White Paper Report

Report ID: 102646

Application Number: HD5126211

Project Director: Kristin Negulescu (kcarpenter@archive.org)

Institution: Internet Archive

Reporting Period: 5/1/2011-9/30/2011

Report Due: 12/31/2011

Date Submitted: 1/31/2012

Final Report: Linked Open Data in Libraries, Archives, and Museums Summit (LOD-LAM)

Report Date: January 31, 2012

Grantee Institution: Internet Archive
Grant Number: HD-51262-11

Project Director: Kris Carpenter Negulescu, Director Web, Internet Archive
Project Web site: http://lod-lam.net/summit/
Summit Organizer & Facilitator: Jon Voss (@jonvoss), Founder, LookBackMaps/Historypin Strategic Partnerships Director

Section One: Project Description and Activities

Thanks to the generous support of the National Endowment for the Humanities and the Alfred P. Sloan Foundation, the Internet Archive convened a two-day summit in San Francisco, June 2-3, 2011 to foster greater public access to metadata in the world's libraries, archives, and museums through increased adoption and implementation of Linked Open Data.

The International Linked Open Data in Libraries, Archives, and Museums Summit ("LOD-LAM") convened leaders in their respective areas of expertise from the humanities and sciences to catalyze practical, actionable approaches to publishing Linked Open Data, through:

- Identification of the tools and techniques for publishing and working with Linked Open Data.
- Drafting of precedents and policy for licensing and copyright considerations regarding the publishing of library, archive, and museum metadata.
- Publishing of definitions and promotion of use cases that give LAM staff the tools they need to advocate for Linked Open Data in their institutions.

Participants had to apply to attend. The ideal attendee was a programmer, administrator, lawyer, LAM professional, or other professional with at least a working understanding of Linked Open Data, if not some direct experience with the technology or policies involved. Participants needed to have authority in their position to implement policy or technology, or influence decision makers in their institution or sector. A selection committee was formed to review applications and recommend participants. The selection committee looked for people that had organized others in their field around Linked Open Data and who had a wide sphere of influence. They actively sought representative candidates from a broad range of institutions with diverse levels of leadership and technical expertise. Application submissions opened at 8am, PST February 1, 2011, and closed 5pm PST, February 28, 2011. Over 150 applications were received for 50 available slots. Given the enthusiastic response, the event budget was reworked to accommodate up to 100 participants. Participants were selected and notified by March 7, 5pm PST. Ultimately, 100 participants from 85 institutions attended the summit.

The following individuals served on the organizing committee that helped recruit and select the participants:
Lisa Goddard (@lisagoddard), Acting Associate University Librarian for Information Technology, Memorial University Libraries.
Martin Kalfatovic (@UDCMRK), Assistant Director, Digital Services Division at Smithsonian Institution Libraries and the Deputy Project Director of the Biodiversity Heritage Library.

Mark Matienzo ([@anarchivist](#)), Digital Archivist in Manuscripts and Archives at the Yale University Library.
Mia Ridge ([@mia_out](#)), Lead Web Developer & Technical Architect, Science Museum/NMSI (UK)
Tim Sherratt ([@wragge](#)), National Museum of Australia & University of Canberra
MacKenzie Smith, Research Director, MIT Libraries.
Adrian Stevenson ([@adrianstevenson](#)), Research Officer, UKOLN; Project Manager, LOCAH Linked Data Project.
John Wilbanks ([@wilbanks](#)), VP of Science, Director of Science Commons, Creative Commons.

The LOD-LAM Summit utilized the Open Space Technology meeting format to give this group of expert innovators the time and space to freely identify and address the most pressing issues related to forwarding Linked Open Data in libraries, archives, and museums. The summit was convened to address one specific question; "How do we expand adoption of Linked Open Data amongst Libraries, Archives, and Museums?" and was based on two primary principles - passion and responsibility: passion to jump in and play an active role; and responsibility to lead, and follow through with action.

The summit opened with a session in which the participants collaboratively created the agenda for breakout sessions for the first day. Because the LOD-LAM Summit was action-oriented, a similar process occurred on the second day, but with a focus on deliverables, documentation, and collaboration thought to be achievable during the calendar year following the summit.

The primary means of communication publicizing the event and encouraging applications to attend was facilitated via email lists. Information about the event was sent to libraries, archives, museums and memory institutions via the following lists:
- W3C Library Linked Data group
- LOD-LAM group
- Freebase user group
- Lotico Semantic Web group
- International Internet Preservation Consortium (IIPC) Members list
- NDIIPP list
- Society of American Archivists list
- Digital Library Federation list
- DIGLIB list

The primary means of communication and dissemination for the project (pre & post meeting) occurred via:
1) Listserv & Tweets: http://groups.google.com/group/lod-lam, #LODLAM
2) Event Web site: http://lod-lam.net/

3) Regional Meet-ups and report-outs by event coordinators and event attendees at leading conferences and meetings in the United States and abroad (See Section Seven below).

