

LD4P2 Project Background and Goals

Linked Data for Production: Pathway to Implementation (LD4P Phase 2) builds upon the foundational work of [Linked Data for Production \(LD4P\) Phase 1](#) and [Linked Data for Libraries Labs \(LD4L Labs\)](#), to begin the implementation phase of the cataloging community's shift to linked data for the creation and manipulation of their metadata.

A collaborative project among four institutions (Cornell, Harvard, Stanford, and the University of Iowa School of Library and Information Science) and the Library of Congress and the Program for Cooperative Cataloging (PCC), this phase of LD4P will have seven goals:

- the creation of a continuously fed pool of linked data expressed in BIBFRAME from a core group of academic libraries
- development of a cloud-based sandbox editing environment in support of an expanded cohort of libraries to create and reuse linked data
- the development of policies, techniques and workflows for the automated enhancement of MARC data with identifiers to make its conversion to linked data as clean as possible
- the development of policies, techniques, and workflows for the creation and reuse of linked data and its supporting identifiers as libraries' core metadata
- better integration of library metadata and identifiers with the Web through collaboration with Wikidata
- the enhancement of a widely-adopted library discovery environment (Blacklight) with linked-data based discovery techniques
- the orchestration of continued community collaboration through the development of an organizational framework called LD4, ensuring continued exchange of ideas and techniques across a distributed developing community.

Collaboration will be key in this phase of LD4P, both internal and external. The partners will be collaborating on the development of the cloud environment and Blacklight, the Library of Congress, and the Program for Cooperative Cataloging will collaborate with the project through training in the use of the BIBFRAME Editor, Harvard will foster a close relationship with the PCC in the development of policy decisions, and the core institutions will collaborate with Wikidata in the publishing, linking, and enriching linked data through the Wikimedian-In-Residence program.

In Linked Data for Production Phase 1, the partners proposed the development of a communal work environment based in linked data; the strengthening and expansion of the BIBFRAME ontology to cover the multiple formats (e.g., books, music, maps, etc.) that libraries must catalog; the tools needed to perform the work itself; and the development of lightweight workflows (Tracer Bullets) to prove that the transition to linked data was both possible and practical. In the LD4L-Labs work, the partners piloted the development of an editing tool to support cataloging using BIBFRAME and variations, together with selected extension ontologies; the integration of linked data authority lookup and management into cataloging; and the use of linked data for discovery and visualization.

The focus of the second phase of LD4P is implementation. Building upon the expertise, structure, and workflows developed during the first phase of LD4P and LD4L-Labs, the four partners (Cornell, Harvard, Stanford, and the University of Iowa School of Library and Information Science) will implement a prototype environment, from metadata acquisition/creation through to discovery. An important enhancement in this phase will be collaborating with the Program for Cooperative Cataloging (PCC) and the Library of Congress to expand the number of libraries moving to implementation of linked data. Sub-grants for committed libraries will help them defray transition costs.

The choice of working with the PCC was deliberate. Within the United States, we work within the concept of a virtual, distributed "national library" for the creation of high-level metadata. The PCC provides the community with a forum for the development of policy and training programs for member libraries. The full buy-in of the PCC, along with their ability to provide training and support, will be key to expanding the transition to linked data from the core libraries within LD4P to the broader academic library community.

Discovery will also be a key development in LD4P Phase 2. By focusing on linked-data enhancements to current discovery systems such as Blacklight, LD4P hopes to take immediate advantage of linked data for library patrons through such developments as the addition of knowledge panels, authority-based browse, and semantic search.