP12: Designación de información coincidente en el Formato MARC 21 de Autoridades**

Este documento también fue presentado por la Deutsche National Bibliothek y proponía cómo expresar que dos registros de autoridad contienen información coincidente y pueden ser el mismo. En la discusión surgió la preocupación sobre la naturaleza temporal de ese tipo de información. El representante de OCLC comentó que este tipo de información no se hace explícita en OCLC, el algoritmo que utilizan no necesita un campo de este tipo. La representante de la Biblioteca Nacional



de Medicina comentó que cada institución puede comportarse de una manera diferente en este sentido y que no vía la necesidad de que este tipo de información esté recogida en MARC. El representante de la British Library sugirió que esta posibilidad debería extenderse también al formato bibliográfico, lo que parecía lógico a la Biblioteca del Congreso. Desde OCLC se insistió en que solo en ciertos casos podría resultar interesante este tipo de información, pero no en cada registro que coincide con otro. Mathew H. Wise, chair de MAC, recordó que en el documento alemán se señalaba que se estaba hablando en todo momento de información temporal. Se acordó que se presentara como propuesta en la siguiente reunión del Comité.

#### **4.6. DP 2016-DP08: Eliminación de la restricción sobre el uso de las fechas en el campo 046, subcampo \$k, del Formato MARC 21 para registros bibliográficos**

El documento, presentado por OLAC (Online Audiovisual Catalogers), proponía la eliminación de la frase “las fechas contenidas en el subcampo \$k pueden no estar codificadas en otro lugar del formato” que, actualmente, se encuentra recogida en la definición del subcampo \$k.

El representante alemán habló de que sería necesario recrear la historia del campo y rastrear por qué se registra información redundante. En general, no surgieron muchas objeciones. Se sugirió que se volviera a plantear como propuesta, pero realizando un estudio exhaustivo de cómo se recogen las fechas a lo largo del formato.

#### **4.7. DP 2016-DP10: Definición del campo 347 (Características del Archivo digital) en el Formato MARC 21 de Holdings**

Este documento fue presentado por CONSER (Cooperative Online Serials Program), del Program for Cooperative Cataloging, y proponía la definición del campo 347 (características del archivo digital) también en el formato de Holdings para registrar información técnica específica relativa a la codificación digital de tipos de datos como texto, imagen, audio, video u otros en el recurso. La discusión giró en torno al uso de los subcampos \$3 (usado habitualmente) y \$8 (preferible para un uso automatizado de la información, aunque no incluido de forma habitual en los Sistemas de gestión bibliotecaria). Se sugirió su planteamiento como propuesta para la próxima reunión del Comité.

Finalmente, la Biblioteca del Congreso comentó que BIBFRAME seguía progresando en cuanto a su desarrollo y animó a asistir a la jornada sobre actualización de BIBFRAME del domingo. Por su parte, el chair del Comité recordó que los próximos encuentros del Comité tendrían lugar en la Conferencia anual de ALA de 2016 en Toronto los días 25 y 26 de junio.



## 5. FORO DE ACTUALIZACIÓN DE RDA

### 5.1. Progresos de RDA en relación a la gobernanza y la estrategia / Gordon Dunsire

Gordon Dunsire, chair del RDA Steering Committee, comenzó hablando de los cambios que se habían acordado en la reunión de noviembre de 2015 en Edimburgo. Por un lado, se había producido un cambio de denominaciones: el Committee of Principals for RDA había pasado a denominarse RDA Board y el Joint Steering Committee for development of RDA a RDA Steering Committee. Por otra parte, se había cambiado el logo y se había diseñado una nueva web, en la que el RDA Board cuenta con un sitio específico: <http://www.rda-rsc.org/>.

Ha cambiado el número y origen de los miembros del RDA Steering Committee también. Habrá ahora un representante por cada una de las seis regiones definidas por Naciones Unidas: África, Asia, Europa, América Latina y el Caribe, Norteamérica y Oceanía. Tanto Europa como Norteamérica que contaban con tres representantes cada una, tendrán que seleccionar un único representante a partir de ahora. Europa estará representada por un miembro de EURIG, pero hasta el próximo encuentro de RDA, que tendrá lugar en Frankfurt en 2016, la British Library ejercerá la representación de CILIP y la suya propia y, junto a la Deutsche Nationalbibliothek, representarán a Europa hasta tanto se produzca la elección del representante europeo por EURIG.

La estructura final del RDA Steering Committee sería la siguiente:

- Miembros ya designados: Chair of RDA Board, ALA Publishing, Chair, Secretary, Examples Editor.
- Miembros designados para el período de transición: Europa, Technical Team Liaison, Translations Team Liaison.
- Miembros todavía no designados y en una fase preliminar de organización del período de transición: Norteamérica, Asia, América Latina y el Caribe, África, Oceanía y la representación del Wider Community Engagement.

Los grupos de trabajo actualmente activos son, con carácter permanente: Technical y Translations. Temporales y para tareas concretas: Aggregates, Music, Places, Capitalization instructions, Fictitious Entities, RDA/ONIX Framework, Relationship Designators. Además, se está constituyendo en estos momentos un grupo de trabajo sobre Materiales raros. Finalmente, se está trabajando en la formación de un grupo de trabajo sobre Archivos.

La estrategia de RDA en la actualidad tiene como objetivo tres comunidades diferentes: comunidades internacionales, comunidades ligadas al patrimonio cultural y comunidades en torno a los datos enlazados.

Durante los próximos dos/tres años se trabajará en la transición, señalaron, hacia el nuevo modelo de gobernanza y se procederá a llevar a cabo una revisión y reorganización del RDA Toolkit. Por otro lado, los modelos conceptuales necesitan consolidarse en torno al nuevo FRBR Library Reference Model del que se publicará un borrador en 2016. Una vez aprobado por IFLA se espera que sea



publicado en 2017. Su publicación impactará sobre RDA lo que obligará a nuevas reorganizaciones del RDA Toolkit.

También se señaló que se habían desarrollado otras actividades más específicas:

- **Integración internacional.** ALA Publishing ha diseñado una nueva política de traducción por la que la traducción de RDA a una lengua implica necesariamente tener que traducir el RDA Reference, o sea, las etiquetas y definiciones de las entidades, elementos y vocabularios RDA para que sean añadidos al RDA Registry. Como resultado, actualmente en el RDA Registry ya se encuentran dichas traducciones al chino, francés, alemán, español e inglés. Esto permite a las aplicaciones la interoperabilidad automática de los datos a partir de las agencias nacionales utilizando diferentes lenguas y muestra como una única URI para un tipo de soporte puede asociarse con etiquetas y definiciones en diferentes lenguas.
- **Refinamiento local de elementos y valores.** RDA se está desarrollando para permitir refinamientos y extensiones de elementos y vocabularios RDA para aplicaciones locales. Por ejemplo, la Library of Congress ha propuesto “audio belt”, implementación local de una categoría de recursos básico en la ontología RDA/ONIX. Estos refinamientos se dan en un mapeo legible por máquina en el RDA Registry.
- **Modelos y ontologías básicas.** RDA/ONIX es la base para todos los tipos de soportes y tipos de contenidos que determinan una distinción clara entre contenido y soporte. Está disponible en formato de datos enlazados legible por máquina. El RDA Steering Committee y los grupos de trabajo están investigando el impacto de los cambios en el modelo FRBR y ontologías asociadas. El nuevo FRBR Library Reference Model es compatible con la extensión orientada a objetos del modelo CIDOC-CRM desarrollado para la comunidad museística, lo que permitirá el desarrollo de RDA para las comunidades relacionadas con el patrimonio cultural. Por otra parte, FRBRoo ha sido desarrollado por la Biblioteca Nacional de Francia como modelo orientado a objetos en relación a las publicaciones periódicas.
- **Procedencia.** El tratamiento de los datos de procedencia ha sido problemático en RDA. Se está trabajando en la clarificación de los elementos de notas para diferenciar aquellos que proporcionan datos adicionales para el valor del elemento asociado de aquellos que tratan sobre el valor del elemento asociado. Además están trabajando en la arquitectura completa de los elementos. Esto llevará su tiempo, por lo que, explicó, que no podía prometer que estuviera disponible el año próximo.

Finalmente, animó a involucrarse y participar en los grupos de trabajo de RDA, a estar al tanto del Foro de Desarrollo de RDA, a participar como representantes en el RDA Steering Committee y en el RDA Board, a seguir las discusiones en la lista de discusión RDA-L y en el proyecto RDA Vocabularies en Github, así como a participar en los próximos Jane-athons que tendrán lugar en Francia, España, Latvia, etc. La estrategia de RDA depende, recordó, de la participación activa de las comunidades RDA y de sus miembros individuales.

## 5.2. Actualización de RDA Toolkit / James Hennelly

A continuación, James Hennelly, editor del RDA Toolkit, habló de las tres actualizaciones llevadas a cabo desde la Conferencia anual de ALA de junio de 2015 y de la próxima que se producirá en febrero de 2016, con reajustes en las traducciones alemana y española. La traducción al finés se publicó en



diciembre. En marzo se publicará una nueva actualización en la que se espera incluir la traducción al italiano y traducciones al catalán y al noruego están en camino. Se van a añadir también las traducciones francesa e italiana de las políticas alemanas. Por otra parte, la segunda edición de DCRM2 está en camino y probablemente se publicará en 2017, comentó.

RDA Toolkit es cada vez más internacional y han aumentado de forma significativa las suscripciones desde Europa. También se está incrementando su grado de complejidad, por lo que se hace necesaria una reorganización para que sea una herramienta de verdadera utilidad para todo el mundo. Será posible establecer preferencias por lengua y mostrar aquellos links que se quieran seleccionar, por ejemplo.

También habló de la nueva web del RDA Board. El contenido del antiguo JSC está archivado y es posible su consulta. Respecto a la reorganización y revisión del Toolkit, comentó que estaban en una etapa de preplanificación y que los trabajos comenzarían a finales de 2016. Se incluirán historias de usuarios.

Comentó que en abril se publicará un nuevo libro sobre RDA titulado *RDA Essentials*, escrito por Thomas Brenndorfer. También comentó que este año no se publicaría ninguna versión impresa de RDA. Finalmente, recordó que el lunes 11 de enero se celebraría otro evento sobre RDA también en el Convention Center de Boston en el marco del Midwinter Meeting de ALA.

## 6. GRUPO DE INTERÉS SOBRE GESTIÓN DEL CATÁLOGO

### 6.1. Que estás haciendo, ¿qué?!? Metadatos, migración y fusión en Ex Libris / Kevin Clair, Erin Elzi

Kevin Clair y Erin Elzi, de la Universidad de Denver, hablaron sobre el proceso de migración a un nuevo Sistema de Gestión de Bibliotecas (Alma, de Ex Libris), que iban a llevar a cabo a mediados de 2016, así como sobre la fusión de dos catálogos, que incluyen metadatos no MARC de material de archivo. Hasta ahora, se habían desarrollado procesos de exportación de metadatos MODS y MARC desde ArchivesSpace al SIGB Sierra, y contaban [scripts de exportación desde ArchivesSpace](#), pero este proceso tendrá que cambiar con la migración a Alma. Por otra parte, las Bibliotecas de la Universidad de Denver y la Iliff School of Theology habrán de fusionar sus catálogos, pero mientras la primera utiliza Sierra, la segunda utiliza Horizon como SIGB, lo que introduce un nuevo factor de complicación a todo el proceso.

Dos departamentos están implicados en el Proyecto global, comentaron, el de Catalogación y el de Materiales de archivo y Colecciones especiales. En el verano de 2014 se produjo un proceso de migración del segundo a ArchivesSpace, lo que permitió mejorar la eficiencia en las estrategias de exportación de datos, así como establecer conexiones con el SIGB. A posteriori han desarrollado un flujo de trabajo que parte de ArchivesSpace, para después utilizar MarcEdit para preparar los registros bibliográficos para la migración a Sierra. El personal de ambos departamentos ha trabajado conjuntamente y con mucha implicación en el proyecto e incluso el personal del Departamento de



Catalogación ha ayudado en la formación del personal de los Archivos en RDA y MARC y se ha llegado a acuerdos en cuanto al uso de terminología de género/forma.

