

Linked Data Summit - Interoperability of Library Data

November 7-8, 2022
Library of Congress

Summary

Promoted by the BIBFRAME Interoperability Group (BIG), Library of Congress (LC), Linked Data for Production (LD4P), OCLC, Share-VDE (SVDE)

Libraries have long been dependent on the exchange and reuse of their metadata as a means of providing discovery and access to an explosion of resources. As stated by Sarah Thomas on the formation of the Program for Cooperative Cataloging, [our goal] *is to increase the timely availability of bibliographic and authority records by cataloging more items, by producing cataloging that is widely available for sharing and use by others, and by cataloging in a more cost-effective manner.* With the substitution of “data” for “records,” the same goals are true in an open, linked data environment.

The primary focus of this summit was the interoperability of library data between libraries and their service providers in order to enhance the exchange and reuse of library data. The interoperability of library data with other linked data providers outside the community is equally as important to resolve, but was not the thrust of this meeting. It is hoped that a consistent approach to the interoperability of library data between libraries and their service providers will provide a smoother path for the integration of library data with data from other communities as well.

This meeting was preceded by two earlier meetings held in 2021, the Linked Data Summit (Library of Congress, OCLC, Share-VDE and Stanford University Libraries) in August and the [PCC BIBFRAME Data Exchange Meeting](#) in September. A follow up meeting is currently planned for Fall 2023.

Agenda and Presentations

Day 1

Topic 1: Data Interoperability - Discussion on the interoperability of ‘data’ (defined as entities, persistent identifiers, etc) that are required for the creation, maintenance, consumption and for the discovery of library resources across the networked library community and among the involved data pools.

Presentations:

- [BIG presentation](#): Overview, Current Work and Forecasting (Melanie Wacker and Ian Bigelow)
- [Report from Exchange Shapes/Strawperson Group](#) (Nathan Putnam)

Topic 2: Model Interoperability - Discussion on the interoperability of data models, specifically the various 'flavors' of BIBFRAME that exist today. While there are additional data models like IFLA LRM and RDA to consider, these will not be the main focus.

Presentations:

SVDE-SEI Working Group referring also to the Opus and Hub parallel developments (Ian Bigelow and Oddrun Ohren):

- [Query_ Link or Copy - The Opera and Interoperability - 2022-11-07.pdf](#)
- [Model interoperability.pdf](#)

Day 2

Topic 3: Record Interoperability - Discussion on the interoperability of BIBFRAME "records". A record (i.e. dataset) is defined as all of the elements (i.e. set of triples that compose an instance) used to describe a specific library resource within BIBFRAME.

Presentation:

SVDE and BIBFRAME interchange structure (Jim Hahn UPenn in-person presentation, also delivered at BFEW 2022 [hahn.pdf](#) [Video](#))

Topic 4: Discovery - Comparison of existing entity discovery platforms. Review and discussion of preliminary further use cases for discovery of library resources in a linked data library environment.

Presentations:

- [Readout of linked data use cases](#) & [Miro Board Exercises](#) from Discovery working group session
- SVDE entity discovery interface [live demo + introductory slides](#) (Anna Lionetti)

Outcomes and Next Steps

Meeting Threads

1. Focus on prototype and testing (e.g., production use cases).
2. Focus on validation of data both on import and export.
3. Focus on BF data aggregation and what variations can co-exist.
4. Focus on openness and trust.
5. Focus on discovery and the end user more broadly.

Proposals for Concrete next Steps (due at least by the next meeting)

1. Move SHACL validation into SINOPIA Stage environment.
2. Exploration of Shapes in the validation of PCC BF data.
3. Reduce the time lag between NAR creation to validation/availability for non-LC NACO members from three days to one day.
4. Devise a test environment that blends Hubs and Opus to see what the effect will be.
5. Develop a set of canonical test tools to show the behaviors associated with linked data that people / projects might produce.
6. Document the differences between the Hub and Opus to help catalogers.
7. Aggregate BF data and analyze the impact of variations on data coherence and discovery
8. Report on practical issues/decisions needed for production implementation. (Those institutions going into production in the next 12 months)
9. Create a small group composed of BF implementers to discuss parameters for setting up community control of a common flavor of BF. Also, determine if this should be the interchange model or merely standard BF (Meeting organizers will form the group)
10. Evaluate trust and the use of the PCC Data Pool. Specifically, what sources of linked open data are acceptable.
11. Create a demo of the discovery of digital materials
12. Further exploration of the interchange model
13. Validate the use cases from the Use Case Group in production. (Those institutions going into production in the next 12 months)
14. Create a list of definitions for the terms we use
Examples:
 - a. Trust - + possible checklist
 - b. Openness
 - c. Copy Cataloging vs. Sharing
 - d. BIBFRAME record
 - e. BIBFRAME Cataloging (or new term)
 - f. Significantly (i.e. an implementation differs significantly from another)

- g. Substantial
- h. Implementation
- i. Production

15. Define a common set of criteria to determine if data meets our trust threshold.

16. Share the Discovery use cases and determine how we can measure success

Background reading

- [Use Case Working Group Final Report 2021 December.pdf](#)
- [Data Exchange Working Group Final Report 2022 March.pdf](#)
- [BIBFRAME Workshop in Europe 2022](#)
- [Control or Chaos: Embracing Change and Harnessing Innovation in an Ecosystem of Shared Bibliographic Data](#)