Chair, Matthew Wise (NYU) welcomed everyone to the meeting of the MARC Advisory Committee (MAC). Introductions were made and the 2014 Annual Meeting MAC minutes were approved.

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

Presented by the British Library. This paper, a follow-up from Discussion Paper No. 2014-DP06, proposes to define indicator 1 in Bibliographic field 037 (Source of acquisition) in order to record sequencing information. Additionally, subfields $3$ and $5$ are defined so that materials and institution or organization to which a source of acquisition applies can be recorded. The proposal was accepted as written, with one abstention.

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

Presented by the Canadian Committee on MARC (CCM). This discussion paper presents options for recording the RDA data element Format of Notated Music in the Bibliographic format, with the possibility for recording in the Authority format as well. Options explored include defining a field in the 33X block, 348, or 5XX block.

- The Format of Notated Music is a significant data element for finding, identifying, and selecting resources consisting of notated music.
- Precise encoding of the Format of Notated Music is useful in enabling retrieval as it is often a determining factor in selecting notated music. For instance, performers of instrumental chamber music will require resources that include parts, whereas someone who is coaching such a group may find a study score or full score of more value.
- CC:DA is looking at extent of manifestation vs. extent of expression so this proposal may be in line with that. Although the Music Library Association could not come up with a specific use case, they felt it would be better to leave the door open for encoding in the Authority format.
- 348 was preferred due to the strong parallel the 34x fields for audio recordings have with Format of Notated Music. The German National Library’s approach makes use of field 655 (Genre/Form) by attempting to link to the authority record to control the listings. However, it was suggested that CCM move forward with defining 348 $a$/b in parallel to 336-338 and add subfield $0$, while at the same time explaining the potential use of 655 in a different context.

**Proposal No. 2015-04: Broaden Usage of Field 088 in the MARC 21 Bibliographic Format**

Presented by Matthew Wise on behalf of Alaska Resources Library and Information Services (ARLIS). This paper, a follow-up to Discussion Paper No. 2014-DP07, proposes the broadening of MARC Bibliographic field 088 (Report number) to allow for the coding of series numbers...
commonly found in technical reports and government publications. The proposal was unanimously accepted with corrections to MARC coding in examples.

**Proposal No. 2015-02: Adding Dates for Corporate Bodies in Field 046 in the MARC 21 Authority Format**
Presented by the British Library. This paper proposes to accommodate dates associated with corporate bodies in field 046 (Special Coded Dates) by defining a non-repeatable subfield $q (Establishment date), defining a non-repeatable subfield $r (Termination date), expanding the 046 field definition and scope to specify that the period of activity is a type of date which can be associated with corporate bodies, and amending the definitions for 046 subfields $s and $t to make them more indicative of their scope. The proposal was unanimously accepted as written.

**Proposal No. 2015-03: Description Conversion Information in the MARC 21 Bibliographic Format**
Presented by the Library of Congress. This proposal originated from a need to encode data related to conversion processes by Stanford’s BIBFRAME project. LC had originally looked at field 883 (Machine-generated Metadata Provenance) but concluded that the scope of this field was limited to machine-generated data in a field in the record so a new field (884) would be necessary to note machine-converted data for a whole description/record.

There was some question as to whether this field would be useful in a shared cataloging environment beyond the specific conversion projects, but ultimately, a reasonable use case had been defined so a new field was warranted. Another concern was whether it was necessary to define a subfield to flag a record to indicate the record had been updated and brought up to standard post-conversion; however, consensus was that current MARC coding would suffice for that purpose.

The proposal was accepted, with one abstention, and the following revisions: (1) change the scope definition to read “Used to provide information about the origin of a MARC record which has been converted by machine from another metadata structure.”; (2) incorporate the addition of subfield $b (Version number of transformation); (3) modify subfield $k label to “identifier of source metadata;” (4) modify subfield $k definition by replacing “control number” with “identifier” and by striking the word “bibliographic;” and (5) change subfield $d (Conversion date) to subfield $g (Generation date).

**Proposal No. 2015-05: Definition of New Code for Leased Resources in Field 008/07 in the MARC21 Holdings Format**
Presented by the British Library. This paper proposes the definition of a code q in MARC Holdings Format 008/07 (Method of Acquisition) to represent those resources that were acquired through lease. The proposal was unanimously accepted, with a minor correction to the proposal number in the header of the body.

**Proposal No. 2015-06: Defining New Subfield in Field 382 for Coding Number of Ensembles in the MARC 21 Bibliographic and Authority Formats**
The Music Library Association sponsored this paper. This paper proposes to define a subfield ($e) for recording number of ensembles in Bibliographic and Authority fields 382 (Medium of
Performance). Some clarification was sought on whether this new subfield would somehow disable the utility of the algorithm pairing subfield $n$ (Number of performers of the same medium) and subfield $s$ (Total number of performers). The proposal passed unanimously as written.