Mientras tanto, el proceso de fusión de catálogos y migración a Alma ha continuado su curso. La Iliff School of Theology y las Bibliotecas de la Universidad de Denver comparten el mismo campus físico, aunque se encuentran ubicadas en edificios diferentes y cuentan con colecciones también diferentes. Las colecciones no se solapan mucho y mientras Iliff cuenta con 250.000 registros bibliográficos, las bibliotecas de la Universidad de Denver cuentan con cuatro millones setecientos mil registros bibliográficos y unos cuarenta y nueve mil usuarios. La primera prioridad era el acceso. Un problema a solventar era el acceso Proxy. Además, era necesario realizar mucho trabajo respecto al tema de las licencias.

Respecto al proceso de migración a Alma, existían muchos sistemas diferentes en las Bibliotecas de la Universidad de Denver que no estaban trabajando juntos, demasiadas cajas de búsqueda diferentes: ARchivesSpace, Islandora, Summon, Serials Solutions, Encore vs Classic. Además, los flujos de trabajo utilizados hasta ahora habían dejado de ser efectivos y flexibles.

La Gestión del Proyecto de migración por parte de la Universidad de Denver la ha llevado a cabo personal de diferentes departamentos de Iliff y de la Biblioteca de Derecho, más personal de Ex Libris, señalaron. Cada semana se celebraban encuentros entre estos dos grupos. Los temas más importantes que debían gestionar eran la extracción de los datos, su depuración y la migración final desde Horizon y Sierra, así como las cuestiones relativas a la gestión de registros electrónicos.

Respecto al futuro del Proyecto, se va a desarrollar un enfoque más centralizado. Se ha detectado que es preciso refinar los procesos de exportación y las tareas de depuración de registros con MarcEdit, establecer una única opción de búsqueda sobre todos los registros de las bibliotecas de la Universidad de Denver, establecer un flujo inverso de datos extrayendo metadatos desde Alma a ArchivesSpace, y, finalmente, publicar metadatos bibliográficos en diferentes contextos.

## **6.2. GMD o No GMD: aproximación de una biblioteca a la conversión a RDA / Jim Kalwara**

Jim Kalwara, de las Bibliotecas de la Mississippi State University, habló de la implementación en 2013 de RDA no solo en la propia Universidad sino también en las ochenta bibliotecas que forman parte del consorcio del Estado denominado Mississippi Library Partnership. Para llevar a cabo la transición desde AACR2 a RDA se establecieron guías y políticas para el consorcio, pero se tomó una decisión contraria a la de muchas bibliotecas de todo el mundo: no omitir las Designaciones Generales del material en los registros bibliográficos de su catálogo y retenerlos durante el proceso de conversión a RDA.

La Designación General de Material (GMD, General Material Designation) define aquellos términos que describen el formato físico de las obras (MARC bibliográfico: 245 \$h). Sin embargo, en RDA desaparece y se sustituye por tres tipos de datos que ofrecen más granularidad a la información



tradicionalmente recogida en la designación general de material, como el tipo de contenido, el tipo de medio y el tipo de soporte (MARC bibliográfico: 336-338 y 007).

Pero, ¿por qué decidir mantener la información relativa a la Designación General de Material? Fundamentalmente, decidieron hacerlo porque apoya las tareas del usuario y ayuda a diferenciar materiales con el mismo título, facilita el descubrimiento al presentar la información de forma breve en el OPAC y ayuda a contextualizar la información recogida en los campos 33x, entre otras razones. En 2014 se comenzó a recopilar información sobre su uso por el personal de referencia y por los usuarios, sobre su utilidad en el caso de los recursos electrónicos y la necesidad de varios tipos de “recursos electrónicos” o los problemas que acarrearía a los sistemas su sustitución por la información recogida en los nuevos campos 33X.

Se desarrolló una política para su conversión de cara a que los registros bibliográficos existentes pasaran a ser conformes con las nuevas reglas RDA, lo que implicaba añadir los campos 33X (en Connexion Client, Browser y en Sirsi), el desarrollo de las abreviaturas y la actualización de la Mención de publicación siempre que fuera necesario. También se procedió a evaluar los términos desactualizados de la Designación General de Material y a desarrollar nuevos términos más útiles. Finalmente se creó una nueva lista de términos con su mapeo correspondiente a los nuevos campos 33X y se comprobó cómo se visualizarían en el catálogo.

Al crear registros bibliográficos originales se debía cumplir con RDA, pero se añadían las Designaciones Generales del Material tras la creación del registro en OCLC Local Save File y en SirsiDynix. El proceso de mantenimiento no es perfecto, comentaron, y todavía es necesario hacer correcciones y seguimiento y tener en cuenta nuevos tipos de ítem. Los procesos de conversión sacan a la luz información incorrecta codificada en los campos 33X y en el 007 y cómo se muestra en el catálogo. Todos estos procesos lo que han venido a demostrar es la granularidad de RDA.

### **6.3. Preparación de materiales impresos para su almacenamiento en un depósito externo: la experiencia de la Universidad de Notre Dame / Nastia Guimaraes**

Las bibliotecas de la Universidad de Notre Dame cuentan con unos ocho mil estudiantes universitarios, tres mil quinientos licenciados, cuatro millones y medio de volúmenes impresos y alrededor de un millón de ebooks. En septiembre de 2012 pasó por un proceso de rediseño organizativo a gran escala, al que se añadieron nuevos cambios en el otoño de 2015.

Ya antes del proceso de renovación tenían necesidad de contar con capacidad de almacenamiento adicional para materiales impresos, por lo que solventar este problema se convirtió en una iniciativa bibliotecaria estratégica y la Universidad adquirió un depósito con una capacidad para, aproximadamente, un millón seiscientos mil volúmenes.

El equipo de gestión del proyecto estuvo formado por un gestor del proyecto contratado en 2014, al que se unió en octubre de 2015 un bibliotecario especializado en servicios técnicos a tiempo



completo, a los que, poco tiempo después, se sumaron dos gestores más. Con anterioridad al inicio del proyecto en sí se tuvieron en cuenta otras experiencias similares de universidades de Estados Unidos, como Harvard, que compartieron sus conocimientos y experiencia en el tema. Además, se realizaron algunas actividades más como: análisis del entorno ambiental, visitas a dos depósitos de bibliotecas universitarias situados fuera de sus respectivos campus, informe para los administradores de la biblioteca con recomendaciones y buenas prácticas o identificación del grupo gestor del Proyecto, incluido el Jefe de Proyecto.

Para establecer los criterios de selección de transferencia de los materiales al nuevo depósito externo se creó un Comité, integrado por personal bibliotecario y personal docente e investigador, que decidió incluir títulos de publicaciones periódicas cerradas y para materias científicas, monografías sin registro de circulación durante los últimos diez años.

Para la preparación de la colección se contrató a personal para la asignación de códigos de barra y se consiguió en seis meses alcanzar la cifra de doscientos diecisiete mil ejemplares con código de barras. También se asignaron catalogadores a proyectos específicos de mantenimiento del catálogo y realizaron la depuración de registros a partir del proceso de asignación de códigos de barras. Asimismo, se incorporaron en esta fase voluntarios de otras bibliotecas.

Los desafíos a los que hubo que enfrentarse en esta fase previa fueron: la detección de errores en los registros del catálogo provocó que fuera necesario trabajar en el arreglo de datos con problemas, como por ejemplo títulos que no contaban con un registro en el catálogo. Por otra parte, el tiempo asignado a esta fase fue demasiado estrecho. Pero también hubo tareas realizadas de especial orgullo como los cuatrocientos cincuenta mil ejemplares preparados para la transferencia en seis meses, el desarrollo de un Sistema de gestión de inventario a medida en cinco meses o que el proyecto de transferencia tuvo poco impacto en los usuarios.

Se realizó a cabo un informe final global con la totalidad de los títulos transferidos, se cambió la localización de los holdings a todos ellos, y se depuraron problemas de catalogación y de localización de ejemplares perdidos en el proceso de transferencia. Para la preparación de los materiales se recogió información del grosor de cada ítem a transferir, lo que sirvió de ayuda para la evaluación de los espacios necesarios, y se grabó información sobre su estado de conservación.

Respecto al desarrollo de un Sistema de gestión de inventario a medida, se decidió su creación dado que no encontraron ninguna herramienta en el mercado que satisficiera todas sus necesidades. Establecieron los requerimientos necesarios y contrataron un desarrollador para ello. Se trata de una herramienta Open source (<https://github.com/ndlib/annex-ims>) y quieren desarrollar el código con posibles socios.

Si pudieran comenzar de nuevo, comentaron, contratarían más personal, comenzarían antes a llevar a cabo todo el trabajo preparatorio y ampliarían el estudio de capacidad a una muestra más amplia para conseguir unos resultados más precisos en relación a necesidades de equipamientos.



## 7. GRUPO DE INTERÉS EN LA TRANSICIÓN DE LOS FORMATOS MARC

### 7.1. Datos enlazados abiertos para el Descubrimiento: BIBFRAME + Schema.org en una vista experimental Bento / Jim Hahn

Gracias a la financiación con una subvención interna a la innovación de la Biblioteca de la Universidad de Illinois en Urbana-Champaign, investigadores consiguieron transformar y enriquecer alrededor de 300,000 registros bibliográficos correspondientes a e-books de su catálogo desde MARC a BIBFRAME. Esta financiación permitió que estudiantes graduados de los Departamentos de Informática y de Biblioteconomía y Ciencias de la Información trabajaran por horas en el proyecto.

Los objetivos eran enriquecer el modelo BIBFRAME con datos enlazados que conectaran a otros proyectos de datos enlazados abiertos, así como contribuir a modular vistas Bento de los datos del catálogo en el modelo BIBFRAME.

El proceso consistió en extraer datos en MARC XML, transformarlos a BIBFRAME XML, enriquecerlos con datos enlazados e indexarlos con Google Custom Search Engine. Para el proceso de extracción de MARC XML se generaron consultas para reunir identificadores bibliográficos únicos de todos los ebooks de la biblioteca y se utilizaron estos identificadores para generar MARC XML que, a su vez, se introdujo en el código de transformación de la Biblioteca del Congreso marc2bibframe. Con Schema.org se validaron datos enlazados embebidos en HTML. Han desarrollado dos prototipos de interfaz de búsqueda: un motor de búsqueda a medida que saca a la luz los datos estructurados en la lista de resultados y una vista tipo Bento para la búsqueda de ebooks.

Entre las limitaciones del Proyecto señalaron que el modelo actual solo incluye relaciones Obra a Manifestación y Manifestación a Obra (hay relaciones Obra a Obra y Manifestación a Manifestación que no se han modelado aquí), o que es un prototipo para la transformación de información sobre ebooks, pero no se ha implementado la importación de nuevos datos sobre ebooks. Se trata tan solo de un prototipo.

Desde el público se planteó la pregunta de si se habían visto en la situación de contar con una Persona sin registro en VIAF y tener que crear la URI o el identificador. La respuesta es afirmativa y explica que utilizaron ORCID, pero fue complicado. Es algo que tienen que analizar con más detalle, señalaron.

Más información sobre el Proyecto se puede localizar en: <http://sif.library.illinois.edu/bibframe/>.

### 7.2. De MARC a MODS XML y a los amigables datos enlazados RDF / Steven Anderson

Steven Anderson, de la Boston Public Library, habló sobre la experiencia de la Boston Public Library en la transformación de registros MARC a MODS XML para terminar convirtiéndolos en RDF.



Comenzó advirtiendo que su punto de vista era, no el de un experto, sino el de alguien dando traspies en la oscuridad.

[Digital Commonwealth](#) es una organización sin ánimo de lucro que proporciona recursos y servicios para la creación, gestión y diseminación de materiales del patrimonio cultural custodiados en bibliotecas, archivos, museos y sociedades históricas de Massachusetts y su sitio web proporciona acceso a miles de documentos digitalizados por las instituciones que forman parte de ella. Cuenta con una colección de 228.000 ejemplares, en 521 colecciones, de 138 instituciones diferentes.

