Share Virtual Discovery Environment in Linked Data (SHARE-VDE)

Concise overview

What is SHARE-VDE?

SHARE-VDE is a library-driven initiative to establish procedures for the identification and reconciliation of entities, the conversion of data to linked data, and the creation of a virtual discovery environment based upon the three-layered structure of the BIBFRAME data model. As one of the first prototypes of a virtual discovery portal within a continually evolving workspace, SHARE-VDE aims to establish an effective environment for the use of linked data by libraries within a global context.

Who is responsible for it?

SHARE-VDE is a collaborative endeavour based upon the needs of libraries, formulated by Casalini Libri (a provider of bibliographic and authority data and a member of the Program for Cooperative Cataloging), and @Cult (a provider of Integrated Library System (ILS) and Discovery tools), with initial input from 16 North American University Libraries. The initiative is be governed by the library community.

What are its principal objectives?

To establish standards and infrastructures to "future-proof" library data, ensuring that they are compatible with the structure of linked data and the Semantic Web.

To work together with the library community in order to create a collaborative linked data environment.

To enrich library data with additional information and relationships previously unexpressed under MARC21 in order to increase the discoverability of all types of resources.

The project aims to create an environment that is useful both to library patrons and librarians. Advanced discovery interfaces will improve the user experience and deliver wider search results to library patrons, while librarians will benefit from advances such as the iterative development and support for cataloguing functions according to Semantic Web standards and the integration of these processes with local systems and tools for implementation in a collaborative environment.

SHARE-VDE also hopes to help reveal the wealth of data within existing collections, which often remain hidden or unexpressed in a traditional catalogue.

How does it work?

Library data benefit from being expressed as linked data through entity identification, reconciliation, and data enrichment. Attributes are used to uniquely identify an entity such as a Person or Work with variations in expression reconciled to form a cluster of data referring to that same entity. These entities are then reconciled against traditional authority files and other external sources to enrich the data and create a network of information and resources. A database of relationships, open to the entire SHARE community, has been compiled for the project, together with a common knowledge base of clusters that is accessible in RDF (Resource Description Framework). The database uses the model of the Semantic Web while allowing participating libraries to continue handling their own data as independently as possible.

The main areas of focus are:

- Enrichment of MARC records with URIs;
- Conversion of MARC data to RDF using the BIBFRAME vocabulary and other ontologies;

- Publication of that data following the BIBFRAME structure of Person/Work, Instance, Item for bibliographic and authority data;
- Batch/automated data updating procedures;
- Batch/automated dissemination of data to libraries;
- Progressive implementation of workflows according to the order established by the SHARE-VDE community, among these: cluster knowledge base management, copy cataloguing, original cataloguing, API (Application Programming Interface) for ILS's, retro conversion for local acquisition and administration systems based on MARC formats, and reporting.

What has been done so far?

Following discussions with participating libraries, the enrichment of existing catalogue records, their conversion into BIBFRAME, and their publication as linked data were undertaken as part of phase 1 and 2 of SHARE-VDE. During this period, tools and processes were enhanced and refined at each step of the way to allow for customizations according to the needs of each library. At the same time, over 100 million bibliographic and authority records were enriched, reconciled, and converted to linked data applying BIBFRAME and other ontologies as necessary.

What comes next?

The SHARE-VDE project is governed by discussion, analysis, and the collaborative configuration of options for the creation, enhancement, and sharing of data for all types of resources held in Libraries, Archives, and Museums (LAM).

Building on the foundations laid by the first two phases of the project, the focus of the third phase is now on the implementation of production level processes and the progressive implementation of additional workflows according to priorities determined by the SHARE-VDE community.

What are the advantages?

The advantages for libraries are many. Foremost, by implementing linked data, libraries will be able to integrate their data into the semantic web while maintaining ownership and control of their own data and benefiting from the streamlined administration of a collaborative environment and an extensive data pool. The shared platform increases discoverability of resources in library catalogues and reveals data that would otherwise have remained hidden in archives, allowing end users to access a wealth of information that may be both imported and exported by participating institutions. Usage and sharing of data are encouraged and valuable support is provided to cultural heritage collections.

Who are the participating libraries?

The institutions involved in phase 1, 2, or both phases of the project are the following (in order of State): Stanford University, University California Berkeley, Yale University, Library of Congress, University of Chicago, University of Michigan Ann Arbor, Harvard University, Massachusetts Institute of Technology, Duke University, Cornell University, Columbia University, University of Pennsylvania, Pennsylvania State University, Texas A&M University, University of Alberta, University of Toronto.

How to participate?

Institutions wishing to find out more or interested in participating in the SHARE-VDE project may contact the group at info@share-vde.org.