Business Meeting

Library of Congress Report
The Library of Congress is embarking on a BIBFRAME Pilot with 25 catalogers working in BIBFRAME Editor. The records created will be transformed into MARC. There are some things in MARC that will never be in BIBFRAME Vocabulary.

Other
The representative from the German National Library wondered if the MARC Advisory Committee needed a public relations person. MAC meetings are open, but there is a desire to attract more interest in our work.

MARC Formats Transition Interest Group
Carolyn Hansen (Co-chair of MARC Formats Transition Interest Group) announced that the group recently updated their charge to reflect the desire to identify and learn about MARC transition projects; create a community to discuss best practices, challenges, and successes around MARC transition; and raise awareness around BIBFRAME as a standard for transition. Vicki Troemel (MARC Formats Transition Interest Group Co-chair) introduced the speakers.

BIBFLOW: An IMLS Project
Xiaoli Li (Co-head of Content Support Services Department, University of California Davis) provided an update on the two-year IMLS grant-funded collaborative project between UC-Davis and Zepheira, “Reinventing Cataloging: Models for the Future of Library Operations.” This research project will address the impact that adoption of BIBFRAME (BF) will have on technical services workflows in academic libraries by developing a roadmap that the library community can use in their transition.

After mapping out a very complicated network of programs, services, and workflows that touch bibliographic data at UC-Davis, Li concluded that Linked Data (LD) represents an evolutionary leap for libraries and not a simple migration. Based on identified workflows, articulated use cases, and initial assessment on the Kuali-OLE product (their chosen library management system), Li and her team have developed a preliminary Linked Open Data (LOD)/BIBFRAME implementation model which allows for the gradual, phased shifting of library work efforts from a MARC to LOD/BF environment.

Upcoming plans include working closely with Zepheira to enhance Scribe with more profiles; collaborating with Kuali-OLE to make the product Linked Data ready; developing and testing data transformation tools and services; using Blacklight to show how LD really works for the discovery layer; and developing automatic conversion tools to create two versions in each datastore (MARC and triplestore LD) with a desire to eventually phase out the MARC record store altogether.
Li closed by sharing her vision (“Xiaoli’s dream”) to break down silos by holding data in triplestores. She hopes that the OPAC is not a one-size-fits-all, but rather, can feature a customizable interface that can be tailored to certain community’s needs. And finally, she hopes our systems can accommodate patron-added tags and data.

She encouraged anyone interested in providing feedback or following the project to comment on the project blog at http://www.lib.ucdavis.edu/bibflow/.

**Experiments in BIBFRAME: A modular approach**

Nancy Fallgren (Metadata Specialist Librarian, National Library of Medicine) discussed NLM’s project to develop a modular BIBFRAME (BF) model that can accommodate the descriptive standards that experts have already developed in an effort to divorce BIBFRAME from its “MARC replacement” status. She posited that the BF model should not be based on any existing standard; rather, descriptive standards should be used to drive the development of the model and these standards should be equivalently mapped behind-the-scenes.

This practical approach would entail establishing a BF Core Vocabulary that is informed by descriptive metadata schemes across the commonly used vocabularies which could then be extended. The fictitious BF/RDA Profile User Interface would use BIBFRAME as a communication tool only. The cataloger would see RDA elements and input values accordingly. Fallgren then presented a proof of concept, illustrating the challenges of describing versatile resources from a number of disparate sources.

Progress to date includes: the drafting of a BF Core Vocabulary by Zepheira, mapping PCC RDA BIBCO standard and CONSER standard record to Zepheira’s BF Core, and mapping the PCC core map to RDA/RDF constrained vocabulary which is currently in review. Next steps include: reviewing and refining vocabulary with Zepheira; sharing the BF Core version 1 observations with the BF community, PCC, and JSC; testing the mapping and conversion of legacy metadata; testing the generation of new bibliographic data with the developed RDA profile for Zepheira’s Scribe tool; and collaborating with Jackie Shieh of George Washington University. In spring 2015, they’ll begin parallel testing with the Library of Congress.