Recientemente tuvieron que introducir 40.000 libros digitalizados por Internet Archive en Digital Commonwealth, de tal modo que fueran accesibles y se pudieran buscar junto al resto de objetos digitales. Los metadatos de estos libros provienen de MARC y tienen datos descriptivos detallados, sin embargo, el repositorio está basado en MODS. Era necesario, por tanto, que hablaran el mismo lenguaje. Se decidió que era necesario traducir los registros MARC a MODS para poder incluirlos en Digital Commonwealth, para lo que se utilizó un script a medida. Por ejemplo, utilizaron Geomash para enlazar a TGN/Geonames para la denominación normalizada de nombres geográficos y para conseguir las coordenadas geográficas o analizaron las fechas utilizando BplEnrich para introducir las en la estructura MODS en un formato w3cdtf.

Los registros MARC van a continuar existiendo, de hecho, su sistema de gestión de bibliotecas utiliza MARC y cuentan con millones de registros en ese formato. Y, además, su sistema de gestión de bibliotecas, POLARIS, va a continuar siendo su fuente de metadatos “original” para este tipo de ítems por ahora.

Datos enlazados y RDF son el futuro. La tecnología mejora y ni MODS XML ni MARC son ya los formatos ideales para compartir datos. La Biblioteca del Congreso cuenta con una representación RDF de MODS, por lo que parecía que iba a resultar fácil el poder “unirse a la fiesta”.

Diversas instituciones como la Biblioteca Pública de Boston, la de New York, la Universidad de Emory, Amherst College y la Universidad de Indiana se unieron para construir un camino que les permitiera introducirse en el mundo de los datos enlazados RDF:

<https://wiki.duraspace.org/display/hydra/MODS+and+RDF+Descriptive+Metadata+Subgroup>.

Están utilizando un montón de namespaces, no un único namespace, para que otros fuera del mundo bibliotecario puedan ser capaces de leer y comprender los datos bibliotecarios. Entienden que sí existe un predicado y se usa para un concepto que te estás planteando crear, es preferible usarlo antes de crear tu propia “isla”.

BIBFRAME es impresionante, pero han preferido utilizar predicados utilizados a gran escala siempre que fuera posible (esto es, múltiples namespaces). Existen muchos huecos en los vocabularios existentes que será genial que BIBFRAME rellene, pero en ocasiones los términos que han mirado son demasiado restrictivos y, a menudo, muy complicados.



El proyecto todavía no ha terminado. MODS es complejo y su mapeo requiere tiempo, por lo que es probable que no se complete hasta mediados de 2016. Steven Anderson terminó su intervención afirmando que MARC21 va a seguir en nuestros catálogos al menos durante algún tiempo.

Desde el público se plantearon algunas cuestiones cómo la razón de no utilizar sencillamente BIBFRAME, a lo que contestó que hay demasiados predicados en BIBFRAME y es demasiado complejo, es mucho más descriptivo. También se preguntó sobre el número de instancias que estaban usando, a lo que respondió que más de cinco, pero menos de diez.

La Chair del Grupo de Interés en la transición de los formatos MARC señaló que tenemos serios problemas de comunicación en la comunidad y que este ejemplo era ilustrativo de cómo hay gente trabajando en estas áreas sin esperar a que la Biblioteca del Congreso diga: ¡Ya estamos aquí! ¡Ya hemos terminado BIBFRAME!

### **7.3. Datos enlazados para Archivos / Gloria Gonzalez (Zepheira), Dustin Stokes (Atlas Systems)**

Zepheira y Atlas Systems comenzaron a trabajar en colaboración hace algunos años para encontrar métodos de incrementar la accesibilidad de materiales raros y archivísticos. Durante los meses de septiembre y octubre de 2015 han trabajado juntos en 6 grupos focales durante nueve días, con un total de 16 participantes.

Entre sus descubrimientos iniciales, señalaron que las herramientas y métodos utilizados para crear descripciones archivísticas varían ampliamente y existe una pluralidad de aplicaciones de normas archivísticas. Se necesita estructurar más profundamente la forma en que los archivos recogen datos sobre cómo la gente encuentra y usa las colecciones archivísticas, señalaron. Para ello, la colaboración es crítica para desarrollar software de código abierto para los archivos y las descripciones de sus documentos.

La comunidad archivística y la de los materiales especiales y raros comparten enfoques centrados en el objeto, señalaron. Se necesitan estrategias de datos enlazados para rastrear la cronología de la propiedad y la custodia de los materiales archivísticos y raros, dijeron.

Atlas Systems y Zepheira van a contactar con instituciones que estén experimentando con RDF para aprender más acerca de sus objetivos y casos de estudio, así como con grupos que ya estén trabajando en ontologías de datos enlazados para archivos, como el Grupo de Expertos de ICA, así como con aquellas personas que estén trabajando en el mapeo de EAD a otro tipo de formatos, como Europeana.

Ya existen algunas soluciones posibles como el plug-in ArchivesSpace para realizar transformaciones de EAD a Datos enlazados, pero es solo el principio y necesitan más ayuda, por lo que animaron a involucrarse en el proyecto y a contestar la siguiente encuesta: <https://goo.gl/fdvN8D>.



Finalmente, anunciaron que a finales de enero publicarían un informe que se podrá solicitar en la dirección anterior.

## 8. GRUPO DE INTERÉS EN METADATOS

Tras los primeros veinte minutos de la sesión dedicados a la aprobación de las actas de la sesión anterior celebrada en el ALA Annual 2015, a realizar un recordatorio de actividades programadas para 2016, así como a los informes de los enlaces de la Music Library Association, de LITA y del CC:DA, se pasó a las presentaciones del Metadata Interest Group de la Association for Library Collections & Technical Services (ALCTS), División de ALA, que se dedicaron al borrador de principios para evaluar estándares de metadatos presentado en octubre de 2015 y al tema de la gestión de derechos de copyright en el ámbito de las bibliotecas digitales.

### 8.1. Borrador de principios para la evaluación de estándares de metadatos / Jennifer A. Liss, Kevin Clair

Jennifer A. Liss y Kevin Clair, ambos integrantes del ALCTS/LITA Metadata Standards Committee (MSC) presentaron el borrador de documento elaborado por el Comité, *Principles for Evaluating Metadata Standards*, que tiene como objetivo ser utilizado por archivos, bibliotecas y museos, para el desarrollo, mantenimiento, selección, uso y evaluación de todo tipo de estándares de metadatos. En enero de 2015 se presentó a discusión una lista de principios y tras un período abierto para comentarios se presentó el borrador de principios en octubre de 2015.

Se trata de siete principios básicos, a saber:

1. **Metadatos y estándares de metadatos deben formar parte de la red.** Estándares, sistemas y servicios de metadatos deben ser tan abiertos como sea posible y tanto las definiciones como el modo en el que están modelados deberían estar disponibles en la web. Deberían hacer uso de modelos de referencia y permitir conectar información de diferentes fuentes, lo que redundaría en una mayor visibilidad de los recursos de las instituciones de la memoria en la web.
2. **Metadatos y estándares de metadatos deben ser abiertos y reutilizables.** Acceso y uso deben ser abiertos y deben combinarse con otros estándares.
3. **La creación de metadatos debe beneficiar a las comunidades de usuarios.** Se necesitan casos claros de uso para permitir a los profesionales establecer el nivel de descripción necesario en cada situación dada.
4. **Los estándares de metadatos deben apoyar los nuevos métodos de investigación.** Los metadatos bibliográficos tradicionales solo son de ayuda en parcelas reducidas de la investigación. Los nuevos estándares de metadatos deben favorecer los métodos de investigación emergentes, tales como la lingüística computacional o el análisis de redes, por poner dos ejemplos, y servirles de apoyo. Es imprescindible que sustenten las tareas del usuario definidas en el modelo FRBR.



5. **Los estándares de metadatos deben contar con una comunidad activa que contribuya a su mantenimiento y gobernanza.** Solo tendrán valor aquellos estándares que cuenten con amplias comunidades que los utilicen. Durante los últimos cuarenta años los estándares de metadatos utilizados en bibliotecas, archivos y museos han sido muy estables, pero no parece que en el futuro esto vaya a seguir siendo así. Contar con una comunidad amplia y activa que contribuya a su evolución será fundamental para su supervivencia y desarrollo. En este sentido, la transparencia va a resultar fundamental en el mantenimiento de los estándares.
6. **Los estándares deberían ser extensibles, integrables e interoperables.** Contribuirá al éxito de un estándar su capacidad de ser utilizado de forma eficaz en conjunción con estándares de otras comunidades. Será recomendable que los estándares proporcionen opciones de implementación muy ligeras para permitir su adopción a un nivel amplio, incluso si eso significa la pérdida de datos en la transferencia en algunas ocasiones.
7. **Los estándares de metadatos deberían seguir las reglas “graceful degradation” (degradación sin pérdida de operatividad) y “responsive design” (diseño adaptativo).** El diseño de un estándar debería permitir su adaptación a nuevos e inesperados usos, así como a usos sofisticados, como la inferencia ontológica.

Al final de la presentación se animó a los asistentes a contribuir con sus comentarios al desarrollo de estos principios tanto a través de la cuenta de Twitter @metawarebuzz, como de la web [metaware.buzz](http://metaware.buzz).

## 8.2. Hacer lo correcto en temas de derechos de propiedad intelectual / Emily Gore, DPLA (Digital Public Library of America)

Emily Gore, directora de contenidos de la Digital Public Library of America, comentó el desarrollo de un conjunto de declaraciones de derechos normalizadas que ha llevado a cabo un comité internacional integrado por representantes de la Digital Public Library of America, Europea y Creative Commons para que materiales digitalizados del patrimonio cultural estén disponibles como URIs, además de la definición de requerimientos técnicos adicionales para la implementación. A comienzos de 2016 se lanzará [rightsstatements.org](http://rightsstatements.org), señaló.

La Digital Public Library of America es una red global de centros y servicios de contenidos. Cuenta con alrededor de 11 millones de registros, 27 centros más 9 en desarrollo, más de 1600 instituciones que contribuyen con sus contenidos, utiliza 9 schemas de metadatos principales (no todos XML) y cuenta con un Perfil de Aplicación de Metadatos, DPLA Metadata Application Profile.

Para todos sus contenidos, la DPLA ha contabilizado 87.000 declaraciones de derechos diferentes, aunque las más utilizadas son: copyright o “Todos los derechos reservados” (45%), Otros (25%), Creative Commons (3%), Public Domain (5%), “Copyright de imágenes digitales” (7%), Missing Rights Statement (7%) y Documentos gubernamentales (8%).



Desde la DPLA se ha trabajado en colaboración con sus colegas europeos de Europeana en el desarrollo de la Estructura de Licencias de Europeana. Cuando se comenzó a trabajar en Europeana los registros MARC que recolectaban no recogían declaraciones de derechos, con lo que ello suponía de confusión. En la actualidad, Europeana ha desarrollado una búsqueda facetada por copyright.

El Proyecto en el que han trabajado de forma conjunta DPLA y Europeana sobre el tema de las declaraciones de derechos de autor se denomina Rights Statements y ha sido financiado por la Knight Foundation. En febrero de 2016 se lanzará [Rightsstatements.org](http://Rightsstatements.org), con el objetivo de alcanzar interoperabilidad a nivel internacional.