Please refer to the Appendices to view:
- Appendix A: List of Summit Participants
- Appendix B: List of Summit Sessions - Days One & Two
- Appendix C: LOD-LAM Summit Application
- Appendix D: Raw Summit Notes and Reference Materials
- Appendix E: Proposal for a 4 star classification-scheme for linked, open, cultural metadata
- Appendix F: Sampling of live resourcing on Twitter during the first day

Section Two: Key Accomplishments

The first International Linked Open Data in Libraries, Archives, and Museums Summit ("LODLAM") was the first of its kind gathering of leading practitioners, industry leaders, and proponents of Linked Open Data in libraries, archives and museums to define the field, and address barriers to adoption. For the critical period of one year following the international summit, LOD-LAM continues to foster greater collaboration and leadership across disciplines within libraries, archives, and museums through:
- continued advocacy for Linked Open Data at key conferences and annual meetings;
- dissemination of tools, research and best-practices to enable librarians, curators and archivists to champion Linked Open Data at their own institutions.

Original Work Plan & Assessment of Results

Goal #1: LOD-LAM will utilize the Open Space Technology meeting format to address one specific question or theme; "How do we expand adoption of Linked Open Data amongst Libraries, Archives, and Museums?" The participants will have a passionate interest in Linked Open Data and represent a wide array of libraries, archives, and museums, assuring that the most important questions are documented and addressed.

> **Result: ACHIEVED.** The Open Space Technology format was adopted. The meeting was facilitated by Jon Voss, an expert in open space technologies and their application to the humanities. The format included an initial session in which the participants collaboratively created the agenda for breakout sessions for the first day. A similar process happened on the second day, with a more detailed focus on actionable items and publication.
>
> **The LOD-LAM summit attracted the participation of over 100 people engaged in linked data from 85 different organizations from around the**

globe. [Please refer to Appendix A for a full list of the attendees]. While it might be difficult to measure the passion of the group quantitatively, the output of the meetings can serve as a proxy along with the applications submitted by each participant. No one stood on the sideline. Each participant was engaged and actively contributed to the work of the meeting and the meet-ups and projects that followed.

During the summit participants examined the barriers and incentives to linking cultural and scientific metadata and attempted to address the following questions in the context of the humanities:

1. The tools and techniques for publishing and working with Linked Open Data are many and varied. What are the best practices and tools that make publishing Linked Open Data feasible?
2. Licensing and copyright questions abound regarding the publishing of library, archive, and museum metadata. What precedents exist and how can we simplify the set of decisions/tradeoffs each institution must consider when publishing and consuming linked data?
3. Due in part to a lack of concrete examples, many who are experimenting with Linked Open Data struggle to describe it in lay terms, even as it has become a catchphrase and popular topic at a wide variety of humanities conferences. How can we create simple definitions and descriptions that will give humanists the tools they need to explain Linked Open Data, and its importance to their institutions?

Goal #2: **Institutional decision makers can gain a firm grasp on the techniques, technology and terminology of Linked Open Data.**

Result: IN PROGRESS. There are several resources that were added by attendees to the LOD-LAM meeting (listed below) that support this goal. Sessions at the Summit helped identify ways to best summarize the technology and terminology of Linked Open Data. But, we are still lacking simple, easy to understand introductory tutorials to Linked open data that address terminology, technology AND techniques.

Key resources stemming from the summit include but are not limited to:
- **Intro to LODLAM Video** (http://lod-lam.net/summit/2011/09/15/intro-to-lodlam-talk-live-from-the-smithsonian/)
- **LODLAM.net community blog with ongoing news and resources** (http://lodlam.net)
- **LOD-LAM Google Group** (http://groups.google.com/group/lod-lam?pli=1)
- **LODLAM Reading Lists** (http://lod-lam.net/summit/2011/04/25/lodlam-reading-lists/)

- **LOD-LAM Tutorials From DLF Workshop**
  (https://docs.google.com/?tab=oo&authuser=0#folders/0B7v0ey4WR_Gn
  YmI5MTNkOGEtY2UwYi00NDEzLWJhZjltMDIyNGI3ZTVmYWUw)
- **LOD-LAM Zotero** (https://www.zotero.org/groups/lod-lam)

Goal #3: Participants will develop a collaborative strategy to forward the adoption and promotion of Linked Open Data in the humanities in the United States.

Result: IN PROGRESS. This was a very ambitious goal and one that has not yet been fully met. However, significant progress was made toward this end. Specific proposals for how to integrate OAI-PMH and linked data and for a 4 star classification-scheme for linked open cultural metadata that would be used to simplify decisions regarding licensing and rights were put forth during the summit. Many other aspects of adoption and promotion were addressed beginning with efforts to define Users and Use Cases, methods of enhancing and integrating data via crowdsourcing, the strengths and weaknesses of existing publishing tools, as well as barriers to integration of linked data sets and enhanced consumption models.

Some of the most inspired discussions stemmed from demonstrations that came at the end of Day One of the event when attendees shared announcements, ideas, working prototypes, and even live services and applications built on top of multiple services that had not previously been broadly promoted. The innovative, and often "under the table" or "backroom/basement" work happening at small and large scales in isolation within museums, libraries, and archives served to illustrate the real opportunity if implementations could be made public and more accessible to the mainstream Web user a la Wikipedia. Attendees all agreed that what was needed most were more very public reference implementations working across data sets and institutional sectors, including the for-profit sector. What prevented that most was a lack of clarity around licensing issues. This event went further than any that preceded it in crystallising LOD rights issues and in proposing a concrete licensing framework for publishing and consumption of linked open cultural metadata.

Goal #4: Actionable steps and strategies for sharing, publishing and licensing structured data on the Web will be widely distributed and publicized.

