



Linked data

The future of library cataloging

A new way to advance libraries

As the volume and variety of online information—both valuable and suspect—increases, good metadata is more important than ever. Our communities and institutions need people with the knowledge and expertise to connect library resources to the wider knowledge ecosystem.

Linked data is a way to organize and connect data on the web so it can be easily, automatically, and programmatically shared and used by various systems and services. Linked data breaks up the valuable, library-focused data locked in MARC records and publishes it using uniform resource identifiers (URIs).

Library metadata is very good for “known item” searching and finding works based on subjects, titles, and authors. But finding unexpected and potentially valuable connections takes time, and often requires advanced knowledge. It’s not always easy to find good, new connections that can lead to powerful insights.

Imagine a student investigating a specific subject for their graduate work. They find a thesis that excites them and informs their research. Today, their information search might extend to other works written by that author or materials cataloged under related subjects. **But what about pieces by co-authors? Articles in other databases or published online? Or articles put out by the same publisher? Or from the same university? What about topics that are related more loosely by contextual topics, time periods, locations, or organizations that aren’t, strictly speaking, within the same subject heading?**

With linked data, libraries can pave new paths—both within and outside traditional library platforms—that lead users to unexpected discoveries. And it can bring people from other services and systems back to the unique, important materials that are often available only in libraries.

From enhanced, library-focused discovery services to better visibility across the web, linked data will improve results for libraries and their users.

Library work at a point of inflection

Libraries dedicate a lot of training, time, and resources to the description of their materials. Today, that metadata is created and managed primarily in MARC records and other fixed-format data resources, making it difficult to use in other environments, including across the web. Similarly, outside data that could potentially be useful to libraries and their users is hard to get into library systems.

Imagine that your library wants to support the culture and heritage of a local community. Today, you could create a special library landing page featuring that work, add links to resource lists, or manually promote the materials to specific, interested people. But with linked data, you can automate the creation of knowledge cards that connect to outside authorities, services, data hubs, and content. Your library can build and improve a single source that showcases all the great innovations and discoveries that come from the communities you support, helping your library attract attention, exert influence, and extend reach.

Evolving library data into linked data frees the knowledge in library collections and connects it to knowledge streams that inform our everyday lives—on the web, through smart devices, and while using technologies like artificial intelligence (AI). And that’s a two-way street, because library data, while obviously vital for libraries, also provides value for other knowledge work like research, education, and cultural enrichment.

Linked data enables library staff to provide greater context for information and build rich connections across library resources, their communities, and beyond.

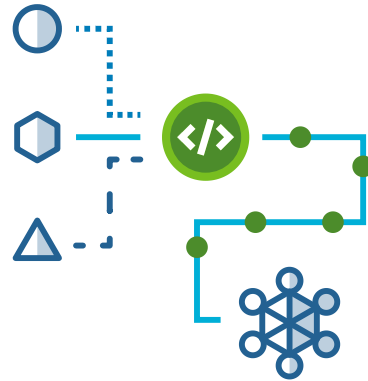
The move to linked data facilitates greater contextualization of information, makes resources more discoverable, provides opportunities to improve workflows, and supports easier integration of library data into other systems and services.

How does linked data benefit libraries?

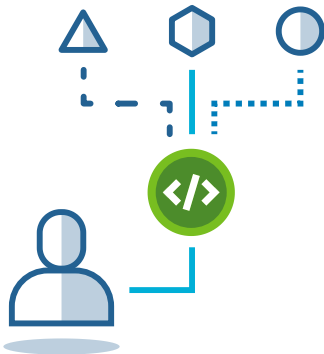
As an “industry-agnostic” metadata technology, linked data provides advantages that are difficult, time-consuming, and expensive to achieve using current methods.



Linked data connects and scales library knowledge. This helps libraries showcase their value by connecting local resources to wider information streams.



Linked data fuels serendipitous discovery. It adds context to the description process and creates connections that can lead to unexpected discovery experiences, like a new approach to solving a complex research question.



Linked data empowers library experts to focus on high-value knowledge work. This expands librarians' role as knowledge workers, enhancing discovery while embedding the library more deeply on campus and in the community.

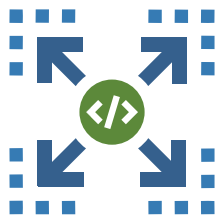


Linked data equips libraries to actively contribute to the global information ecosystem. By harmonizing library metadata with other data formats and technologies, linked data makes it consumable by machines and emerging technologies and, ultimately, available in more places.