Las categorías definidas de derechos son tres: con copyright, sin copyright, indeterminado. Para las primeras, obras con copyright, se definen diferentes subcategorías:

1. In Copyright - InC (Con Copyright)
2. In Copyright – EU Orphan Work (Con Copyright – Obra huérfana en la Unión Europea) (InCOWEU) (No hay un Proyecto de estas características en los Estados Unidos todavía sobre obras huérfanas)
3. In Copyright – Rights-holder(s) Unlocatable or Unidentifiable (Con Copyright – Detentador de los derechos de autor no localizable o identificable) (InCRUU)
4. In Copyright – Educational Use Permitted (Con Copyright – Permitido el uso educativo) (InCEDU)
5. In Copyright – Non-Commercial Use Permitted (Con Copyright – Solo se permite el uso no comercial) (inCNC)

Para las obras sin copyright se han definido las siguientes:

1. No Copyright – NonCommercial Use Only (Sin copyright – Uso no comercial únicamente)
2. No Copyright – Contractual Restrictions (Sin copyright – Restricciones contractuales) (noCCR)
3. No Copyright – Other Known Legal Restrictions (Sin copyright – Otras restricciones legales conocidas)
4. No Copyright – United States (Sin copyright – Estados Unidos)

Y para las indeterminadas:

1. No Known Copyright (Sin copyright conocido)
2. Copyright Not Evaluated (Copyright no evaluado)

Lo que se persigue es utilizar estas declaraciones de derechos junto a las licencias Creative Commons.

En la Guía para el diseño técnico se intenta reflejar que las declaraciones de derechos en desarrollo NO son licencias. Se ha desarrollado como un vocabulario para su reutilización y el modelo de datos se ha publicado como vocabulario SKOS: <http://bit.ly/rights-data-model>. Se desarrollarán documentos de buenas prácticas sobre la publicación de datos enlazados. Puede ser utilizado con



múltiples esquemas de metadatos y puede incluir de forma opcional URIs para representaciones legibles por humanos.

Una vez se hayan completado las declaraciones y se haya lanzado [rightstatements.org](http://rightstatements.org), DPLA ofrecerá recomendaciones a todas las instituciones participantes en su red. La primera sesión de formación para la implementación en los diferentes centros se llevará a cabo en abril en el DPLAFest que se celebrará en Washington DC.

En el grupo de trabajo han participado, entre otros, Paul Keller (co-chair), Emily Gore (co-chair), Mark Matienzo (DPLA), Antoine Isaac (Europeana) o Dave Hansen (UNC/DPLA).

## 9. ACTUALIZACIÓN SOBRE BIBFRAME DE LA BIBLIOTECA DEL CONGRESO

### 9.1. Actualización sobre el programa piloto BIBFRAME de la Library of Congress / Beacher Wiggings y Sally McCallum

En 2014 se publicó la versión 1.0 del vocabulario BIBFRAME, desarrollado a partir de elementos de RDA, MARC y AACR. Ya cuentan con herramientas de conversión de MARC a la versión 1.0 de BIBFRAME, un editor BF, un editor de perfiles BF, así como herramientas de búsqueda que se pueden localizar en la siguiente página web: <http://www.loc.gov/bibframe/tools/>. Asimismo, se han elaborado diversos informes sobre autoridades, relaciones, perfiles, anotaciones, etc.; se ha debatido y se sigue debatiendo profusamente a través de la lista de distribución BIBFRAME y se ha recibido retroalimentación y consejo de expertos en la revisión del vocabulario.

Se pretende mejorar el vocabulario BIBFRAME y hacerlo evolucionar hacia una versión 2.0. En este sentido, se han publicado toda una serie de borradores de especificaciones y documentación para el debate sobre, entre otros, títulos, acontecimientos, agentes, ítems, identificadores y notas, metadatos administrativos, etc.: <http://www.loc.gov/bibframe/docs/index.html>. Se está valorando la revisión del modelo básico en relación al ítem, por una parte, agentes (personas, familias, organizaciones, jurisdicciones), conceptos (temas, lugares, épocas, acontecimientos, obras) por otra, o temas como la necesidad de etiquetas normalizadas, etiquetas de texto libre, conceptos abstractos, objetos del mundo real, denominaciones adecuadas (¿identidades? ¿entidades? ¿términos genéricos?).

En relación a la adecuación de los principios RDF en BIBFRAME se plantea el debate de la funcionalidad en oposición a las reglas definidas en RDF. Se hace preciso distinguir entre tipos de datos y propiedades del objeto; proporcionar URIs, etiquetas o ambos; distinguir tipos por clase; definir propiedades recíprocas, si es necesario; utilizar las propiedades `rdfs:label` y `rdf:value`; evitar la proliferación de propiedades, clases abstractas y propiedades, y el exceso de limitación; separar metadatos administrativos de los datos descriptivos; o utilizar vocabularios externos estables, ampliamente utilizados y normalizados.

Algunos ejemplos de esta controversia entre funcionalidad y reglas podrían ser: se señala que el nombre de una propiedad debe ser un nombre si el valor que se espera es un literal y un verbo si se



trata de un recurso. Sin embargo, en la realidad se permiten frases a partir de un nombre (copyof – copia de), a partir de un verbo (identifiedBy – identificado por), frases preposicionales (continuesInPart – Continúa en parte), etc.

En la actualidad y a corto plazo, la Biblioteca del Congreso se plantea la publicación de una propuesta de vocabulario 2.0 (modelo, principios, denominación, mapeo MARC), una reingeniería de la herramienta, el inicio de un segundo proyecto piloto con el nuevo vocabulario 2.0, continuar con los proyectos LD4P (Linked Data for Production), así como explorar otros medios. En este sentido, se acaba de publicar un nuevo informe sobre metadatos técnicos para recursos audiovisuales: <http://www.loc.gov/bibframe/docs/pdf/bf-avtechstudy-01-04-2016.pdf>.

## 9.2. Una aproximación a BIBFRAME por parte de un distribuidor / Tiziana Possemato, Casalini Libri

Tiziana Possemato, de Casalini Libri, explicó los trabajos que había llevado su empresa en relación a datos enlazados y BIBFRAME. La empresa Casalini Libri, como miembro del [Program for Cooperative Cataloguing](#) (PCC), ha creado más de 20.000 registros bibliográficos en RDA de publicaciones en lenguas romances en formato MARC. Realizan control de autoridades de nombres, títulos y encabezamientos de series y contribuyen al mantenimiento de las bases de datos de autoridades NACO y SACO.

En los últimos tiempos han estado llevando a cabo proyectos sobre datos enlazados y la transición de un entorno MARC a un entorno de datos enlazados, con especial énfasis en relación con BIBFRAME. En este sentido, se han centrado en tres grandes áreas de actividades:

1. **Enriquecimiento de registros MARC para simplificar la conversión a BIBFRAME.** En este sentido, están mejorando un sistema de gestión de URIs en su módulo de catalogación para gestionar los identificadores de cada punto de acceso. Estas URIs se pueden utilizar de diversas formas. Es posible elegir en el proceso de exportación/conversión de datos el número de URIs disponibles para cada encabezamiento, cómo asociarlas al encabezamiento y cómo mostrarlas, utilizando el subcampo \$0 en los registros bibliográficos y el campo 024 en registros de autoridad, como propiedad RDF en las triplas que se producen en el proceso de conversión.
2. **Utilización de una estructura que permita automatizar la conversión de MARC a RDF.** Este proceso de conversión se realiza dentro del propio módulo de catalogación, con el uso de un software integrado en dicho módulo, aunque este mismo proceso también se puede realizar independientemente del sistema de gestión bibliotecaria. La conversión automática a RDF se realiza a través de la [estructura ALIADA](#), aplicando el vocabulario BIBFRAME.

ALIADA (Automatic publication under Linked Data Paradigm of Library Data) ha partido de un Proyecto cofinanciado por la Unión Europea y el programa de financiación de la Innovación para 2007-2013 (FP7), en el que han participado dos empresas tecnológicas (@CULT, SCANBIT), dos museos (ARTIUM de Vitoria y el Museum of Fine Arts Budapest) y un instituto de investigación



(TECNALIA). Las ontologías utilizadas en la estructura ALIADA son FRBRoo, DCMI Metadata Terms, RDF Schema, elementos RDA y BIBFRAME.

3. **Creación de una capa FRBR/BIBFRAME desde los registros bibliográficos y de autoridades.**

Dado que los catálogos en la actualidad describen, fundamentalmente, manifestaciones/instancias, era necesario rediseñar el modelo de datos con un sistema que, a partir de los datos existentes, produjera una capa de Personas, creando un único punto de acceso de nombre para las Personas, y Obras asociadas a cada Persona y enlazadas a los títulos de las Manifestaciones/Instancias. Para ello se han utilizado APIs específicas que han recuperado de ficheros de autoridad externos, como VIAF, estos clusters (agrupaciones) de Personas y Obras.

En este sentido, se ha creado una estructura en tres capas siguiendo BIBFRAME, con una primera capa Persona/Obra en RDF y salvada en un SPARQL endpoint, una segunda de Manifestaciones/Instancias, con datos indexados en un motor de búsqueda capaz de producir nuevas agregaciones de datos en facetas y una tercera de Ítems, a partir de los datos locales de los ejemplares del catálogo de la biblioteca.

En este momento, continúan con la investigación en torno a estas tres áreas de actividades, comentaron y expresaron su interés en la cooperación con otras instituciones y profesionales para compartir experiencias, actividades y resultados.

### 9.3. Visión de conjunto del proyecto LD4P (Linked Data for Production)

#### 9.3.1. De-“Framing” (desmontando a) Derrida: Archivos y anotaciones / Jennifer Baxmeyer, Princeton University

En 2015, la Princeton University compró la biblioteca personal del filósofo francés Jacques Derrida, con unos 16.000 libros y otro tipo de materiales. En estas publicaciones se encuentran muchas anotaciones, inscripciones, marcas, subrayados, símbolos utilizados por el filósofo en su lectura. Jennifer Baxmeyer, comentó que, dentro del proyecto LD4P, intentaron estudiar todo este tipo de anotaciones para tratar de extraer conocimiento relevante respecto a, por ejemplo, personas que le hubieran donado sus obras o cómo era su proceso de lectura.

La participación de la Princeton University en el proyecto LD4P se ha centrado en la evaluación de BIBFRAME en relación con esta colección de Derrida y la identificación de áreas susceptibles de ser expandidas o modificadas, fundamentalmente en relación con la semántica de las anotaciones e inscripciones de manuscritos y las relaciones entre anotaciones e inscripciones originales y transcritas. Como productos finales dentro del proyecto se han producido descripciones de recursos de la colección de Derrida con inscripciones de autores, así como una ontología de extensión para la descripción de recursos de materiales de colecciones especiales.



### 9.3.2. BIBFRAME y colecciones artísticas: Arte y Museos / Melanie Wacker y Amber Billey, Columbia University Libraries

Columbia University cuenta con una colección artística de más de diez mil obras de arte, parte de las cuales se encuentran dispersas por el campus y parte almacenadas en depósitos. Como contribución al [proyecto LD4P](#), eligieron trabajar en la descripción de esta colección utilizando BIBFRAME ya que se presupone que puede trabajar con varios modelos de contenido y acomodarse a diferentes necesidades en relación a la descripción de recursos.

Los integrantes del proyecto han sido tanto bibliotecarios como especialistas en este tipo de colecciones, contaron Melanie Wacker y Amber Billey, y, entre sus objetivos, estaban evaluar la idoneidad del modelo y vocabulario BIBFRAME para la descripción de colecciones artísticas, identificar y documentar necesidades actualmente no cubiertas por BIBFRAME, desarrollar un perfil para la descripción de obras de arte o desarrollar flujos de trabajo que permitan conectar datos enlazados con funciones bibliotecarias fundamentales como las adquisiciones, el inventario o la circulación.

No se utiliza ningún sistema de gestión de colecciones para la descripción de estas piezas, indicaron, tan solo se describen utilizando una hoja de cálculo, pero sí se utilizan vocabularios controlados para la mayor parte de los elementos (como el fichero de autoridades NAF de la Biblioteca del Congreso, el Getty Thesaurus of Geographic Names (TGN), el Art and Architecture Thesaurus (AAT). Entre los datos que se recogen se encuentran información descriptiva (creadores, títulos, fechas, etc.), historia de la obra (notas de exposiciones, procedencias), información sobre la adquisición (fecha, donante, etc.), información sobre condiciones de preservación o información sobre su localización y circulación.