Result: ACHIEVED. One of the most seminal accomplishments of the LOD-LAM summit was the drafting and dissemination of the **4-star classification-scheme for linked open cultural metadata.** This proposal clarifies licensing and rights options for those institutions interested in exposing and/or consuming cultural metadata. To view the full draft of the proposal, please refer to Appendix E.

Goal #5: **Provide limited support of a joint effort on at least one use case that can convincingly illustrate the potential of Linked Open Data in libraries, archives, and museums.**

> **Result: IN PROGRESS.** A working group was convened at the summit to focus on the pilot project Civil War Data 150. Participants included representatives from universities, libraries, archives, and museums. The group considered ways to continue moving forward with the project and how the approach could also serve as a model for other domains. Since the summit, we've learned a great deal about the various roadblocks to demonstrating the use of Linked Open Data by defining and attempting to implement a series of engaging apps that rely on a variety of data sources using Linked Data to consume and integrate them successfully. Results of the collaborative should begin to emerge in the next 12-18 months as the anniversary progresses.

Organizational Timeline and Critical Responsibilities

--Identify prospective participants. (COMPLETE - March 2011).
--Invite participation, announce event and funders, and build website (Feb/Mar 2011).
--Create Collaborative Web Space and Google group (Mar 2011).
--Hold Meeting (Hotel Kabuki, San Francisco. June 2-3, 2011).
    Day 1. Open Sessions (See Appendix for full agenda)
    Day 2. Action Items and documentation (See Appendix for full agenda)
--Disseminate Outputs and Findings (Ongoing, June 2011-present)

The full proceedings of the meeting were published on www.lod-lam.net within six months after it was held. Results were posted in waves as follow-up meetings and discussions occurred fleshing out the ideas originally documented back in June in San Francisco.

Section Three: Audiences

The primary audiences for this event were cultural heritage/memory institutions (libraries, archives, museums), nongovernmental organizations, and not for profits who publish and/or consume cultural linked data or that would like to be doing so. Organizations addressing scientific and other data needs were also invited along with commercial entities actively engaged in advancing Linked Open Data globally.

The funding provided by NEH was used to sponsor the involvement and participation of representatives of cultural heritage institutions from the United States whose mission is in part to support promotion and dissemination of the humanities and humanities research.

Section Four: Evaluation

The LOD-LAM Summit was a success on many, many levels. The event brought together a diverse group of library, archives, and museum professionals from the United States and around the globe to identify shared interests and needs regarding development and implementation of Linked data strategies. Public response to the project was decidedly positive, and participation in dialogues and actionable outputs has continued amongst a broad spectrum of attendees and some new recruits to the work, 7+months post event. On the other hand, the LOD-LAM Summit might have missed the mark in the following ways:

- Response to the event was so overwhelmingly positive, the organizing committee decided to expand the forum from 50 to 100 attendees. While this was the right decision for this meeting and achieved tremendous results in terms of cross fertilization of LOD-LAM best practice, nationally and globally, it might have hindered the scope of hands on, actionable work that could happen *at* and, especially, following the event. This is hard to imagine given how much great work was produced, but in order to incubate production level, cross institutional, cross-sector, ongoing collaborations you need very small, focussed sessions with almost a singular goal and set of tasks in mind, and, ideally, some dedicated funding to fuel immediate participation. A smaller forum might have spurred more concrete outputs beyond best practice documentation, recommendations, and proposals. To this point, event organizers did coordinate efforts with the Stanford Linked Data Workshop, a week-long intensive with a group of about 20 participants, some of whom overlapped with the LOD-LAM Summit.

- There are still heated technical debates over what are appropriate standards and tools for implementation and adoption of LOD and specifically of LOD-LAM [(e.g. http://faviki.wordpress.com/2011/10/20/the-challenge-of-building-the-semantic-web/)](http://faviki.wordpress.com/2011/10/20/the-challenge-of-building-the-semantic-web/). The event organizers had hoped to produce a concrete list of recommended tools and implementation approaches, a 'How To' of LOD-LAM if you will for newbies to the space, as a result of the meeting. Break-out sessions addressed the full range of issues and perspectives representing the debate and the needs and desires of those new to LOD, but there was no opportunity to define and assemble the 'How To' in detail at the event. Had there been the opportunity to fund a one or two day, working collaboration immediately following the summit with a much smaller subset of the participants whose goal it was to produce a "strawman" proposal for the 'How To' and to demonstrate publication and consumption of LOD between museums, libraries, and archives in the US using available tools and best practices, there would have been a long list of willing participants and the goal would have been achieved to the full degree originally envisioned. This might be a format to consider for future working meetings in which the actionable output of the meeting would benefit from execution in stages including hands on, face to face, participation over multiple days. Careful consideration should be given to the idea of an "incubator" style approach to future LOD-LAM funding by which small group forums can be assembled to address specific areas of need.

Section Five: Continuation of the project

The LOD-LAM website, Google group and hash tag (#LODLAM) will serve as continuation of the collaborative web space that was created in advance of the summit, and will provide an ongoing location for tracking LOD-LAM related initiatives.