Some of these benefits will require time. The data stored in MARC and other traditional record stores will need to move into linked data formats. But as OCLC develops the necessary services and processes to do so, some benefits can be realized today by connecting linked data to the great cataloging work that libraries have been doing for decades.

OCLC's commitment to linked data

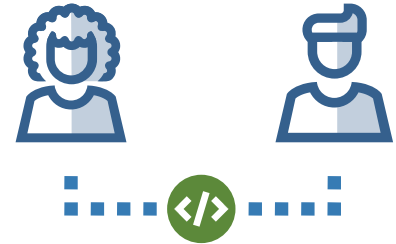
OCLC is uniquely positioned to provide the three things that are vital for a successful transition to library linked data:



Scalability: The capacity to support linked data that improves library work both at individual institutions and across massive, interconnected global workflows



Sustainability: The ability to steward this process over years and decades, with a proven track record of community building, metadata expertise, technology, and data curation



Inclusivity: A commitment and requirement to work with every type of library, archive, and museum, of any size, anywhere in the world

At OCLC, we're building the infrastructure, tools, and community-informed expertise to support libraries in this metadata evolution.

Our approach to collaborative management of library data at scale has proven successful for 50+ years. We have researched, developed, and standardized many current best practices. And we have supported the profession's ethics and values while advancing new technologies and services.



OCLC is developing linked data products, services, and datasets that connect data and provide the resources needed to curate data at scale across all library workflows.

This transformation will flourish through collective innovation and collaboration of many libraries, publishers, and library services organizations. We'll continue to partner on efforts around this shared endeavor with WorldCat® as the foundation. This unique data resource allows us to translate today's data formats into linked data, integrating their value into the workflows of thousands of libraries as opposed to one at a time.

A long-term strategy that serves the needs of all libraries

Many in our profession believe that, eventually, libraries will all need to transition to a linked data model. But today, very few libraries are ready to move to cataloging and management systems based entirely on linked data. On the one hand, some large institutions with dedicated funding are doing ground-breaking work in this area. Other libraries are experimenting with linked data in terms of specific discovery applications. But for now, the majority are just beginning to explore what linked data might mean for them.

Any replacement of MARC-based cataloging with linked data services won't happen overnight. Libraries will need to evaluate needs and resources to navigate the transition, operating in a hybrid environment for an extended period.

Our library members represent institutions of every size, type, geography, and funding. And our goal is to help each library proceed at its own pace, based on its own needs and resources. We're committed to the long-term success of this transition, and to providing ongoing community support, training, and collaboration. With WorldCat as the foundation, our long-term strategy focuses on expertise, data, infrastructure, and tools.



Expertise

We're harnessing our collective expertise. We'll continue to bring libraries, OCLC teams, and other community partners together so that library-focused metadata expertise can influence and enhance broader knowledge ecosystems.

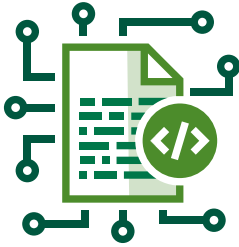
What we're doing

We're continuing and building on decades of linked data research and experiments like "Project Passage." The results of work like an international survey for linked data implementers, explorations with archives and special collections, and research on related linked data models and the transition from MARC have helped inform our current activities.

We're bringing unique teams of technologies, systems experts, researchers, community representatives, and others together in new ways to explore how this technology can create new avenues for research, discovery, visibility, and development with libraries at the center.

Benefit to libraries

OCLC's 50+ years of library data stewardship, research, and metadata capabilities are being focused on this new opportunity. Libraries can be confident that the same care and professional ethics that have guided our work for decades will continue in this area. And library leaders and workers know that as a library-focused organization that does not operate for profit, we have their best interests—and those of the people they serve—at the heart of these endeavors.