Se realizó una revisión de la literatura existente sobre museos y obras de arte y datos enlazados, desarrollo de esquemas de metadatos no-MARC y modelos para recursos artísticos y visuales, mapeos de las descripciones a MARC o estudios sobre conducta en la búsqueda de información relativa a este tipo de recursos. En este sentido, existen ya vocabularios controlados y tesauros publicados como datos enlazados (los tesauros Getty, como el AAT o el TGN; los ficheros de autoridades de la Biblioteca del Congreso o el vocabulario FAST); ontologías como CIDOC-CRM, VRA-RDF y VRA como perfil BIBFRAME; mapeos (Schema.org, Art Properties data in VRA, Art Properties data in LIDO) y proyectos asociados a estos temas como la iniciativa de datos enlazados abiertos [American Art Collaborative](#).

Se han encontrado con algunos temas a solventar: falta de propiedades en BIBFRAME (por ejemplo, inscripciones y firmas), problemas con la diferenciación entre acontecimientos como materias y como períodos, o el tratamiento de información restringida como valores de seguros, localización exacta, información sobre el donante, etc. En un futuro inmediato se comenzará a experimentar con herramientas como Karma y Vitro, se continuará con el desarrollo de casos de uso y con la evaluación de ontologías existentes para la descripción de obras de arte y de BIBFRAME 2.0, tan pronto sea publicado, y se trabajará en la modelización de los datos, ontologías, desarrollo de perfiles y en la creación/transformación de los datos.



### 9.3.3. Proyectos LD4P en la Universidad de Cornell / Chew Chiat Naun

La representante de la Universidad de Cornell comenzó su intervención haciendo un descargo de responsabilidad: ¡Nada de esto es real!, comentó.

La Universidad de Cornell participa en el proyecto LD4L a través de sus LD4L Labs y del proyecto LD4P y han trabajado en el pasado con VIVO, incluido el trabajo con el editor VITRO en la descripción de colecciones especiales relativas al hip hop. Sus principales áreas de interés en este sentido se centran en la extensión de BIBFRAME, autoridades alternativas y la creación de datos RDF nativos, en dos subproyectos diferentes: los vocabularios RBMS y la colección de LPs del DJ Afrika Bambaataa, pionero del desarrollo de la música hip hop.

En relación al primero de ellos, los [vocabularios RBMS](#), se trata de vocabularios muy consolidados, mantenidos por una comunidad muy activa, todavía no en formato de datos enlazados, aunque durante el último año se ha comenzado a investigar para trasladarlos a un entorno de datos enlazados. Se ha trabajado en este sentido experimentando con Vitro.

El segundo se trata de un proyecto de catalogación original de la colección de LPs del DJ Afrika Bambaataa, que incluye obras raras, flyers, anotaciones, relaciones con otros materiales de colección, permite utilizar fuentes externas, así como la ontología para música interpretada desarrollada por la Universidad de Stanford y la Music Library Association

### 9.4. Actualización de OCLC: BIBFRAME y el programa de datos Enlazados / John Chapmann

John Chapmann, de OCLC, comenzó comentando que, en la actualidad, están trabajando en la modelización de BIBFRAME, en el desarrollo de servicios de producción y en la experimentación en temas de visualización.

1. **Modelización BIBFRAME.** Están trabajando junto a la Biblioteca del Congreso y otras instituciones para finalizar el estándar BIBFRAME y comenzando a explorar cómo trabajar con él a gran escala. En 2015 han publicado, junto a la Biblioteca del Congreso, el informe [Common Ground: Exploring Compatibilities Between the Linked Data Models of the Library of Congress and OCLC](#). Partiendo de este informe se ha procedido a la actualización del análisis técnico de BIBFRAME, se han llevado a cabo trabajos preliminares en relación a la importación y exportación con BIBFRAME y se han evaluado compatibilidades, incluyendo los identificadores de obra de OCLC y los enlaces a otras entidades en las descripciones BIBFRAME, la reconciliación del vocabulario BiblioGraph con BIBFRAME o la modelización de tipos de recursos especializados, como materiales audiovisuales desde la perspectiva de los curadores profesionales.

Frente a la modelización en BIBFRAME 1.0 en Obra (que incluye Obra y Expresión) e Instancia (que incluye Manifestación e Ítem), el modelo OCLC para obras creativas ofrecía una visión



diferente basada en las entidades Obra Creativa (que incluye Obra y Expresión), Modelo de Producto de la Obra Creativa (Manifestación) y Producto Individual de la Obra Creativa (Ítem). Ambos modelos se reconcilian en BIBFRAME 2.0 con la definición de tres entidades:

- BIBFRAME 2.0: Obra (incluye Obra y Expresión), Instancia (Manifestación) e Ítem.
- Modelo OCLC: Obra creativa (incluye Obra y Expresión), Modelo de Producto de la Obra Creativa (Manifestación) y Producto Individual de la Obra Creativa (Ítem)

Se han definido identificadores para los diferentes casos de uso.

2. **Desarrollo de Servicios de Producción.** En septiembre comenzó un proyecto piloto denominado *Person Entity Lookup Pilot*, dedicado a encontrar un modo de identificar y encontrar personas utilizando identificadores comunes. Uno de los principales problemas relativo a esto es que muchos identificadores carecen de menciones sobre relaciones.

Entre los participantes en este proyecto piloto se encuentran la Biblioteka Narodowa de Polonia, Cornell, la Deutsche Nationalbibliothek (DDB), Harvard, Biblioteca del Congreso, Biblioteca Nacional de Medicina (NLM), Stanford, UCDavis, Pepperdine, swissbib y Drexel. Han desarrollado ya varias Apps de demostración: Autocomplete, Autocomplete/Search Hybrid, Search.

3. **Experimentos en relación a la visualización.** OCLC está experimentando con un nuevo sistema de descubrimiento aún en desarrollo denominado "EntityJS" en el que se usan datos enlazados. En él es posible buscar a través de las diferentes entidades.

## 9.5. Actualización de Zepheira: Acelerando la Biblioteca Visible o "No esperen – Repetimos" / Eric Miller

Desde la Conferencia Anual de ALA de junio de 2015 se ha seguido trabajando en el establecimiento de conexiones estratégicas, así como en la implicación y el compromiso de la comunidad, comentó en primer lugar Eric Miller, de la empresa Zepheira.

Como en anteriores ocasiones señaló que las bibliotecas son mucho más que un catálogo y que los datos enlazados son solo tecnología. Y también como en otras ocasiones recurrió a una cita de Chuck Gibson, de la Worthington Public Library, sobre las expectativas de visibilidad de las bibliotecas en la web: "cuando mi comunidad busca en la web algo que tenemos en la biblioteca, deberíamos mostrarnos allí como una opción". Asimismo, volvió a recordar de nuevo que las bibliotecas no están "hablando" el lenguaje que la web entiende. Las bibliotecas cuentan con una gran cantidad de contenido, colecciones especiales y recursos cuya información está encerrada en una tecnología cerrada y en unos vocabularios encasillados.



Para conseguir llegar de la descripción a la visibilidad y al descubrimiento es preciso publicar las descripciones de recursos partiendo de esquemas como VRA, RDA, EAD o MARC21 en vocabularios que la web entienda como BIBFRAME o Schema.org. El proceso pasaría por las fases de transformación de datos (BIBFRAME), enlazado, publicación (Schema.org, BIBFRAME, DC, etc.) y consumo (Google, Bing, Yahoo, Yandex, etc.)

Para llegar a la visibilidad de las colecciones de las Bibliotecas en la web es preciso comenzar con el establecimiento de acuerdos y declaraciones de intenciones. El aprendizaje viene a partir del trabajo conjunto entre diferentes entidades que consiguen obtener un impacto a través de la unión en la acción. Existen ya sistemas locales que son capaces de publicar en la web como SirsiDynix, Innovative Encore, VTLS (Innovative) Virtua, Ex Libris Aleph, Evergreen, CDS/ISIS, OCLC, SirsiDynix Symphony, Baratz Absys, etc.

Cada biblioteca es única y cuenta con materiales únicos. La colaboración es fundamental para llegar a contar la historia completa. Entrelazar lo mejor de las bibliotecas con lo mejor de la red requiere aumentar el poder de la comunidad bibliotecaria para poder llegar a beneficiarse de todo lo que la web permite. Es necesario unirse a la conversación mediante el libhub.org, contar lo que significa la visibilidad para tu biblioteca, escuchar a otros y ver los pasos que han dado para incrementarla.

## **10.MARC ADVISORY COMMITTEE (2)**

En la segunda sesión del MARC Advisory Committee, celebrada el domingo 10 de enero, se aprobaron dos propuestas y se debatieron el resto de documentos que quedaban pendientes.

### **10.1. Propuesta 2016-01: Codificación de posiciones del campo 007 para las reproducciones digitales de grabaciones sonoras en el Formato MARC 21 para registros bibliográficos**

Se trataba de una propuesta presentada por el Comité canadiense sobre Intercambio de Metadatos (CCM) centrada en definir nuevos valores para algunas posiciones del 007 del Formato MARC21 para registros bibliográficos con vistas a obtener un mejor acomodo para las reproducciones digitales de grabaciones sonoras. La propuesta fue aprobada con algunas revisiones menores, tras sugerir el Chair del Comité que la British Library y la Music Library Association llegaran a un acuerdo fuera de la mesa sobre el término “vibraciones” en la definición. Se aceptó la nueva formulación de la definición de “grabación sonora” aportada por la British Library.

### **10.2. Propuesta 2016-02: Defining Subfield \$r and Subfield \$t, and Redefining Subfield \$e in Field 382 of the MARC 21 Bibliographic and Authority Formats**

La Music Library Association presentaba esta propuesta centrada en la definición de dos nuevos subcampos para el campo 382 (Medio de interpretación), uno para describir el número total de



conjuntos y otro para describir el número de instrumentos solistas en dichos conjuntos. La propuesta fue casi unánimemente aprobada, con la salvedad de una única abstención.

### **10.3. DP 2016-DP01: Definición de los subcampos \$3 y \$5 en el campo 382 (Medio de interpretación) del Formato MARC 21 para registros bibliográficos**

El campo 382 del Formato MARC21 para registros bibliográficos, relativo al Medio de Interpretación, no tiene definidos los subcampos \$3 ni \$5. El documento presentado por la Music Library Association planteaba la necesidad de definir ambos subcampos en el campo 382 de cara al tratamiento de las compilaciones de recursos musicales.

Dado que se había planteado en los comentarios enviados que el papel asociado al subcampo \$3 se podía establecer mediante el uso del \$8c siempre que hubiera un punto de acceso o un campo 505, el Chair comentó que hacía ya tiempo que la Music Library Association había rechazado su uso, por lo que esta opción no se consideraba. En el debate se planteó que el \$5 en Estados Unidos se utiliza para proporcionar información sobre la procedencia y la representante de OLAC comentó que ella no estaba en contra del uso que se planteaba en el documento del \$5, pero pidió que no se le denominara “procedencia”. Finalmente, se estableció que el documento volviera al Comité como propuesta para su próxima sesión.

### **10.4. DP 2016-DP02: Clarificación de los valores en la posición 20 del campo 008 (Formato de Música) en el Formato MARC 21 para registros bibliográficos**

La posición 20 del campo 008 ha sufrido varios cambios durante los últimos años para adaptarse a los valores proporcionados por RDA en su regla 7.20.1.3 (Registro del formato de música notada). El documento, presentado por la Music Library Association, planteaba la necesidad de adaptar la terminología empleada en dicha posición a la recogida en RDA.

El representante de la Music Library Association expresó su acuerdo con algunos comentarios enviados (por ejemplo, los de la BNE) relativos a que MARC21 debería permanecer agnóstico a cualquier Código de catalogación, dado que no todo el mundo emplea RDA como tal y para su presentación como propuesta revisarían otras definiciones para tratar de hacerlo. El Chair recordó que no se ofrecía ninguna definición de “partitura para piano” y el representante alemán consideró que sería necesaria. La representante de OLAC planteó que debería incluirse un nuevo código para las “partituras para piano” que podría ser, por ejemplo, \$p. Finalmente, se aceptó que volviera como propuesta a la próxima reunión del Comité teniendo en cuenta todos estos aspectos.