To encourage the widest distribution of the findings and to encourage continued discussions and concrete actionable results, the group identified shared topics such as the anniversary of WWI, and the US Civil War as unifying themes that would help inspire collaboration across institutions and disciplines. The group also identified discipline-specific conferences that would be held throughout the year following the event at which to present the findings and to educate a wider audience about the basic technological and legal issues behind Linked Open Data, use cases, and specific ways in which libraries, archives, and museums of any size can begin to publish structured data. Those events are listed in **Section Seven** below. Upcoming events in the US include:
- Linked Data Day - NY (METRO, NYPL Labs, and NYU http://www.metro.org/en/art/488) FEB 2012
- Museums and the Web - April 2012

Section Six: Long Term Impact

Linked Open Data: Innovation in the Humanities

Within the humanities, where vast stores of data in libraries, archives, and museums have long lived in isolation, Linked Open Data is proving to offer incredible potential for collaboration. The ability to publish structured data in increasingly accessible ways is beginning to show promise for connecting disparate archives, improving and sharing bibliographic and archival data, and vastly improving discoverability. We're already beginning to leverage the power of crowdsourcing to improve upon metadata, geotag historical photos, and link information.

Yet despite the possibilities, only a handful of institutions in the US had begun to embrace Linked Open Data by 2011, and even fewer had been awarded grants to forward these efforts. The US lagged far behind Europe and Australia/New Zealand in terms of institutional innovation and adoption of Linked Open Data. It was not without good reason, yet the obstacles were by no means insurmountable.

The following provides a brief scan of Linked Open Data within the humanities in the US post-summit. Members from each of these sectors and all the projects mentioned (with the exception of the NARA project) were present at the LOD-LAM summit in June 2011.

Individual institutions within the US are beginning to experiment with creating Linked Data and releasing Open Data, and though it's very early on, several examples demonstrate the scope and uses actively under development in the humanities:

- **Library of Congress**: **The Bibliographic Framework Initiative General Plan** (http://www.loc.gov/marc/transition/news/framework-103111.html), announced in October 2011, is an ambitious effort to move the U.S. library community toward a modern method of exchanging bibliographic data including the use of the World Wide Web Consortium's (W3C) Resource Description Framework as a data model, which is the preferred method for publishing linked data, and implant libraries in an environment conditioned by the technologies of the Semantic Web and linked data principles. (http://www.libraryjournal.com/lj/home/892637-264/library_of_congress_provides_details.html.csp)

- **NARA**: On December 23, 2011, The National Archives of the United States launched the Citizen Archivist Dashboard, a portal for public participation in crowdsourcing activities. The public can now contribute tags, transcriptions, images, and collaborate on articles that help make records more accessible. (http://www.archives.gov/citizen-archivist/)

- **Smithsonian**: In late September 2011, Smithsonian Libraries' digital projects librarian Keri Thompson and lead developer Joel Richard, along with Trish Rose-Sandler of the Missouri Botanical Garden, presented "Building the New Open Linked Library: Theory and Practice." The talk gave a high-level overview of the redesign of the Smithsonian Libraries' website, a brief summary of Linked Data, how the Libraries' website redesign centers around the concept of Linked Open Data, and some of the unique things that happen when open data is made available on the web, specifically with the Biodiversity Heritage Library (BHL). (http://smithsonianlibraries.si.edu/smithsonianlibraries/2011/10/lita-national-forum-2011-notes.html)

- **Stanford University**: The Stanford University Libraries and Academic Information Resources (SULAIR) with the Council on Library and Information Resources (CLIR) conducted a week-long workshop with 20 particpants on the prospects for a large scale, multi-national, multi-institutional prototype of a Linked Data environment for discovery of and navigation among the rapidly, chaotically expanding array of academic information resources. Due to the diversity of knowledge, experience, and views of the potential of Linked Data approaches, the workshop participants focused on two primary goals: building common understanding and enthusiasm and on identifying opportunities and challenges to be addressed in the definition, development and operation of a LOD prototype. A technology plan was also produced as an output of the workshop six months

following the event. (www.clir.org/pubs/reports/pub152/**LinkedData**Workshop.pdf) Stanford also published millions of bibliographic records as LOD in 2011. Stanford Libraries plans to continue an alliance with Metaweb/Freebase in transcoding bibliographic facts to URIs, and will continue to send HighWire micro-data to Schema.org, which then finds its way to DBpedia. They are also exploring linked data collaborations with the British Library, the British Museum, and JISC.

- **Emory University**: The University started with a focus on state of the art creation and application of linked data sets to support health services and outcomes research. They have more recently expanded their efforts to include their digital collections in the humanities and other major disciplines.

- **University of Richmond**: The Digital Scholarship Lab project "Hidden Patterns of the Civil War" has virtually intregated a number of interrelated projects on the sectional crisis, slavery, and emancipation during the Civil War era, with a particular emphasis on the histories of the city of Richmond and the state of Virginia. Grouped as "texts" and "maps," these projects use digital tools and digital media to uncover and represent patterns that are not easy to find when examined solely in isolation. (http://dsl.richmond.edu/)

- **Archives of Michigan**: The Archives of Michigan initially conceptualized the Civil War Data 150 project as an opportunity to leverage state efforts going into the Civil War Sesquicentennial and build on that momentum to begin making data connections across local, state, and federal archives. SeekingMichigan.org, a joint project with the Archives of Michigan and Library of Michigan, now provides a unique digital combination of presentation and discovery of state holdings, including a large collection of Civil War related assets.