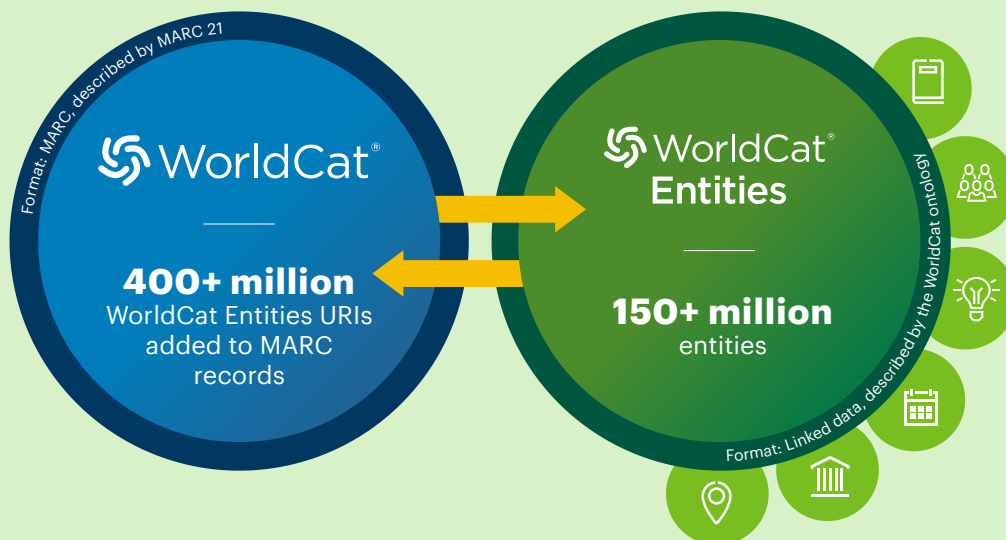
Data

We're protecting and enhancing library data at scale.

We're publishing existing library information as linked data in ways that immediately support better description and discovery of library resources across the web.

What we're doing

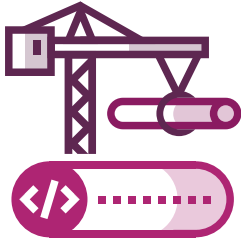
WorldCat Entities is a set of authoritative data that establishes the context for bibliographic data curation with uniform resource identifiers (URIs) for a variety of entities, including Works, Persons, Places, Organizations, Events, and more. The connections built in WorldCat Entities are semantic—like linking a work to locations it's set in. You can already explore more than 150 million WorldCat Entities at entities.oclc.org.



We're using the Dewey Decimal Classification® (DDC®) system and Faceted Application of Subject Terminology (FAST) data to create linked data URIs for established classifications and subject headings. With these URIs, libraries will be able to integrate Dewey® linked data into other data workflows at the library, enabling exploration of hierarchical and equivalence relationships among data sets and new connections that enhance discovery of library resources.

Benefit to libraries

WorldCat already represents the single largest collection of library metadata in the world. By immediately enhancing this existing resource with linked data elements, OCLC can improve library discovery without requiring all libraries to switch to a linked data model for cataloging. And ongoing linked data work will inform and improve the hundreds of millions of records and billions of holdings that we and our members have spent decades creating and cooperatively improving.



Infrastructure

We are creating the technology and systems necessary for libraries to succeed at linked data. Building on a scalable and stable foundation, we're evolving WorldCat as the sustainable infrastructure to facilitate linked data integration and interoperability at a global scale.

What we're doing

We're adding WorldCat Entities URIs to WorldCat records as \$1 links in 100, 600, 647, 651, 700, and 758 MARC fields. This enrichment establishes a bridge between MARC data and linked data, connecting data across local systems and workflows.

We researched, tested, and designed the WorldCat ontology to be a data model that provides the critical, authoritative information needed to describe library linked data. It's designed to meet library workflow needs, create more and better connections between resources, and support library visibility and partnerships.

Benefit to libraries

Linked data offers an entirely new way of creating, storing, and connecting library metadata. OCLC is integrating linked data within existing systems and services while simultaneously developing new tools that will help make the transition easier for our members.



Tools

We're integrating linked data into library workflows today while planning for tomorrow, too. We're creating a set of sophisticated new tools to add valuable linked data elements to existing records and workflows while maintaining parallel MARC services and applications for the foreseeable future.

What we're doing

We've enhanced WorldShare Record Manager with the ability to look up WorldCat Entities within existing cataloging workflows and then add linked data identifiers to records. This functionality will also be added to other OCLC services in the future.

WorldCat Entities URIs are now included in MARC records exported using OCLC cataloging tools. Individual exports of MARC records with linked data identifiers through WorldShare® Record Manager, Connexion®, the WorldCat Metadata API, and Z39.50 protocol provide record-by-record access to linked data identifiers. And bulk output is available in WorldShare® Collection Manager.