### **10.5. DP 2016-DP03: Números de distribuidores de materiales musicales e imágenes en movimiento en el Formato MARC 21 para registros bibliográficos**

Este documento, presentado conjuntamente por la Music Library Association y OLAC (Online Audiovisual Catalogers), señalaba la necesidad de distinguir los números de música y videograbaciones de los distribuidores de los de los editores en el campo 028 (Número de editor), para lo que sugerían definir un nuevo indicador 6 en dicho campo, así como cambios menores aclaratorios en el campo 037 (Fuente de adquisición), de cara a la introducción de estos datos.

Se plantearon algunas dudas sobre si sería fácil diferenciar los números de los editores de los distribuidores y se señaló que dicha cuestión sería más bien objeto de un documento de buenas prácticas. Se aceptó que fuera presentado de nuevo como propuesta en la siguiente reunión del Comité.

### **10.6. DP 2016-DP04: Extensión del uso del subcampo \$0 (número de control del registro de autoridad o número normalizado) a campos de enlace**

Presentado por la British Library, este documento señalaba la necesidad de extender el uso del subcampo \$0 en los campos de enlace 760, 762, 765, 767, 770, 772, 773, 774, 775, 776, 777, 780, 785, 786, 787 en el Formato MARC 21 para registros bibliográficos.

En general, había bastante acuerdo en la utilidad de incluir URIs para expresar los tipos de relaciones en \$0 en los campos 76X a 78X, a excepción del PCC (Plan for Cooperative Cataloguing) Task Force que señaló su preferencia por utilizar URIs de forma independiente para posibilitar el procesamiento automático. Asimismo, se señaló que el uso del \$0 para los designadores de relación junto a URIs podría ser problemático porque se estaría posibilitando su uso de dos modos totalmente diferentes que las máquinas no pueden entender. El Chair propuso que British Library y el PCC Task Force llegaran a un consenso sobre la propuesta y, en caso de que la British Library siguiera teniendo cuestiones que plantear debería volver al Comité no como propuesta sino como Documento de Discusión de nuevo.

### **10.7. DP 2016-DP05: Expansión de la definición del subcampo \$w (Número de control del registro bibliográfico) a los números normalizados**

También presentado por la British Library, este documento proponía la extensión de la definición del subcampo \$w (número de control del registro bibliográfico) a los números normalizados tanto en el Formato MARC21 para registros bibliográficos como en el de Autoridades. Tras una discusión sobre las consecuencias, por ejemplo, al fundir registros en OCLC o la posible confusión entre contenido primario y secundario, se planteó que se presentara como propuesta en la siguiente reunión del Comité con la ayuda del PCC (Plan for Cooperative Cataloguing) para la reelaboración del texto explicativo referido al subcampo \$w.



### **10.8. DP 2016-DP06: Definición de los subcampos \$2 y \$0 en el campo 753 del Formato MARC21 para registros bibliográficos**

Este documento, presentado por OLAC (Online Audiovisual Catalogers) y GAMECIP (Game Metadata and Citation Project), proponía añadir los subcampos \$2 (Fuente del término) y \$0 (Número de control del registro de autoridad o número normalizado) al campo 753 (Información técnica sobre acceso a archivos de ordenador). Dado que existía consenso sobre la necesidad de estos cambios y no había ninguna objeción a ellos, la representante de OLAC preguntó si no podría aprobarse ya directamente como propuesta, lo que se valoró de forma positiva, por lo que el documento fue aprobado ya como propuesta sin tener que esperar a la siguiente reunión del Comité.

### **10.9. DP 2016-DP11: Puntuación en el Formato MARC21 de Autoridades**

La Biblioteca Nacional de Alemania presentó este documento que expresaba la necesidad de codificación para indicar la ausencia de puntuación redundante en la codificación de campos y subcampos a través de la posición 18 de la Cabecera.

La British Library expresó su preocupación respecto a esta posibilidad en el Formato de Autoridades. El representante del Committee on Cataloging: Description and Access (CC:DA) también expresó su preocupación en relación a esta solución dada la debilidad de dicha posición ya en el Formato para registros bibliográficos. Para la British Library la única posibilidad sería considerar dos posiciones en vez de una (incluido/no incluido), para indicar si es acorde con la puntuación RDA. Desde la Library of Congress se opinó que sería mejor no hacer referencia a ningún código de catalogación específico o prácticas nacionales. Se sugirió que volviera como propuesta pero teniendo en cuenta todas estas consideraciones.

### **10.10. DP 2016-DP15: Tipo de Medio y Tipo de Soporte en el Formato MARC21 Autoridades**

Presentado también por la Biblioteca Nacional de Alemania, este documento proponía un modo de codificar el tipo de medio y el tipo de soporte establecido por RDA en el Formato MARC de Autoridades.

La discusión se centró en la necesidad de introducir información que pertenece a la entidad Manifestación en el contexto de un registro de autoridad, ya que dicha información debería quedar confinada al registro bibliográfico y la información sobre el tipo de soporte también en la información sobre el holding. Desde la Biblioteca Nacional de Alemania se argumentaba que existen ciertos recursos que tienen un carácter único (“singletons”), para los que sería fundamental que dicha información constara en el registro de autoridad, lo que no afecta a aquellas manifestaciones que implican varios ítems. Una determinada agencia bibliográfica podría necesitar utilizar un registro de autoridad para un tipo de manifestación de carácter único con el fin de poder ser utilizada como materia en otro registro bibliográfico, por ejemplo.



No se llegó a ningún consenso sobre el documento, por lo que se solicitó que se volviera a presentar como un Documento para debate y no como propuesta en la siguiente reunión del Comité.

### **10.11. DP 2016-DP16: Extensión del nivel de codificación en el Formato MARC21 de Autoridades**

Este documento, también propuesta de la Biblioteca Nacional de Alemania, planteaba la necesidad de extender la posición 17 de la Cabecera con siete nuevos valores para expresar el grado de exhaustividad del registro, así como los derechos y restricciones aplicables. Se reconocía la necesidad alemana, pero se entendía que era información local de la propia agencia bibliográfica, por lo que el Chair sugirió la posibilidad de utilizar un campo local para introducir este tipo de información. También sugirió que se proporcionaran definiciones de cada uno de los siete niveles propuestos para ver si podrían ser utilizados de forma generalizada.

Desde la British Library se sugirió como posible solución definir una posición en la cabecera que indicara una práctica local en el campo 042, en lo que estaba de acuerdo una representante de la Biblioteca del Congreso. Se sugirió que se estudiara esta posibilidad y que se volviera a presentar como Documento a Debate para la próxima reunión del Comité, quizás con la definición de un nuevo subcampo para el contenido, el desarrollo de indicadores y el posible uso del \$2 añadiendo una nueva lista a la [MARC Authentication Action Code List](#).

## **11. GRUPO DE INTERÉS DE DIRECTORES DE DEPARTAMENTOS DE CATALOGACIÓN ALCTS CaMMS**

### **11.1. Presentación**

En la presentación del acto, Daniel Lovins and Jackie Shieh, co-Chairs del grupo, explicaron que el objetivo del acto era plantear el mejor modo de avanzar y elegir el mejor camino para la evolución del trabajo con datos de autoridades. La Información que se proporciona debe ser precisa y rigurosa, señalaron. Durante la última década diferentes proyectos han conseguido implementar información relativa a autoridades en la web.

Jackie Shieh, Coordinadora de Descripción de Recursos en la George Washington University, hizo un llamamiento: debemos movernos desde el control de autoridades a la gestión de identidades y para ello necesitamos pensar en cuestiones como la formación, las infraestructuras necesarias, etc.

### **11.2. Autoridades, Entidades, Objetos de la vida real y... ¿Gatos?: De la creación de autoridades a la gestión de identidades / Nancy Lorimer**

La charla se centró en temas como la importancia de las URIs en los datos de autoridad, el enlazado de URIs con datos en el proceso de transformación de MARC a BIBFRAME, así como en las diferencias entre conceptos como gestión de identidad, basado en objetos de la vida real (RWOs: Real world objects) y los datos tradicionales de autoridad.



Hasta ahora algunos proyectos que se han desarrollado nos indican cómo se mostrarían los datos en un entorno de datos enlazados, recordó, pero es necesario que avancemos para poder trabajar completamente en un entorno de características semejantes. Uno de los principales problemas para ello es el mantenimiento de identificadores para las autoridades, aunque Nancy Lorimer, de la Universidad de Stanford, avisó que no iba a ser capaz de ofrecer una respuesta final a estas cuestiones.

Comenzó hablando sobre cómo, de acuerdo con la [definición de Wikipedia](#), una entidad puede ser prácticamente cualquier cosa. Sin embargo, las entidades definidas en BIBFRAME son muy precisas, a saber: Works (Obras), Instances (Instancias), Items (Ejemplares); Agents (Agentes), Corporate bodies (Entidades corporativas), Geographic locations (Localizaciones geográficas) (MADS or Geonames), Concepts (Conceptos), Events (Acontecimientos).

En relación a los identificadores, Nancy Lorimer comenzó hablando de que nombrar a un gato es algo complicado, haciendo alusión al poema de T. S. Eliot *The Naming of Cats*, y, muchas veces continuó, llamamos a nuestro gato de diferentes maneras según el momento. Quizás, dijo, consideréis que estoy tan loca como el Sombreroero, pero un gato debe tener al menos tres nombres diferentes, como, por ejemplo: Kit, Kit-kit, Shu-shu. Una gran carcajada estalló en la sala.

Si tenemos tres nombres diferentes debemos escoger uno como el nombre principal (1XX en MARC) en el registro de autoridad y hacer referencias cruzadas (4XX) a los otros dos. Modelado en RDF y en BIBFRAME 1.0 tendríamos la entidad bf:Agent, subclase de la subclase de bf:Authority, o sea, en otras palabras, bf:Agent es un nombre.

En BIBFRAME 2.0 tendremos, dijo, la entidad bf:Person, subclase de bf:Agent, pero no hay un bf:Authority, ya que se refiere a un objeto del mundo real (RWO). Sin embargo, nuestro gato, no es una persona.

```
--<bf:Agent rd:about=http://id.loc.gov/rwo/agents/nr15000000>
  <rdfs:label>"Shu Shu (Cat)"</rdfs:label>
  <rdfs:label--"Kit (Cat)" </rdfs:label>
  <madsrdf:isIdentifiedByAuthority
rd:resource=http://id.loc.gov/authorities/names/nr15000000>/
--<bf:Agent >
```

La Biblioteca del Congreso ha comenzado a crear de forma discreta y callada identificadores para estas personas. Pero, ¿cómo conseguimos estos identificadores y cómo los integramos en los flujos de trabajo relacionados con los metadatos? Tradicionalmente esto se hacía mediante el uso de fuentes autorizadas y esta consulta de fuentes autorizadas ya está incorporada en nuestros flujos de trabajo tradicionales. El flujo de trabajo de las Bibliotecas de la Universidad de Stanford relativo a la catalogación de monografías permite incorporar al catálogo multitud de entidades autorizadas, pero el personal no tiene tiempo para revisarlas todas y uno de los desafíos es cómo trabajar con estos identificadores.



En las bibliotecas de Stanford utilizan una única herramienta (SearchWorks) para la búsqueda de información en múltiples repositorios y bases de datos de la institución. En unas, como Symphony, se utiliza MARC21 como esquema de metadatos y en otras como el Repositorio Digital, MODS. La mayor parte de ellos producen este tipo de identificadores.