- **W3C Library Linked Data Incubator Group**: The final report (http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/) of the W3C Library Linked Data Incubator Group, released in October 2011, urges the library community to reconceptualize metadata and publish it to the Web using linked data technologies so that it will play well with non-library datasets on the Web. Several authors of that report attended the LOD-LAM summit. The "Use Cases" report (http://www.w3.org/2005/Incubator/lld/XGR-lld-usecase-20111025/) describes library applications which showcase the benefits of adopting Semantic Web standards and linked data principles to publish library assets such as bibliographic data, concept schemes, and authority files. The "Datasets, Value Vocabularies, and Metadata Element Sets" report provides a snapshot of key resources available for creating library linked data today. (http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/)

- **Linked Open Data Cloud in September 2011**
  (http://richard.cyganiak.de/2007/10/lod/)

Section Seven: Grant Products

The outputs of the LOD-LAM Summit can be summarized as follows:
1. **Presentations, Proposals and Visualizations**
   - **4 Star Classication Scheme** (http://lod-lam.net/summit/2011/06/06/proposed-a-4-star-classification-scheme-for-linked-open-cultural-metadata/#comment-2269), http://lod-lam.net/summit/2011/06/08/mackenzie-smith-on-open-licenses-for-data/
   - **Open Discussions - Users, Uses, Services** (http://lod-lam.net/summit/2011/06/03/users-uses-service/)
   - **Library Linked Data cloud – a teaser** (http://semantic.ckan.net/group/?group=http://ckan.net/group/lld)
   - **Open discussions - Vocabularies**: http://lod-lam.net/summit/2011/12/18/time-and-tide/, http://lod-lam.net/summit/2012/01/12/vocabulary-alignment-meaning-and-understanding-in-the-world-museum/,
   - **Beyond OAI-PMH** (http://lod-lam.net/summit/2011/06/02/beyond-oai-pmh-report/)
2. **Reference Resources**
   - **Intro to LODLAM Video** (http://lod-lam.net/summit/2011/09/15/intro-to-lodlam-talk-live-from-the-smithsonian/)
   - **LODLAM Reading Lists** (http://lod-lam.net/summit/2011/04/25/lodlam-reading-lists/)
   - **LODLAM Zotero Group**

3. **Post Summit Meet-ups**
   - **LODLAM Meet-up - NYC** (Voss, Reside) July 2011
   - **LODLAM Meet-up - Washington DC** (Kapsalis, Pilsk, Kalfatovic - Smithsonian) September 2011
   - **LODLAM Meet-up - London, UK** (Ridge) October 2011
   - **LODLAM Meet-up - Emory University** (Voss) November 2011
   - **LODLAM Meet-up - Wellington, NZ** (Wray, Neale, Sherratt) December 2011

4. **Report-outs by event coordinators and attendees at conferences and meetings in the US and abroad**:
   - **Linked Data and Libraries 2011** - London, UK (Stevenson)
   - **SAA 2011** - Chicago (Stevenson, Ferel)
   - **1st International Workshop on Semantic Digital Archives** - Berlin (?)
   - **ALA 2011** - New Orleans, Dallas (Hellman, Harper, Coyle, Voss, Frick)
   - **Museums Computer Network (MCN) - Atlanta** (Chun, Moad, Voss)

- **Digital Library Federation Fall Forum - Baltimore** (Frick/Voss/Harper/Carpenter/Zumwalt)

5. A sampling of outputs from attendees on external sites
- **John Wilbanks**, Weak Ties, Linked Data, and Citation (http://del-fi.org/post/6145070263/weak-ties-linked-data-and-citation)
- **Adrian Stevenson**, Linked Data and Libraries: Report on the LOD LAM Summit (http://www.meanboyfriend.com/overdue_ideas/2011/07/linked-data-and-libraries-report-on-the-lod-lam-summit/)
- **David Weinberger**, Multiple posts and video interviews from the Summit (http://www.hyperorg.com/blogger/tag/lodlam/)
- **Micki McGee**, TravelBlog: The Erotic Life of Data, or LOD-LAM and the Pursuit of Compatible Data (http://fordhamdh.blogspot.com/2011/07/travelblog-erotics-of-data-or-lod-lam_05.html)
- **Rurik Greenall**, LOD-LAM Summit Experiences (http://brinxmat.posterous.com/lod-lam-summit-experiences)
- **Eric Hellman**, Our Metadata Overlords and That Microdata Thingy (http://go-to-hellman.blogspot.com/2011/06/our-metadata-overlords-and-that.html)
- **Lori Jahnke**, I Saw the Internet Today (http://www.medicalheritage.org/2011/06/i-saw-the-internet-today…/)
- **Laura Smart,** Linked Open Data Libraries Archives Museums (LOD-LAM) Summit (http://library.caltech.edu/laura/?p=89)
- **Takeda Hideaki,** (in Japanese) LOD-LAM Summit (http://researchmap.jp/jovjnh3a8-466/#_466)

APPENDIX A: Final List of LOD-LAM Summit Participants

Full Name	Affiliation	Country
Josh Greenberg	Alfred P. Sloan Foundation	United States
Jill Vermillion	Apple, Inc.	United States
Perian Sully	Balboa Park Online Collaborative	United States
John Deck	Berkeley Natural History Museums	United States
Andrew Ashton	Brown University Library	United States
Lisa Dawn Colvin	California Digital Library, University of California, Office of the President	United States
Laura Smart	Caltech	United States
Rachel Frick	CLIR: Digital Library Federation	United States
Lori Jahnke	College of Physicians of Philadelphia	United States
Dean B. Krafft	Cornell University Library	United States
John Wilbanks	Creative Commons	United States
Jonathan Rees	Creative Commons	United States
Thomas Baker	Dublin Core Metadata Initiative	United States
Bradley Allen	Elsevier	United States
Micki McGee	Fordham University	United States
Robb Detlefs	Gallery Systems	United States
Eric Hellman	Gluejar, Inc.	United States
Shawn Simister	Google	United States
Jamie Taylor	Google/Freebase	United States
David Weinberger	Harvard Library Innovation Lab & Berkman Center	United States

