

### Streszczenie

*W czwartek 21 stycznia 2016 braliśmy udział w dorocznej konferencji METRO - Metropolitan New York Library Council - która miała miejsce w Baruch College w Manhattanie. Konferencja ta, jak i poprzednie, była doskonałym przeglądem najnowszych inicjatyw, pomysłów, rozwiązań i projektów w dziedzinie humanistyki cyfrowej w społeczności GLAM. Poniżej przedstawiamy omówienie wybranych prezentacji w języku angielskim.*



The annual METRO (Metropolitan New York Library Council) conferences are about the best sources of the latest inventions, projects and ideas in the GLAM community, concentrated in one day of intense briefings. This year was no exception - the conference that took place January 21, 2016 at the Baruch College in Manhattan. On the conference a number of "Project briefings" were presented - the intent was to show the projects in progress and discuss their workings, issues and plans, not necessarily the completed works. It was impossible to attend so many parallel briefings; we have selected two in each sessions, and report on them here as a sampling of the conference.

The full [schedule of the conference](#) is available, as well as a listing of the [project briefings](#) .

### Keynote

The inaugurating keynote was delivered by Kari Lämsä, the manager of Library Number 10 and the Urban Workshop in Helsinki. He presented a fascinating case of modern library and a

testbed for a library transformation to 21 century. [ [Presentation slides](#) ]

The author started with an observation, that only some 25% of the patrons check-out materials from the library and the library events are now organized in 80% to 90% by its customers. He then described the transformations, beginning with the furniture - traditionally uncomfortable library chairs to various places - some flexible, some comfortable, some creating a tranquil space, all movable by the customers from one place to another. Another change was to treat paperback editions as consumables - no reservations, no fixed order on shelves, read until fall apart.

The big change compared to traditional library was a creation of different spaces, for example:

- An e-Lounge to read/view/listen to e-books, e-audio, e-magazines, e-newspapers, e-movies.
- An urban office for small business owners, with bookable soundproof screens for work or small group meetings.
- A Maker Space, where customers select pilot projects and staff helps them do it by themselves, equipped with 3D printer, 3D scanner, vinyl cutter, 3D cutter, sewing machines, digitizing station and more.

The experience gained in implementing those projects has been utilized in the design of the new, [Helsinki Central Library](#), which is under construction now.

## **Documenting Art Collections in Gilded Age New York: An online exhibition using Google Open Gallery**

Victoria Pilato from Frick Art Reference Library presented current digital project called Phase IV of Documenting the Gilded Age series using Google Open Gallery. This project was possible due to grant that allowed the Frick Art Reference Library to digitize 80 art collection catalogues and some archival materials. Google Open Gallery was chosen to showcase images over text. [

[Presentation slides](#)

]

This presentation discussed usability of Google Open Gallery[1], its limits and possibilities. Overall, the Google Open Gallery is simple to use but rather time consuming, uses minimal

captions and metadata. The result is good and can be checked at the [project website](#) .

## **Making the Leap Towards Linked Data**

The American Museum of Natural History[2] has a vast collection, of which only a small fraction is exhibited and any one time. It includes 3.5 million artifacts and specimens and 7.6 linear kilometers of archives, organized in five divisions (Anthropology; Paleontology; Invertebrate and Vertebrate Zoology; Physical Science) plus a research library. The museum also sport close to a petabyte (thousand terabytes) of digital data, in form of archival finding aids, library catalogs and digitized collections.

In the presentation “Making the Leap Towards Linked Data” Museum’s Iris Lee and Jennifer Cwiok together with CodedCulture consultant Nick Krabbenhoeft introduced us to the complexities of a huge project which final goal is to create interconnections between the data and metadata and make them widely accessible. [ [Presentation slides](#) ]

Each of the divisions has some system of storing information and accessing it, mostly home grown. The archival data is stored in EAD[3] format, but also the EAC[4] (Encoded Archival Context) is used, as well as text documents and spreadsheets. DSpace[5] is in usage for the Museum scientific papers, Omeka[6] for web publishing, the archival data in Archivist’ Toolkit and eventually in Archives Space[7], xEAC[8] is the tool to store the EAC data, while proprietary OPAC keeps the online catalog and the digitized collection of typed specimen and other catalog cards is kept on another proprietary Central Archives application.