En el proceso de trabajo con los datos heredados (Legacy data) para la mejora de los registros bibliográficos y de autoridades es preciso añadir también URIs en los registros MARC. Además, es bastante probable que la mayoría de las bibliotecas nunca cambien a un entorno de datos enlazados. Entre los beneficios de añadir URIs a los registros MARC antiguos se pueden citar fundamentalmente dos: un identificador se relaciona directamente con el término en el registro bibliográfico y se puede elegir fácilmente encabezamientos que no cuenten con identificadores para intervenir en ellos. En cuanto a los inconvenientes, señaló, este proceso consume mucho tiempo, si los enlaces a los identificadores se rompen no lo puedes saber, si necesitas recargarlos tienes que pasar por todos y cada uno de los registros bibliográficos y, además, tienes que tener mucho cuidado con los servicios de actualización de los registros bibliográficos.

Por otra parte, también se pueden añadir URIs a los registros de autoridad, que es algo que las Bibliotecas de la Universidad de Stanford han estado haciendo recientemente. Entre las ventajas de hacerlo, citó que el identificador se graba una única vez, hay un uso más eficiente del tiempo que consume la máquina, es fácilmente actualizable, utiliza los registros de autoridad de la forma en que fueron pensados y, además, se pueden incluir todos los identificadores que se quiera. Entre los inconvenientes está que solo funciona con vocabularios para los que tengas registros de autoridad y, en ocasiones, el registro de autoridad no siempre se corresponde exactamente con la cadena de caracteres completa en el registro bibliográfico, por lo que puede necesitarse intervención.

Nuevas posibilidades se atisban, dijo: identificadores nuevos sin registros de autoridad obteniendo identificadores locales e internacionales a través del proceso de catalogación por copia o en la catalogación original:

BIBFRAME → Crear el registro NAF en MARC → Añadir el identificador a BIBFRAME → Authority Vendor

BIBFRAME → Crear el identificador ORCID/ISNI → Añadir el identificador a BIBFRAME → Authority Vendor

BIBFRAME → Crear el identificador local → Añadir el identificador a BIBFRAME → Authority Vendor

Cuanto antes se añada el identificador en el proceso de la catalogación mejor será el resultado.

Tus URIs podrán hacerse amigas de otras URIs, explicó. Cualquiera puede crear URIs para cualquier cosa ya que son muy fáciles de crear, aunque no resulta tan fácil enlazarlas. Solo para agentes ya tenemos Id.loc.gov, VIAF, ISNI, ORCID... Y, además, existen otros como LCSH (Library of Congress Subject Headings), AAT (Art and Architecture Thesaurus), TGM (Library of Congress Thesaurus for Graphic Materials), MeSH (Medical Subject Headings), RDA (Resource Description & Access), Wikidata, GAMECIP (Game Metadata and Citation Project), etc.



ISNI (International Standard Name Identifier) y VIAF (Virtual International Authority File) resultan de gran ayuda, pero incluso los “peces gordos” pueden tener problemas. Y pone como ejemplo a su propio padre, químico, para el que existen hasta 3 encabezamientos en VIAF. Pero, a pesar de los problemas, señala, es mejor tener identificadores que no tenerlos, pero deben ser identificadores verdaderamente locales.

A modo de conclusión señaló que existen cuestiones que deben resolverse en relación a la infraestructura (¿dónde colocar los identificadores?, ¿cómo actualizarlos?, ¿cómo resolverlos?), a los flujos de trabajo (¿cuándo añadir identificadores?) y a la formación (¿quién puede hacer ese trabajo?).

### **11.3. Contexto Archivístico Codificado: Entidades corporativas, Personas y Familias (EAC-CPF) / Katherine M. Wisser**

Las raíces de los fundamentos de la descripción archivística se pueden rastrear en la obra *Manual for the Arrangement and Description of Archives*, de los tres holandeses S. Muller, J. A. Feith y R. Fruin, conocido como el Manual holandés, que sentó las bases teóricas de los modernos métodos archivísticos allá por 1898, el año de su publicación. En él se establecían una serie de principios, como el principio 8 que señalaba que “varias colecciones archivísticas colocadas en un depósito deben mantenerse cuidadosamente separadas”. El principio de procedencia y respeto de los fondos, es decir, su conexión a la entidad administrativa u oficial a la que los documentos pertenecen debe ser respetado, en cualquier caso, de acuerdo a dicho principio, establecía.

En este sentido, la razón de la creación del estándar [EAC-CPF](#) (Encoded Archival Context – Corporate bodies, Persons, and Families) fue proporcionar una estructura de comunicación para la información contextual relativa a personas, familias y entidades corporativas que permitiera el intercambio de información de registros de autoridad elaborados de acuerdo a la norma ISAAR (CPF) y a la información registrada en el área 3 de Contexto de los registros descriptivos elaborados de acuerdo a la norma ISAD(G), fundamentalmente 3.2.1: Nombre del creador y 3.2.2: Historia administrativa/biográfica.

Entre los principios formulados en [DACS](#) (Describing Archives: a Content Standard) como base para las reglas de esta norma, se incluyen dos fundamentales que establecen igualmente el principio de procedencia como pilar de la descripción. El principio 2 señala que el respeto de los fondos es la base de la descripción archivística, mientras que el principio 8 señala que los creadores de materiales archivísticos, así como los propios materiales, deben establecerse. En [ISAAR \(CPF\)](#) los elementos 5.1 (Área de identificación), 5.2 (Área de descripción), 5.3 (Área de relaciones), 5.4 (Área de control) y 6 (Relación de instituciones, personas y familias con documentos de archivo y otros recursos) recogen la información relacionada con ambos principios.

El estándar EAC/CPF comenzó a gestarse en reuniones mantenidas en Yale y Toronto allá por los años 1998 y 2001. En 2004 ya se contaba con una versión beta y en 2006 se constituyó un grupo de trabajo sobre EAC, del que formaban parte 14 miembros procedentes de nueve países diferentes (Francia, Alemania, Estados Unidos, Canadá, Australia, Gran Bretaña, entre otros), que se reunió en



Bolonia en mayo de 2008. En agosto de 2009 se publicó para revisión un borrador que fue definitivamente publicado en 2010 y adoptado por la Society of American Archivists en enero de 2011.

El concepto de identidad es fundamental en EAC-CPF, pero se pueden definir varios tipos. El más común es la identidad única, por ejemplo, “Milne, A. A. (Alan Alexander), 1822-1956”. Dentro del tipo de múltiples identidades, tenemos dos tipos, muchas identidades en una y una en muchas, como, por ejemplo, Bill Clinton, del que podemos tener “Clinton, Bill, 1946-”, “Arkansas. Governor (1979-1981: Clinton)”, “Arkansas. Governor (1983-1992): Clinton” o “United States. President (1993-2001: Clinton)”. Por último, podemos tener identidades colaborativas como Franklin W. Dixon, pseudónimo colectivo utilizado por varios autores de la serie policíaca para adolescentes *The Hardy Boys* como Leslie McFarlane, autor canadiense, (McFarlane, Leslie, 1902-1977), Stratemeyer, Edward, 1862-1939 o Robbins, David, 1950-.

La estructura básica de la norma se basa en un primer elemento “Control” que recoge información sobre la identidad, creación, mantenimiento, estado, reglas y autoridades utilizadas para crear el registro. Contiene otros subelementos básicos de entre los que recordId (identificador del registro), maintenanceStatus (Estado de mantenimiento), maintenanceAgency (Agencia que lo mantiene) y maintenanceHistory (Historial de mantenimiento) son obligatorios. Además, existen otros siete elementos opcionales. Por su parte, el elemento descripción (cpfDescription) se subdivide en varios subelementos como identidad (nombre o nombres relacionados con la entidad descrita), descripción (elementos de contenido que comprenden la descripción de la entidad CPF que se describe) y relaciones (con otras entidades, recursos o funciones). Cada elemento de la norma ISAAR (CPF) tiene su correspondencia en EAC-CPF. Por ejemplo, el elemento 5.1.2 (Forma autorizada del nombre) de ISAAR (CPF) se corresponde con los elementos <nameEntry> y <authorizedForm>.

Como ejemplos de proyectos que han utilizado EAC-CPF se pueden citar:

- *Connecting the DOTs: Using EAC-CPF to reunite Samuel Johnson and his circle*, proyecto desarrollado por las Universidades de Yale y Harvard en colaboración: <http://osc.hul.harvard.edu/liblab/proj/connecting-dots-using-eac-cpf-reunite-samuel-johnson-and-his-circle>.
- *People Australia*: <http://trove.nla.gov.au/people>.
- *APEX (Archives Portal Europe)*: <http://www.archivesportaleurope.net>.
- *SNAAC (Social networks and archival context)*: [http://socialarchive.iath.virginia.edu/NAAC\\_index.html](http://socialarchive.iath.virginia.edu/NAAC_index.html).

Este último, [SNAC](#) (Social networks & archival context), es un proyecto cooperativo internacional lanzado en octubre de 2015, que tiene como objetivo cambiar el modo en que la historia es imaginada y escrita, según el historiador Ed Ayers. El proyecto financiado por el National Endowment for the Humanities, la Andrew W. Mellon Foundation y el Institute of Museum and Library Services, cuenta con un presupuesto global de dos millones y medio de dólares. Entre los participantes en el proyecto piloto están las universidades de Harvard, Yale, Tufts, Virginia, George Washington, la Library of Congress, los National Archives, la New York Public Library o el Smithsonian, entre otros.



## 12. ACTUALIZACIÓN SOBRE LOS PROYECTOS DE INVESTIGACIÓN DE OCLC

### 12.1. Investigación y membresía en OCLC / Eric Childress

Eric Childress, gestor de proyectos en OCLC Research, comentó los trabajos que están llevando a cabo y que, fundamentalmente, se centran en explorar los desafíos a los que se enfrentan las bibliotecas y los archivos en un entorno tecnológico rápidamente cambiante. Señaló que son tres los núcleos primarios de actividad: investigación y desarrollo de la comunidad, desarrollo avanzado y participación de socios y miembros.

Los temas de trabajo se centran principalmente en torno a las colecciones de investigación, al conocimiento del Sistema bibliotecario de forma amplia, ciencia de datos, estudios de usuarios o aprendizaje progresivo. Entre los últimos informes OCLC publicados se pueden citar: [The Library in the Life of the User: Engaging with People Where They Live and Learn](#), [Shaping the Library to the Life of the User: Adapting, Empowering, Partnering, Engaging](#), [The Archival Advantage: Integrating Archival Expertise into Management of Born-digital Library Materials](#) o [If You Build It, Will They Fund? Making Research Data Management Sustainable](#).

Entre los próximos webinars que van a impartir citó: Works in progress webinar: *I am the content* (12 de enero) o *A Context-driven Approach to Data Curation for Reuse Workshop* (22 de febrero). Se pueden consultar en OCLC WebJunction: [www.webjunction.org/events/webjunction.html](http://www.webjunction.org/events/webjunction.html).

### 12.2. La biblioteca en la vida del usuario: involucrarse con la gente allí donde viven y aprenden / Lynn Silipigni

Nuestro modelo de pensamiento tradicional era aquel en el que veíamos al usuario desde el punto de vista de la vida de la biblioteca, pero ese modelo está cambiando ampliamente y ahora tenemos que pensar en la biblioteca desde el punto de vista de la vida del usuario, comenzó señalando Lynn Silipigni. Las bibliotecas ya no son el único campo de juego en la ciudad, así que tenemos que conseguir convertirnos en una opción que interese a la gente, dijo.

Sabemos que contamos con muchos datos, los recolectamos, pero ahora necesitamos utilizarlos para crear escenarios diferentes, continuó. El contexto es muy importante para poder interpretar el comportamiento de los usuarios, cuya motivación y conducta cambian en función de él. Por ejemplo, los servicios de referencia virtuales resultan imprescindibles en la actualidad, más si cabe en aquellas ocasiones en que la biblioteca no está abierta. En los servicios de referencia virtuales, si el individuo siente que el bibliotecario se involucra en la conversación y la relaja, la comunicación tendrá mucho más éxito.

