Related Projects

Go to LD4L Wiki Gateway

Archived

LD4L 2014, which was the Linked Data for Libraries original grant running from 2014-2016, has been completed. This page is part of the archive for that grant.

Major Related Projects

BIBFRAME

The Bibliographic Framework Initiative (BIBFRAME) is an undertaking by the Library of Congress and the community to better accommodate future needs of the library community. A major focus of the initiative will be to determine a transition path for the MARC 21 exchange format to more Web based, Linked Data standards. Zepheira and The Library of Congress are working together to develop a Linked Data model, vocabulary and enabling tools / services for supporting this Initiative.

The BIBFRAME Initiative is the foundation for the future of bibliographic description that happens on the web and in the networked world. It is designed to integrate with and engage in the wider information community and still serve the very specific needs of libraries. The BIBFRAME Initiative will bring new ways to: 1) differentiate clearly between conceptual content and its physical/digital manifestation(s); 2) unambiguously identify information entities (e.g., authorities); and 3) leverage and expose relationships between and among entities.

In a web-scale world, it is imperative to be able to cite library data in a way that differentiates the conceptual work (a title and author) from the physical details about that work's manifestation (page numbers, whether it has illustrations). It is equally important to produce library data so that it clearly identifies entities involved in the creation of a resource (authors, publishers) and the concepts (subjects) associated with a resource.

Although the BIBFRAME Initiative will instantiate a new way to represent and exchange bibliographic data – that is, replace the Machine Readable Cataloging (MARC) format – its scope is broader. As an initiative, it is investigating all aspects of bibliographic description, data creation, and data exchange. In addition to replacing the MARC format, this includes accommodating different content models and cataloging rules, exploring new methods of data entry, and evaluating current exchange protocols. http://www.loc.gov/bibframe/

Blacklight

Blacklight is an open source Ruby on Rails gem that provides a discovery interface for any Solr index. Blacklight provides a default user interface which is customizable via the standard Rails (templating) mechanisms. Blacklight accommodates heterogeneous data, allowing different information displays for different types of objects. Blacklight uses Apache Solr, an enterprise-scale index for its search engine. Blacklight features faceted browsing, relevance based searching (with the ability to locally control the relevancy algorithms), bookmarkable items, permanent URLs for every item, user tagging of items. Bl acklight is a component of the Hydra Project framework (see below) http://projectblacklight.org

Fedora Commons Repository Software

Fedora (Flexible Extensible Digital Object Repository Architecture) was originally developed by researchers at Cornell University as an architecture for storing, managing, and accessing digital content in the form of digital objects inspired by the Kahn and Wilensky Framework. Fedora defines a set of abstractions for expressing digital objects, asserting relationships among digital objects, and linking "behaviors" (i.e., services) to digital objects.

The Fedora Repository Project (i.e., Fedora) implements the Fedora abstractions in a robust open source software system. Fedora provides a core repository service (exposed as web-based services with well-defined APIs). In addition, Fedora provides an array of supporting services and applications including search, OAI-PMH, messaging, administrative clients, and more. Fedora provides RDF support and the repository software is integrated with semantic triple store technology, including the Mulgara RDF database. Fedora helps ensure that digital content is durable by providing features that support digital preservation.

The Fedora Commons refers to the community surrounding the Fedora Repository Project. This community joins together with common needs, use cases, and projects. The Fedora Commons community is very active in producing additional tools, applications, and utilities that augment the Fedora repository. Many of these creations are available to the entire community as open source.

The Fedora Repository software has been installed by institutions, worldwide, to support a variety of digital content needs. The Fedora Repository is extremely flexible and can be used to support any type of digital content. There are numerous examples of Fedora being used for digital collections, e-research, digital libraries, archives, digital preservation, institutional repositories, open access publishing, document management, digital asset management, and more. The Fedora Repository software is a component of the Hydra Project Framework (see below). http://fedora-commons.org

Hydra Project

Hydra is a repository solution that is being used by institutions on both sides of the North Atlantic to provide access to their digital content. Hydra provides a versatile and feature-rich environment for end-users and repository administrators alike. Hydra is a large, multi-institutional collaboration. The project gives like-minded institutions a mechanism to combine their individual repository development efforts into a collective solution with breadth and depth that exceeds the capacity of any individual institution to create, maintain or enhance on its own. Hydra is an ecosystem of components that lets institutions deploy robust and durable digital repositories (the body) supporting multiple “heads”: fully-featured digital asset management applications and tailored workflows. Its principle platforms are the Fedora Commons repository software, Solr, Ruby on Rails and Blacklight.

The Hydra Project was founded in 2008 by: Stanford University, University of Virginia, University of Hull, Fedora Commons (now part of DuraSpace), and quickly augmented by MediaShelf LLC. These five founding partners are still very active and are the current members of the Hydra Steering Group.