Sue Kriegsman	Harvard University	United States
Susan Chun	Independent researcher and consultant	United States
Charles Moad	Indianapolis Museum of Art	United States
George Oates	Internet Archive	United States
Kris Carpenter Negulescu	Internet Archive	United States
Marcia Zeng	Kent State University	United States
Jerry Persons	Knowledge Motifs LLC	United States
Jane Mandelbaum	Library of Congress	United States
Kevin Ford	Library of Congress	United States
Kathy Jordan	Library of Virginia	United States
Jon Voss	LookBackMaps	United States
Dave Lester	Maryland Institute for Technology in the Humanities (MITH)	United States
Matt Zumwalt	MediaShelf, LLC	United States
William Gunn	Mendeley Research Networks	United States
Diane Hillmann	Metadata Management Associates	United States
David Henry	Missouri History Museum	United States
MacKenzie Smith	MIT	United States
Jenel Farrell	MPR	United States
Jerry Simmons	National Archives and Records Administration	United States
Corey A Harper	New York University	United States
Douglas Knox	Newberry Library	United States
Doug Reside	NYPL	United States

Roy Tennant	OCLC	United States
Brian Tingle	Online Archive of California, California Digital Library	United States
Eric Kansa	Open Context	United States
Cristina Pattuelli	Pratt Institute, School of Information and Library Science, New York	United States
Layna White	San Francisco Museum of Modern Art	United States
Melanie Feinberg	School of Information, University of Texas at Austin	United States
Martin Kalfatovic	Smithsonian Institution	United States
Effie Kapsalis	Smithsonian Institution Archives	United States
Aaron Straup Cope	Stamen Design	United States
Benjamin Albritton	Stanford University Libraries	United States
Ethan Gruber	The American Numismatic Society	United States
Mary Elings	The Bancroft Library, UC Berkeley	United States
Piotr Adamczyk	The Metropolitan Museum of Art, New York	United States
Lily Pregill	The New York Art Resources Consortium	United States
Aaron Rubinstein	Tufts University Digital Collections and University Archives	United States
Francesco Spagnolo	UC Berkeley: The Magnes Collection of Jewish Art and Life (Bancroft Library)	United States
Thea Lindquist	University of Colorado at Boulder	United States
Richard J. Urban	University of Illinois	United States
Dave Pcolar	University of North Carolina at Chapel Hill	United States
Ryan Shaw	University of North Carolina at Chapel Hill	United States
Mark Phillips	University of North Texas Libraries	United States

Scott Nesbit	University of Richmond	United States
Kristin Eschenfelder	University of Wisconsin-Madison	United States
Eric Rochester	UVA Scholar's Lab	United States
Asaf Bartov	Wikimedia Foundation and Project Ben-Yehuda	United States
Karen Coyle		United States
Roger Macdonald		United States
Anra Kennedy	Culture24	United Kingdom
Mia Ridge	Science Museum	United Kingdom
Richard Wallis	Talis	United Kingdom
Jo Pugh	The National Archives	United Kingdom
Adrian Stevenson	UKOLN, University of Bath	United Kingdom
Julie Allinson	University of York	United Kingdom
Tyng-Ruey Chuang	Academia Sinica, Taiwan	Taiwan
Rurik Thomas Greenall	Norwegian University of Science and Technology	Norway
Paul Keller	Knowledgeland, Creative Commons NL, Europeana	Netherlands
Uldis Bojars	National Library of Latvia	Latvia
SungHyuk Kim	Sookmyung W. University	Korea, South
Hideaki Takeda	National Institute of Informatics	Japan

Yves Jaques	FAO of the UN	Italy
Imma Subirats	FAO of the United Nations/E-LIS/CIEPI	Italy
Jodi Schneider	DERI, National University of Ireland	Ireland
Markus Geipel	German National Library	Germany
Adrian Pohl	Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen	Germany
Wenz Romain	Bibliothèque nationale de France (BnF)	France
Mikko Heikkinen	Finnish Museum of Natural History	Finland
Antoine Isaac	Europeana + Vrije Universiteit Amsterdam	European Union
Jindřich Mynarz	National Technical Library	Czech Republic
Misty De Meo	Canadian Museum for Human Rights	Canada
Peter Binkley	University of Alberta Libraries	Canada
Alexander O'Neill	University of Prince Edward Island	Canada
Rob Warren	University of Waterloo	Canada
Ingrid Mason	Intersect Australia Ltd - and - Australian National Data Service	Australia
Tim Sherratt	National Museum of Australia & University of Canberra	Australia
Luke Dearnley	Powerhouse Museum	Australia
asa letourneau	public record office victoria	Australia
Jane Hunter	The eResearch Lab, The University of Queensland	Australia
Gavan McCarthy	The University of Melbourne	Australia

APPENDIX B: Sessions - Days One & Two

APPENDIX C: Summit Application

Summit Application Fields included:
- Full Name
- Email Address
- Twitter Handle
- Website
- Institutional Affiliation
- Short Bio
- Interest in LOD-LAM
- Country
- Sector
- Linked Data Projects
- Username for groups/lists, etc.
- Dietary Restrictions
- Request for Travel Support (Y/N & percentage)
- Request for Accommodation (Y/N)

A copy of the raw output of the application surveys is included with this report.