**Library Linked Data Interest Group**

**Opening up science with VIVO**

Kristi Holmes (Director of Galter Health Sciences Library at Northwestern University and a VIVO Project Engagement Lead) gave some background on the VIVO grant project which began in 2009. VIVO is an open source semantic web application and an information model with an open community. It serves to address the challenges institutions face given that research has become increasingly inter-disciplinary, finding collaborators has become more difficult, and administrators are challenged with the variety of research that is available. VIVO facilitates research networking by taking information about research and researchers and presenting them in a web-based infrastructure.
Linked Data allows for visualizations, research and clinical data integrations, and deep semantic searching across multiple types and sources of data. VIVO harvests verifiable information from institutional datastores, government information, etc., and VIVO profiles allow for a variety of useful activities such as finding collaborators, keeping tabs on competitors, providing a space to build a portfolio, and much more.

Holmes then presented recommendations for research networking by demonstrating various ways that institutions have leveraged the large web of data that is available. VIVO profiles can be used to generate visualizations, can create standard CVs, can be aggregated and searched across, can be used to develop a semantic recommendation engine for team-building, and can populate directory information for people.

Many developments have been made to VIVO Search, a multisite search. ICTS search is structured semantic data for 42 universities. VIVO Searchlight is a little bookmark that fits in the browser and can be used to search on someone in a given institution. For more information, the VIVO community maintains an active wiki page and hosts a number of events.

**BIBFRAME: The way forward for library visibility on the web**

Victoria Mueller (Senior Information Architect and System Librarian, Zepheira) focused on the theme of web visibility for library data. She began by setting the context through a look at the history of Linked Data (LD) on the web, and more specifically, the landscape of library data on the web. She noted that the Web became the most pervasive data management and integration platform because it was effectively designed for human consumption, by hiding most of the data.

Mueller then discussed RDF as a common model for creating web data by constructing triples (subject-object-predicate). BIBFRAME (BF) represents a move toward LD standards and serves a MARC replacement. She noted that it was not all about bib data, but that communities could define BF profiles to meet their needs beyond bibliographic description. As such, the LibHub initiative (Leading, Learning, and Linking) began to increase visibility and discovery of library assets. LibHub seeks to take richly described library assets and reflect that data on the web using BIBFRAME.

She concluded by urging librarians to rethink their role in making memory organizations become the credibility engines on the web, to learn about Linked Data standards and BIBFRAME, and to participate in the LibHub Initiative.

**Linked Data for Libraries (LD4L)**

Nancy Lorimer (Interim Head of Metadata Department, Stanford University) provided an update on the 2-year Andrew W. Mellon Foundation grant, “Linked Data for Libraries” (LD4L), a collaborative project between Cornell University Library, Harvard Library Innovation Lab, and Stanford University Libraries. The project began in January 2014 with a goal to create a linked open data standard by leveraging existing work using existing ontologies and Open Source technology.

Lorimer described the technical framework which included Fedora 4 as repository, Solr for indexing, Blacklight for faceted search and view on objects, and a hydra-head plugin to create,
update, delete, etc. The project participants are focused on bibliographic data (e.g., MARC, MODS, EAD), person data (e.g., VIVO/CAP, ORCID, ISNI, VIAF), and usage data (e.g., circulation, citation, curation). They hope to connect these sources by connecting existing ontologies, linking to global identifiers, and supporting the resulting annotations.

The ontology is based on BIBFRAME (BF). Stanford is currently testing the BF converter with MARC data, but recognizing the need for non-MARC originating data, they are developing VRA flavors, MODS to BF converter, etc. Given the complexity of the ontology, which consists of library resources, people, and much more, they face a number of challenges. In effort to guide them and help address these challenges, use cases – literal stories of potential types of research – have been developed.

The project consists of three phases: Phase 1 - annotations (bib and curation data); phase 2 – authorities (bib and person data); and phase 3 – linked open data (leveraging external data). Phase 1 is near completion by a member of the engineering team; and phases 2 and 3 are in the planning and engineering stage. The project timeline for January through June 2015 is to start piloting instances of LD4L and to develop a test instance for LD4L search application harvesting RDF across the three institutions. In February 2015, Stanford is hosting a 2-day workshop for 25 attendees from 10-12 interested library, archives, and cultural memory institutions. July through December 2015, they aim to implement fully functional LD4L instances, publicly release the code, and much more. The final project outcome is the development of an open-source LD4L ontology as well as a semantic editing, display, and discovery system that is Hydra compatible.

**CaMMS Forum**

**BIBFRAME: notes from the field**

**A Brief Overview of BIBFRAME**

Angela Kroeger (Access & Metadata Project Coordinator, University of Nebraska at Omaha) presented an informative overview of BIBFRAME. She summarized BIBFRAME as a data model for resource description that breaks apart the bib record by looking at resources as an assembly of linked metadata pieces with relationships. It is not based on RDA or FRBR and is still a work-in-progress.