The presenters have sketched out the path to integrate the data access. Some applications would be moved to a common shared tool, for example the archival data would go to Archives Space (which can relatively easily import EAD files from Archivist’s Toolkit). Some will stay in place, maintained as a source, keeping the departmental systems mostly intact. The team then plans to convert the data to RDF[9] triplets and use the Linked Open Data[10] tools to provide both common input and common access to integrated data. The work is in progress; the development team has the [Hidden Collections](#) blog where the issues are discussed and where the progress of the works can be monitored

## **Beyond Citation: Critical Thinking About Digital Research**

Most databases of research papers are closed, proprietary systems. It means not only that not all libraries can afford them, it also means that the selection, content bias and completeness are

usually opaque. [Beyond Citation](#) attempts to remedy those issues, at least partially, by providing a single entry point to many popular databases. Each database has a Reviews and Conversations section - links to articles and blog posts related to the database, Overview, Info from Publisher, History/Provenance and, if database is by subscription only, the Access will show the nearest library with that resource

In their presentation Eileen Clancy and Polly Thistlethwaite from the CUNY Graduate Center guided us through some of the difficulties libraries have in providing their patrons with access to journals and research databases, and introduced the Beyond Citation website, a recommended source for researchers. [ [Presentation slides](#) ]

### **Forming the Digital Art History Lab at the Frick Art Reference Library**

Ralph Baylor presented activities of the Frick Art Reference Library staff members to organize the Digital Art History Lab ( [DAHL](#) ) in order to create and promote using new digital tools addressing the specific needs of art-historical research. DAHL is a committee composed of staff not only from the Frick Art Reference Library but also from different departments: The Center for the History of Collecting, Photoarchive, Public Services, and Book Departments. [

[Presentation slides](#)

]

This presentations outlined few new digital art tools which are now being tested and as well underlined: outreach, collaborations, and software development.

### **We put All Our Data in One Basket, and You Won't Believe What Happened Next! (Linked Data and Public Works at NYPL)**

Matt Miller and Shawn Averkamp from the New York Public Library Labs begin with the quote from the design goals of RDF (Resource Description Framework): "2.2.6 Anyone can say anything about anything", and ask themselves a question: "What would happen if anyone could say anything about any library resource?" The rest of the presentation is an attempt to answer this question and show, how the advanced team in the NYPL Labs approaches the topic. [ [Presentation slides](#)

].

The New York Library, private and independent nonprofit corporation, has some 53 million items and is the second largest public library in the United States and fourth largest in the world. Its

digital collection has some 700,000 items available for study and reuse. The [NYPL Labs](#) is working on various ways of making the Library offerings (especially the digital one) more attractive and informative to the users. After opening the comment section in the digital archives, the logical next step was to think how to use those comments to enrich the presentation. Matt and Shawn have shown us how they plan to use RDF to collect and present the data, and how to integrate it with the existing databases. We have learned about Authority Cleanup, Mongo Database[11], Portland Common Data Model[12] and other components of the project. The pilot project is in progress, the technical details are available on a “[Registry](#)” website.

### **Supporting Critical Practice through collaboration: Public Libraries and LIS Schools as Instructional Partners**

Professor Jessica Hochman of Pratt Institute’s School of Information and Brendan Murphy, Education Coordinator at Brooklyn Connections talked about collaboration between LIS (Library and Information Studies) schools and Brooklyn Public Library. This project focuses on using skills and educating Pratt’s students in library instruction on how to work with mentors from the Brooklyn Public Library in teaching information literacy and critical thinking skills to 4th to 12th grade classes from Brooklyn public schools. [ [Presentation slides](#) ]

School children are using original resources from Brooklyn public library in special topic designed by MSLIS students learning history of Brooklyn. It is a great project with 2000 school children participating.