APPENDIX D: Raw Summit Notes & Resources

http://museum-api.pbworks.com/w/page/40871005/LOD-LAM%20live%20blog

http://piratepad.net/491CHtq3Mj

LOD-LAM crowdsourcing session notes

http://museum-api.pbworks.com/w/page/40872182/LOD-LAM%20crowdsourcing%20session%20notes

LOD-LAM Messy data and same-as session notes

http://museum-api.pbworks.com/w/page/40875992/LOD-LAM%20Messy%20data%20and%20same-as

LOD-LAM crowdsourcing, annotations and machine-learning

http://museum-api.pbworks.com/w/page/40881978/LOD-LAM%20crowdsourcing%2C%20annotations%20and%20machine-learning

Power Point Presentation, July 2011, Initial Summit report given in London

http://lod-lam.net/summit/2011/07/14/report-on-the-lod-lam-summit-at-linked-data-and-libraries-2011/

Users, uses, service
Karen Coyle
June 3, 2011

Yesterday at LOD-LAM we talked about users and what users might want to do with data (and thus what we could create for users from LOD). Here's the mind map of that: user verbs (http://kcoyle.net/img/Research.png)

Library Linked Data cloud – a teaser
Antoine Isaac
June 3, 2011

Here's a graphical rendering of library linked data sets on the CKAN LLD group, courtesy of William Waites:
http://semantic.ckan.net/group/?group=http://ckan.net/group/lld

Beyond OAI-PMH Report http://lod-lam.net/summit/2011/06/02/beyond-oai-pmh-report/

Richard J. Urban, June 2, 2011

Thanks to everyone who participated in the Beyond OAI-PMH session this morning. There seems to be a number of places where others have posted their session summaries, but I thought it would be useful to include that here on the LOD-LAM site as well.
This is my brief (and somewhat tardy) account of our meeting based on the notes that I jotted down. Comments and responses are welcome and any misrepresentations are my own.
Two main themes:
1. We need to leverage existing OAI-PMH installation base for Linked Data, because after all it does fit within the basic requirements (three stars) of Linked Data goals.
  ● Promote adoption of tools like OAI2LOD server
  ● Build Sitemaps using existing OAI services
We should acknowledge current OAI-PMH for their existing contributions to Linked Open Data and emphasize that they already are participating in LOD through OAI-PMH. While there are some additional things we can do to make the metadata we share more LOD friendly, LOD is not a completely new idea. Additional documentation about how OAI-PMH has succeed and failed – and what lessons that holds for the future of LOD – would be welcome.
2. We don't necessarily need an OAI-PMH 3.0
It would be better for the community to look towards broadly adopted web standards. Repositories need to provide what users want in multiple serializations, not limited to XML (let alone a specific XML schema).
Some suggestions for alternatives:
  ● Sitemaps
  ● Open Search
  ● Atom
While there wasn't a strong sense that a new OAI standard was needed, there is a recognized need to provide the existing repositories some guidance about alternative approaches. Such guidance should be promoted by funders to help new and existing projects understand how they can contrbute to the Linked Data cloud. There was also a sense that some features of the current OAI protocal might be included in the development of web services:
  ● ability to acquire incremental sets (what's changed, what's new)
  ● an understanding of the "scope" of what's provided (OAI sets/collections)
  ● a minimal set of shared properties (a stub) that is linked directly to richer representations

- some consensus around shared service models to make discovery and use easier
- ability to request sets based on supplied criteria ("search" not just pre-constructed sets)

Of course the devil is in the details, during the session we had several tangential conversations about the technical details of how to implement some of these alternatives that I haven't fully captured here. To me this indicates that further discussion about these different options and how they might be shaped into a common framework is needed and would be valuable guidance for the community.

Additional Comment

There was also a suggestion that OAI-PMH may still be the best way to share large sets of "records" between partners. Rather than worry about making OAI-PMH more LOD friendly, LAMs may wish to focus their energy on providing other kinds of data as LOD (use cases?)

APPENDIX E: Proposed: 4-star Classification-scheme for Linked Open Cultural Metadata

MacKenzie Smith
This entry was posted on Monday, June 6th, 2011 at 4:35 pm.

One of the outcomes of last week's LOD-LAM Summit was a draft document proposing a new way to assess the openness/usefulness of linked data for the LAM community. This is a work in progress, but is already provoking interesting debate on our options as we try to create a shared strategy. Here's what the document looks like today, and we welcome your comments, questions and feedback as we work towards version 1.0.
****

DRAFT
A 4 star classification-scheme for linked open cultural metadata
Publishing openly licensed data on the Web and contributing to the Linked Open Data ecosystem can have a number of benefits for libraries, archives and museums.
1. Driving users to your online content (e.g., by improved search engine optimization);
2. Enabling new scholarship that can only be done with open data;
3. Allowing the creation of new services for discovery;
4. Stimulating collaboration in the library, archives and museums world and beyond.
In order to achieve these benefits libraries, museums and archives are faced with decisions about releasing their metadata under various open terms. To be open and useful as linked data requires deliberate design choices and systems must be built from the beginning with openness and utility in mind. To be useful for third parties, all metadata made available online must be published under a clear rights statement.
This 4-star classification system arranges those rights statements (e.g. licenses or waivers) that comply with the relevant conditions (2-11) of the open knowledge definition (version 1.1) by order of openness and usefulness: the more stars the more open and easier the metadata is to used in a linked data context. Libraries, archives and museums wanting to contribute to the Linked Open Data ecosystem should strive to make their metadata available under the most open instrument that they are comfortable with that maximizes the data's usefulness to the community..
Note: This system assumes that libraries, archives and museums have the required rights over the metadata to make it available under the waivers and licenses listed below. If the metadata you want to make available includes external data (for example vocabularies) you may be constrained by contract or copyright to release the data under one of the licenses below.