OCLC is developing a new set of tools for libraries. The first of those tools, OCLC Meridian®, allows librarians to create and edit authoritative entities, contribute to a growing dataset of entities, and connect that work meaningfully to end-user discovery and other workflows. It will connect entities to existing datasets so that semantic linkages can be made between information both in and outside the library in ways that are widely accepted, standardized, scalable, and built to support both library workflows and connections to the wider information ecosystem.

Benefit to libraries

Because OCLC provides services that already touch every part of library work, we can provide linked data services, applications, and APIs that integrate seamlessly into today's metadata workflows. OCLC will make sure that new and existing cataloging functions all work together to save library workers' time while adding exciting new opportunities.

An incredible future for libraries

Linked data provides a major new opportunity to connect the unique, powerful work done within the world's libraries to more users, more partners, more communities.

Imagine a member of your faculty comes to you looking for information on a topic related to their research. Today, your results will largely be limited to what's available in your library or through resource sharing options. But with linked data, you can make easy, logical connections from your collection to materials well outside the scope of "standard" library resources. Not only that, but you could point out industry events, related meetings, conferences, and webinars. With linked data, you can create dynamic bibliographies that are updated with new citations, links, and recommendations whenever new connections are established. And you can do it all based on criteria that your users find important.

Library workers create unique, detailed, accurate, and highly valuable metadata about information resources. Using linked data, libraries can showcase their value by connecting local resources to wider information streams—within the library field, between communities, and across the web. Cataloging experts can save time and focus on high-value knowledge work. And library leaders can work with partners to better promote academic publications, local resources, and other content that attracts positive attention, users, and grant opportunities.

From an ethical standpoint, linked data can also empower and equip libraries to actively contribute to a more trustworthy, dependable, and inclusive global information ecosystem. When library metadata is more findable, indexable, and sharable it improves the overall information ecology.

Conclusion ... and an invitation

Wherever you are in your personal understanding of linked data—and wherever your library is in terms of adoption and interest—we invite you to learn, experiment, practice, grow, and innovate with us. While OCLC research on linked data goes back more than a decade, we are just beginning to integrate it into library management tasks.

We are committed to helping libraries of all types and sizes around the world make the transition smoothly, gradually, and with specific benefits at every stage.

Our thanks

We've worked with libraries and partners every step of the way as we researched and built our linked data strategy, projects, ontology, and products. Our members' feedback has been crucial to these developments and has shaped the way we're moving forward.

Our thanks to all the OCLC members and staff who have worked on linked data research and technology projects and products, and to partner organizations who continue to be deeply involved in this transformative work, including:

ALA CORE	ORCID
BIBFRAME Interoperability Group	PCC Linked Data Advisory Committee
Bibliographic Conceptual Models Review Group	PCC Task Group on BIBFRAME-to-MARC Conversion
code4lib	PCC Task Group on MARC Simplification for BIBFRAME Conversion
Confederation of Open Access Repositories	Program for Cooperative Cataloging
Controlled Vocabularies Editorial Group	RBMS Bibliographic Standards Committee
Dublin Core Metadata Initiative	RDA Steering Committee
LD4 Steering Committee	SCT Task Group on Linked Data Training
Linked Data Interest Group	Subject Analysis Committee
Linked Data Technical Review Group	World Wide Web Consortium
MetaBelgica Project	
Metadata Interest Group	

Visit oclc.org/linkeddata

Sign up for announcements, updates, webinar registration, and opportunities to connect with OCLC staff and library colleagues around topics and projects related to linked data.

Because what is known must be shared.®



Suggested citation: OCLC. 2024. "Linked Data: The Future of Library Cataloging." Dublin, OH: OCLC. <https://doi.org/10.25333/71w4-vq71>.

© 2024 OCLC, Inc. All rights reserved. The following OCLC product, service, and business names are trademarks or service marks of OCLC, Inc.: Connexion, DDC, Dewey, Dewey Decimal Classification, Meridian, OCLC, WorldCat, WorldShare, and "Because what is known must be shared." In addition, the WorldCat, WorldShare, and OCLC symbols are service marks of OCLC, Inc. Third-party product and service names are trademarks or service marks of their respective owners. OCLC grants permission to photocopy this publication as needed.