BIBFRAME tools available from the Library of Congress include the MARC-to-BIBFRAME Comparison, MARC-to-BIBFRAME Transformation Tool, and BIBFRAME Editor. Anticipated developments include a search and display tool and a profile creation and editing tool.

Kroeger explored the rationale behind moving away from MARC, specifically that MARC is difficult for machines to interpret, and in order to expose library data on the web, libraries should turn to standards that leverage the Linked Data environment, namely RDA and FRBR. She then demonstrated the Missouri-Kansas Conflict website’s relationship viewer, as an example of how LD data can help make connections.

Kroeger covered the differences between the two existing BF editors available. The Library of Congress’ BF Editor provides a very detailed, complex, highly granular workform while Zepheira’s Scribe is much shorter with no separation on whether the data is at a work, manifestation, or item level.
Finally she highlighted some current developments including OCLC’s work with the W3C Schema bib extent community group’s schema.org extension vocab which evolved into bibliograph.net. This provides a mechanism for working toward one another—BF for description, curation, and data exchange; and schema.org for discovery via common search engines. Zepheira’s LibHub initiative is aiming to publish BF resources on the open web and track increased visibility of library data.

**Stanford and Linked Data** (Nancy Lorimer)

Stanford librarians began their work with linked data in 2011, by attending a Linked Data Workshop. In 2014, the library became a partner in the LD4L project, which has the goal of investigating linked open data connecting academic information resources. The project is a partnership of Stanford University, Cornell University, and Harvard University, and is grant-funded. The project works bibliographic data in many formats (MARC, MODS, VRA, EAD), data on persons from VIVO, ORCID, ISNI, and VIAF, and usage data from Harvard’s circulation statistics. LD4L has as outcomes: creating an ontology compatible with VIVO, BIBFRAME, and other Linked Open Data efforts; creating an open source end result; creating a transparent mapping from MARC to Solr; and testing the MARC-to-BIBFRAME converters.

Two problems encountered in the project, particularly with the MARC-to-BIBFRAME testing, are the urge on the part of librarians to want to recreate MARC in BIBFRAME, and the urge to think of “records” in BIBFRAME. A sort of benefit of the project is that it highlighted inadequately represented data in current library databases, particularly in the categories of technical information and title authorities.

Going forward, Stanford plans to begin creating data in BIBFRAME, though they will ease into the process as opposed to moving 100% to BIBFRAME right away. Stanford is also developing a new project, which will include examining how BIBFRAME works for copy and original cataloging and how it works with musical sound recordings. The examination of musical sound recordings will provide a good example of complexity in BIBFRAME testing, since music recordings are often compilations with multiple works and multiple composers. In MARC, a user needs to read such a bibliographic record as a whole in order to get the relationships between all the works, creators, performers, and attributes. These relationships are not built into MARC, so converting from MARC to BIBFRAME for such materials is more difficult. Stanford will be looking at working with the music cataloging community in creating profiles and working on the problems of musical sound recordings in BIBFRAME.

**George Washington University Libraries and BIBFRAME** (Jackie Sheih)


Linked data is a method of publishing structured data. For librarians to understand linked data in the context of BIBFRAME, an important place to start is learning the basics about BIBFRAME classes and properties. For example, a BIBFRAME class is named with a capital letter and is always singular; a BIBFRAME property is a lower case letter; and a BIBFRAME object includes the word “has” as a verb.
An important part of BIBFRAME testing for George Washington University was analyzing their MARC data. In the process of doing that, they realized they needed information on past cataloging practices to know why data was entered or structured in certain ways. For examples in their BIBFRAME analysis, they chose simple resources having one author, few subjects, etc.

Upcoming goals for BIBFRAME work at George Washington University in 2015 include: community testing of core vocabularies, working on a web interface for practitioners to use in data input, and searching and retrieving native MARC and BIBFRAME.

**Q&A from session**

Q: How concerned should librarians be about obsolete fields, such as the 440 series entry
A: George Washington University updated those types of fields themselves, which helped in converting MARC records to BIBFRAME. It is possible that OCLC may do some updating of obsolete fields as well.

Q: If we use the premise that entity modelling in authority records is more successful than it is in bibliographic records, perhaps authority records will work better and need to be replaced less in BIBFRAME?
A: Perhaps. There isn’t an authorities-type module in BIBFRAME, yet.

Q: How much do catalogers need to know to really understand BIBFRAME?
A: Stanford had two levels of training. One group was training in RDF and XML, and the other group was training at a different level. It is likely that many institutions will use this model. Every cataloger should understand the concepts of linked data, but not everyone needs to know RDF.
A: PCC is also working on BIBFRAME training for catalogers.