Additional partners have formally committed themselves to support and further Hydra's work: University of Notre Dame, Northwestern University, Columbia University, Penn State University, Indiana University, London School of Economics and Political Science, Rock and Roll Hall of Fame, The Royal Library of Denmark, Data Curation Experts, WGBH, Boston Public Library, Duke University, Yale University, and Virginia Tech. Further institutions are working with Hydra and its components. In time we hope that they will choose to join the formal Hydra Partners. Amongst them: Spoken Word Services (Glasgow Caledonian University), University College Dublin, University of Illinois at Urbana-Champaign, The Digital Repository of Ireland, Museum of the Performing Arts (MAE) of the Theatre Institute of Barcelona, Johns Hopkins University, Tufts University, and Cornell University. http://projecthydra.org

ORCID

ORCID is an open, non-profit, community-driven effort to create and maintain a registry of unique researcher identifiers and a transparent method of linking research activities and outputs to these identifiers. ORCID is unique in its ability to reach across disciplines, research sectors and national boundaries. It is a hub that connects researchers and research through the embedding of ORCID identifiers in key workflows, such as research profile maintenance, manuscript submissions, grant applications, and patent applications.

ORCID provides two core functions: (1) a registry to obtain a unique identifier and manage a record of activities, and (2) APIs that support system-to-system communication and authentication. ORCID makes its code available under an open source license, and will post an annual public data file under a CC0 waiver for free download.

The ORCID Registry is available free of charge to individuals, who may obtain an ORCID identifier, manage their record of activities, and search for others in the Registry. Organizations may become members to link their records to ORCID identifiers, to update ORCID records, to receive updates from ORCID, and to register their employees and students for ORCID identifiers. http://orcid.org

Stanford Linked Data Workshop

The Stanford University Libraries and Academic Information Resources (SULAIR) with the Council on Library and Information Resources (CLIR) conducted at week-long workshop from 27 June–1 July 2011 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. As preparation for the workshop, CLIR sponsored a survey by Jerry Persons, Chief Information Architect emeritus of SULAIR that was published originally for workshop participants as background to the workshop and is now publicly available. The original intention of the workshop was to devise a plan for such a prototype. However, such was the diversity of knowledge, experience, and views of the potential of Linked Data approaches that the workshop participants turned to two more fundamental goals: building common understanding and enthusiasm on the one hand and identifying opportunities and challenges to be confronted in the preparation of the intended prototype and its operation on the other. In pursuit of those objectives, the workshop participants produced:

1. a value statement addressing the question of why a Linked Data approach is worth prototyping;
2. a manifesto for Linked Libraries (and Museums and Archives and …);
3. an outline of the phases in a life cycle of Linked Data approaches;
4. a prioritized list of known issues in generating, harvesting & using Linked Data;
5. a workflow with notes for converting library bibliographic records and other academic metadata to URIs;
6. examples of potential "killer apps" using Linked Data: and
7. a list of next steps and potential projects.

A full report on the workshop can be found at http://lib.stanford.edu/files/Stanford_Linked_Data_Workshop_Report_FINAL.pdf

VIVO

VIVO is an open community, an information model, and an open source semantic web application supporting the advancement of research and scholarship by integrating and sharing information about researchers and scholars, their activities, and their outputs both within a single institution and across broad, distributed networks. VIVO is fundamentally interdisciplinary; it enables and promotes the discovery of research and scholarship across traditional boundaries of geography, organization structure and type, academic or clinical or applied domain, technology, language, and culture. There is a diverse set of activities associated with the VIVO project, across federal agencies, academic institutions, professional societies, and data providers, as well as a variety of efforts with the Semantic Web and ontology development communities. VIVO was originally funded by Cornell University and the National Institutes of Health (U24 RR029822) and is currently a community-supported incubator project under the DuraSpace umbrella. http://vivoweb.org

Other Related Projects

In addition to the related efforts we have already mentioned, we are very interested in working with and integrating other Linked Data efforts and existing identifier and taxonomy efforts. Cornell is a contributor to the Social Networks and Archival Context (SNAC) effort [1], and this is exactly the kind of additional contextual value that we would like to bring to information resources. We would also like to look at opportunities to leverage other EAD and finding aid information to create structured context for information resources, for example by linking to historical events. The VIVO/Vitro framework that we will be using also supports linkage to standard external vocabularies (e.g., UMLS [2], AGROVOC [3], NALT [4], and LCSH [5]) and authority services such as LC Name Authority File [6] and VIAF [7]. We are tracking NISO Bibliographic Roadmap [8] process, and we will seek to engage with that effort in any way that is appropriate.

[1] http://socialarchive.iath.virginia.edu/
[2] http://link.informatics.stonybrook.edu/umls/
[3] http://aims.fao.org/standards/agrovoc/linked-open-data
[4] http://www.nal.usda.gov/news/NALT_LOD.shtml
[5] http://id.loc.gov/authorities/subjects.html
[6] http://id.loc.gov/authorities/names.html
[7] http://viaf.org
[8] http://www.niso.org/topics/tl/BibliographicRoadmap/