### **Beyond Boutique: The Shift from Selective to Mass Digitization**

Kevin Schlottmann from Center for Jewish History begun the presentation by contrasting selective digitization (- loss of context, - reinforcement of bias, - who does it serve? etc.) with mass digitization, which mostly avoids those problems. Going further, he discussed the concept of “Radical Digitization” - scan everything, convert to text by OCR[13], do not bother with metadata. [ [Presentation slides](#) ]

Following Kevin, collaborators from different organizations talked about their approaches and issues encountered when trying to broaden the scope of digitization.

Greg Wiedemann from SUNY Albany talked about mass digitization on demand, issues with

metadata and Python tools developed to ease the process - EAD Validator, AutoUpload etc.

W. Jordan Patty from George Mason University spoke about the use of OpenRefine, a formerly Google project transformed into an Open Source tool to transform and correct messy data.

Eric Fritzler from the American Jewish Historical Society presented ways to refine and automate the slow ingest of metadata in existing systems (with MARCXML[14], METS[15] and CSV[16] data) to allow the process to be completed in reasonable timeframe.

Lea Lange from Center for Jewish History has shown us how the use of simple Excell macros can radically simplify and speed-up the transformation of metadata.

### **National Digital Stewardship Residency Project Updates**

The National Digital Stewardship Residency in New York City currently has five residents working in different cultural heritage institutions to develop, and implement digital and preservation projects. [ [Presentation slides](#) ]

We had a chance to listen to:

- Dinah Handel from CUNY Television
- Carmel Curtis from Brooklyn Academy of Music
- Mary Kidd from New York Public Radio
- Genevieve Havemeyer-King from Wildlife Conservation Society
- Morgan McKeehan from Rhizome

All updates were very informative, focusing on digital projects in process and challenges encountered. You can read more about the residents in the [National Digital Stewardship Residency](#) website.

**Marek Zieliński i Iwona Korga**

Artykuł ukazał się 25 stycznia 2016 w *Blogu archiwistów i bibliotekarzy Instytutu Piłsudskiego*

### Footnotes

- [1] Google Open Gallery <https://www.google.com/opengallery/>
- [2] Museum website <http://www.amnh.org/>
- [3] Encoded Archival Description <https://www.loc.gov/ead/>
- [4] Encoded Archival Context - Corporate Bodies, Persons and Families <http://eac.staatsbibliothek-berlin.de/>
- [5] DSpace - software for digital repositories <http://www.dspace.org/>
- [6] Omeka <http://omeka.org/>
- [7] Archives Space archival software <http://www.archivesspace.org>
- [8] Presentation about xEAC <http://www.oclc.org/research/events/2014/01-09.html>
- [9] Resource Description Framework - Wikipedia article [https://en.wikipedia.org/wiki/Resource\\_Description\\_Framework](https://en.wikipedia.org/wiki/Resource_Description_Framework)
- [10] Linked Data and Linked Open Data - Wikipedia article [https://en.wikipedia.org/wiki/Linked\\_data](https://en.wikipedia.org/wiki/Linked_data)
- [11] Mongo database - Wikipedia article <https://en.wikipedia.org/wiki/MongoDB>
- [12] Portland Common Data Model <https://github.com/duraspace/pcdm/wiki>
- [13] Optical Character Recognition - Wikipedia article [https://en.wikipedia.org/wiki/Optical\\_character\\_recognition](https://en.wikipedia.org/wiki/Optical_character_recognition)
- [14] An XML flavor of MARC standard - Wikipedia article [https://en.wikipedia.org/wiki/MARC\\_standards](https://en.wikipedia.org/wiki/MARC_standards)
- [15] Metadata Encoding and Transmission Standard - Wikipedia Article [https://en.wikipedia.org/wiki/Metadata\\_Encoding\\_and\\_Transmission\\_Standard](https://en.wikipedia.org/wiki/Metadata_Encoding_and_Transmission_Standard)
- [16] Comma Separated Values - Wikipedia Article [https://en.wikipedia.org/wiki/Comma-separated\\_values](https://en.wikipedia.org/wiki/Comma-separated_values)

### Explore more blog posts

- [Humanistyka Cyfrowa w New York City](#)
- [Wiki-konferencja USA 2014](#)
- [Doroczna konferencja METRO 2014](#)