Public Domain (CC0 / ODC PDDL / Public Domain Mark)
as a user:
● metadata can be used by anyone for any purpose
● permission to use the metadata is not contingent on anything
● metadata can be combined with any other metadata set (including closed metadata sets)
as a provider:
● you are waiving all rights over your metadata so it can be most easily reused

- you can specify whether and how you would like acknowledgement (attribution or citation, and by what mechanism) from users of your metadata, but it will not be legally binding

This option is considered best since it requires the least action by the user to reuse the data, and to link or integrate the data with other data. It supports the creation of new services by both non-commercial and commercial parties (e.g. search engines), encourages innovation, and maximizes the value of the library, archive or museum's investment in creating the metadata.

Attribution License (CC-BY / ODC-BY) when the licensor considers linkbacks to meet the attribution requirement
as a user:
- metadata can be used by anyone for any purpose
- permission to use the metadata is contingent on providing attribution by linkback to the data source
- metadata can be combined with any other metadata set, including closed metadata sets, as long as the attribution link is retained
as a provider:
- you get attribution whenever your data is used

This option meets the definition of openness, but constrains the user of the data by requiring them to provide attribution (in the legal sense, which is not the same as citation in the scholarly sense). Here, attribution is satisfied by a simple, standard Web mechanism from the new data product or service. By using standard practice such as a linkback, attribution is satisfied without requiring the user to discover which attribution method is required and how to implement it for each dataset reused. Note that there are other methods of satisfying a legal attribution requirement (see below) but here we propose a specific mechanism that would minimize the effort needed to use the data if the LAM community collectively agrees to it. Also note that even this simple (ideally shared) attribution method could prevent some applications of linked data if linkbacks are required by many datasets from many sources.

Attribution License (CC-BY / ODC-BY) with another form of attribution
as a user:
- metadata can be used by anyone for any purpose
- permission to use the metadata is contingent on providing attribution in a way specified by the provider
- metadata can be combined with any other metadata set (including closed metadata sets)
as a data provider:
- you get attribution whenever your data is used by the method you specify

This option meets the definition of openness in the same way as the linkback attribution open, but requires the user to provide attribution is some way other than a linkback, as specified by the data provider. The provider could specify an equally simple mechanism (e.g. by retention of another field, such as 'creator' from the original metadata record) or by a more complex mechanism (e.g. a scholarly citation in a Web page connected to the new data product or service). The disadvantage of this option is that the user must discover what mechanism is wanted by the particular data provider and how to comply with it, potentially needing a different mechanism for each dataset reused. For large-scale open data integration (e.g. mashups) this option is difficult to implement.

Attribution Share-Alike License (CC-BY-SA/ODC-ODbL)
as a user:
- metadata can be used by anyone for any purpose

- permission to use the metadata is contingent on providing attribution in a way specified by the provider
- metadata can only be combined with data that allows re-distributions under the terms of this license

as a provider:
- you get attribution whenever your data is used
- you only allow use of your data by entities that also make make their data available for open reuse under exactly the same license

This option meets the definition of openness but potentially limits reuse of data since if more than one dataset is reused and if each dataset has an associated Share-Alike license. Under an Share-Alike license, the only way to legally combine two datasets is if they share exactly the same SA license, since most SA licenses require that reused data be redistributed under exactly same license. If the source datasets had different Share-Alike licenses originally (e.g. CC-BY-SA and ODC-ODbl) then there is no way for the user to comply with the requirements of both source data licenses so this option only allows users to link or integrate data distributed under one particular SA license (or one SA license and any of the other license or waiver options above). In the LAM domain, where significant value is created by combining datasets, the Share-Alike license requirement severely reduces the utility of a dataset.

Related Material
- Open Knowledge Definition (http://www.opendefinition.org/okd)
- Principles on Open Bibliographic Data (http://openbiblio.net/principles)
- Discovery Open Metadata Principles (http://discovery.ac.uk/businesscase/principles/)
- Linked Open Data star scheme by example (http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/)

(http://lod-lam.net/summit/2011/06/06/proposed-a-4-star-classification-scheme-for-linked-open-cultural-metadata/#comment-2269)

Appendix F: Sampling of live resourcing on Twitter during the first day

> Excellent basic tutorial on publishing LOD. bit.ly/joTaZI #lodlam

🐦 8 months ago   rubinsztajn

> Please take a look at digital-conferences-calend... for listing of digital humanities/tech conferences & meetups #lodlam

🐦 8 months ago   p_sully

> weak ties, linked data, and citation. quick thoughts from #lodlam day 1. bit.ly/lSwodu

🐦 8 months ago   wilbanks