Necesitamos proporcionar servicios a nuestros usuarios que realmente sirvan a los intereses de la gente, para lo que realmente hacen, no para lo que dicen que hacen. Para ello necesitamos hacer preguntas, pero es fundamental que queramos de verdad conocer las respuestas a las preguntas que hacemos y también que queramos hacer algo al respecto, aportar una solución a dichas respuestas. Lynn dijo que es preciso que recolectemos datos y los utilicemos de manera inteligente. Asimismo es



necesario plantear las preguntas de diferentes formas. Como señalaba en el reciente informe publicado por OCLC, [The Library in the Life of the User: Engaging with People Where They Live and Learn](#) (2015), los bibliotecarios tienen la oportunidad de formar parte de las redes sociales de sus usuarios y de ubicar sus recursos en el contexto de sus necesidades de información.

Necesitamos pensar en cómo mostramos la información en nuestras páginas web, ya que la gente no tiene paciencia para estar buscando en diferentes silos de información. Además, debemos contar con que no existe una única forma de buscar información, ya que las personas tienen diferentes modos de aproximarse a la información en función de su propia idiosincrasia y necesidades. También debemos pensar que nuestros usuarios realizan múltiples tareas a la vez de forma constante.

Prosiguió explicando la importancia del espacio, pero no solo del espacio físico, sino también del virtual. Debemos contar con espacios físicos y virtuales que permitan socializar y compartir. Una de las cosas que decía Ranganathan en los años treinta del pasado siglo es que es preciso seguir al usuario desde el mismo momento en que entra en la biblioteca hasta que se va de ella. Esto continúa estando vigente y es lo que debemos hacer, pero no solo en la biblioteca física también en la biblioteca digital. En el informe de OCLC [Reordering Ranganathan](#), extrajimos algunas conclusiones, dijo, de nuestros análisis en este sentido como que es preciso que conozcamos a nuestras comunidades y sus necesidades, es necesario que desarrollemos relaciones tanto a nivel presencial como online y que incrustemos dichos servicios en los flujos de trabajo con los que ya contamos. Cuando la gente habla de las bibliotecas habla de libros, no conoce los servicios y recursos virtuales que las bibliotecas ofrecen por lo que es fundamental que publicitemos nuestros servicios para que los usuarios los conozcan.

Pero no nos debemos quedar ahí, concluyó. Tenemos que evaluar nuestros servicios digitales desde el punto de vista de visitantes y residentes. No es necesario que reinventemos la rueda, hay muchísima información a nuestro alrededor, como métodos de recogida de información, ejemplos de entrevistas, actividades para comprender el comportamiento de nuestros usuarios, y es mejor adaptarla a nuestras bibliotecas que partir de cero. No hay que intentar hacerlo todo, sino ir poco a poco. Y ofreció algunas recomendaciones: aprende sobre lo que no sabes, involúcrate en nuevos modos de acción y utiliza aquello que sí conoces.

### **12.3. Transformar las descripciones bibliográficas en conocimiento procesable / Jeff Mixer**

Jeff Mixer, ingeniero de OCLC, comenzó hablando de una serie de investigaciones recientes de Google sobre evaluación de la confianza, que describen el uso de modelos probabilísticos y aprendizaje automático para evaluar si son ciertas y verdaderas las declaraciones hechas por múltiples fuentes. En 2012 Google publicó un [post en su blog](#) que describía el Gráfico del conocimiento, que mejoraba la búsqueda de cosas, personas, lugares y sugería relaciones relevantes con otras cosas. No contentos con dicho gráfico, ahora trabajan en su sucesor, al que denominan la Cúpula del conocimiento, base de conocimiento desarrollada por Google, que incluye el valor de la confianza derivado de sus investigaciones sobre modelos probabilísticos y aprendizaje automático. La introducción de este valor de confianza permitiría distinguir entre declaraciones de conocimiento



con una alta probabilidad de ser ciertas de otras que son menos probables que sean ciertas. Los tres componentes principales de la Cúpula del conocimiento de Google son *extractores*, que extraen las tripletas a partir de un gran número de fuentes web, *graphs-based priors* que establecen la probabilidad de certeza de cada tripleta y *fusión del conocimiento*.

¿Sería posible construir una Cúpula de conocimiento para las bibliotecas? OCLC ha estado trabajando en la adaptación de este modelo de conocimiento de Google al ámbito bibliotecario y está evaluando un modelo bibliográfico similar al de Google para fuentes de datos bibliográficos y de autoridad, en combinación con contenido proporcionado por usuarios y datos enlazados de otros proveedores, para evaluar y crear una “cúpula de conocimiento” para declaraciones sobre entidades y sus relaciones, que incluya a personas, grupos, lugares, acontecimientos, conceptos y obras. Esta cúpula de conocimiento sería utilizada en aplicaciones y servicios de nueva generación.

Las fuentes con las que cuentan son Worldcat, VIAF (Virtual International Authority File) y FAST (Faceted Application of Subject Terminology), para las que OCLC produce identificadores persistentes y datos enlazados RDF. En el flujo de trabajo de la cúpula de conocimiento, los extractores extraen los datos de dichas fuentes, que pueden incluso ser datos no estructurados, y los mapean a un modelo de datos normalizados para el proceso de fusión. Utilizan Schema.org, pero podrían utilizarse otros. En el proceso de extracción se crean las tripletas de conocimiento. Los extractores pueden ser generales o muy específicos, por ejemplo, extractores de nombres o de tipos de entidades. El resultado final son tripletas “scored” tras el proceso de fusión, es decir, una única declaración con algún tipo de calificación asociada.

Una tripleta RDF es una declaración que relaciona una cosa con otra especificando un sujeto, un objeto y un predicado y utilizando URIs para cada uno de esos tres elementos. El proceso de creación de tripletas de conocimiento a partir de datos extraídos de registros sería el siguiente: Registros MARC → Registros MARC procedentes de Worldcat mejorados (clusters FRBR, con identificadores normalizados y enlaces a vocabularios controlados) → Entidades RDF (personas, organizaciones, lugares, conceptos, acontecimientos, obras) → Tripletas. Las tripletas en esta nueva cúpula del conocimiento bibliotecario proporcionan oportunidades a las aplicaciones para mejorar y apoyar las tareas de edición, visualización, descubrimiento, etc.

OCLC Research está investigando cómo ensamblar y trabajar con todo este tipo de datos en un sistema de descubrimiento experimental, al que denominamos EntityJS. Este proyecto pretende extraer experiencia real utilizando datos enlazados, analizar los procesos de edición y refinamiento de entidades y devolver las tripletas a la cúpula de conocimiento. Quieren poder buscar a través de las entidades que crean y extraen de los registros, mostrar las relaciones entre ellas y pensar en algunas cuestiones que puedan servir al usuario para llegar a otras relaciones adicionales.

Han estado experimentando con un subconjunto de conocimiento: los registros MARC del “[ArchiveGrid](#)”, extraídos de Worldcat. Han creado una aplicación que interactúa con las tripletas. El método de programación se denomina MapReduce. En esta aplicación han incorporado la implicación del usuario con la definición de relaciones del tipo “same as”. La experimentación



continúa con la construcción de un modo de asignar niveles de seguridad a los datos de EntityJS, como entrada para el proceso de fusión.

## 12.4. Trabajar con socios e identificadores de OCLC Research / Karen Smith-Yoshimura

Para conseguir impacto es preciso trabajar con iniciativas financiadas a nivel gubernamental para asegurar su valor, señaló Karen Smith-Yoshimura, de OCLC Research. Un indicador de impacto son los rankings de universidades, que utilizan las citas para evaluar su posición en dichas clasificaciones.

Por su parte, los identificadores son URIs únicas, persistentes y públicas asociadas a un objeto digital, que se pueden resolver a nivel global en las redes, inequívocas a la hora de encontrar, identificar y utilizar un recurso. Como ejemplos se pueden citar LC (Library of Congress), VIAF (Virtual International Authority File), ISNI (International Standard Name Identifier) o Wikidata. Estos identificadores son el pegamento necesario para instituciones y sistemas.

Los investigadores interactúan con muchos sistemas tanto internos como externos. Las instituciones quieren rastrear todos sus resultados de investigación, por lo que la afiliación es uno de los temas principales, los grupos de investigación en los que está involucrado su personal en múltiples instituciones, los informes de evaluación nacionales, su financiación, así como validar la afiliación o desambiguar los nombres de sus investigadores, etc.

La OCLC Research Task Force sobre la representación de organizaciones en ISNI (OCLC Research Partners Task Force on Representing Organizations in ISNI) está encargada de documentar cómo deberían representarse las organizaciones en la base de datos de ISNI. Su objetivo es encontrar modos de mejorar la calidad de los registros ISNI, su codificación, exhaustividad, la interfaz del usuario, la difusión, así como de involucrar a la comunidad. Su principal interés se centra en las instituciones académicas y en las organizaciones con las que interactúan. Para ilustrar casos de uso de han elegido ejemplos de relaciones entre grupos de investigación.

Existen desafíos en las organizaciones ya que son especialmente complejas: se fusionan, se dividen, compran a otras, son compradas y, a menudo, no resulta muy claro que en el caso de que una institución cambie su nombre, eso implique que se haya convertido en una nueva organización. ISNI parte de múltiples fuentes. Unas 500.000 organizaciones tienen ISNIs públicos. ORCID, por ejemplo, utiliza ISNIs para las organizaciones.

La actividad del grupo ha producido recomendaciones para ISNI como aumentar los tipos de relaciones y cómo se muestran, publicar la ontología ISNI, añadir Turtle, N-Triple, JSON-LD como opciones de datos enlazados o crear formularios para las organizaciones como usuarios finales.



### 13.RDA TECH FORUM

El RDA Tech Forum se planteaba como una mesa de discusión de aspectos técnicos relacionados con RDA y con el RDA Toolkit, al que se invitaba a participar tanto a usuarios del Toolkit como a empresas especializadas en gestión de metadatos, software para bibliotecas, etc. El objetivo era poner en común con el personal de RDA Toolkit y también con el equipo de desarrollo de RDA posibles mejoras en el RDA Toolkit o temas relacionados con el RDA Registry, como paso previo a la revisión y rediseño de la herramienta que se va a llevar a cabo. James Hennelly, Gordon Dunsire, Diane Hillmann, John Pippy y Deborah Fritz atendieron las preguntas y sugerencias de los asistentes.

Se preguntó si se iba a volver a incluir el índice en el Toolkit. James Hennelly contestó que se trata de un recurso muy pesado, por lo que se había desestimado, pero que si había personas interesadas en mantenerlo podrían reconsiderarlo y renovarlo. La mayor dificultad estriba en la necesidad de gestionar la información en diferentes idiomas. Lo que sí tienen planteado es incluir el índice impreso.

También se plantearon dificultades en el uso del RDA Toolkit. Se comentó que en las próximas semanas se iba a incluir una nueva herramienta para facilitar la búsqueda en el Toolkit. En relación a esto, se planteó si se había considerado alguna solución de gestión documental específica para facilitar el acceso a la información en el RDA Toolkit. Se hizo hincapié en que su estructura se está haciendo cada vez más complicada fundamentalmente cada vez que se añade una nueva lengua.

Se expuso que sería necesario explicitar los cambios, así como los elementos que hubieran sido eliminados o reubicados. En relación al rediseño, también se habló de que los cambios en FRBR iban a afectar a RDA y sería necesario registrarlos. Ahora, RDA está dividida en capítulos y probablemente se verán forzados a reestructurar los elementos y a dividirlos en piezas más pequeñas de información. La visualización va a ser revisada y se pretende que el diseño sea adaptable. Necesitan manejar correctamente la navegación entre traducciones ya que existen algunas inconsistencias. Con la actualización de febrero ya se empezarán a ver algunos cambios. También están comenzando a realizar refinamientos en vocabularios locales y perfiles de aplicación. El objetivo es facilitar el uso.

Por último, se planteó el tema de la formación en el uso de la herramienta RDA Toolkit y se contestó explicando que existen webinars, presentaciones, etc., por ejemplo, una reciente de James Hennelly titulada *Look into Essentials*. Por otra parte, se comentó que la Biblioteca del Congreso cuenta con muchos materiales muy detallados con ejercicios y que el programa RIMMF es gratuito y permite generar descripciones RDA fuera del entorno MARC, con una intencionalidad puramente formativa.