Share-VDE Statement: Share-VDE's Role in Library Linked Open Data

Approved by the Share-VDE Advisory Council in September 2021

Overview

The Share-VDE (Virtual Discovery Environment) project has been a leader in library linked open data and the use of BIBFRAME since the initial Share-VDE prototype in 2016. By bringing together data from many libraries in Europe and North America, Share-VDE demonstrated the power of BIBFRAME in a heterogeneous environment. Rooted in the BIBFRAME data model developed at the Library of Congress, but expanding to library data from many libraries, Share-VDE shows the power of cooperation. The member libraries have contributed their data, time and resources to develop Share-VDE, working with the development team at Casalini and @Cult.

Data model

As a BIBFRAME node in the library ecosystem, Share-VDE provides enriched data that is interoperable with other BIBFRAME nodes. Share-VDE pulls together MARC authority and bibliographic data from the member libraries, enriches it with authoritative entities and clusters data into BIBFRAME entities. The Share-VDE working groups have reviewed the clustering in detail, and expanded the BIBFRAME model to meet real-world needs and to reflect the data of the libraries engaged. This development of the data model has resulted in the Share-VDE Opus (a type of bf:Work), which groups or clusters all related expressions together and represents the original/creative work, thus facilitating the interoperability with IFLA LRM.

Linked Open Data

The Share-VDE infrastructure is based on the LOD Platform, which was designed to be capable of automating the processes of creating and publishing linked open data, regardless of the data source format. The Sapientia Cluster Knowledge Base is available in RDF (therefore as Linked Open Data) and accessible via a SPARQL endpoint and API queries.

PCC Data Pool

Based on this development work, Share-VDE was selected by the LD4P3 grant to create the PCC Data Pool. Share-VDE worked with LD4P, OCLC and the PCC to bring together all BIBCO and Conser MARC cataloging. The Share-VDE data model and clustering algorithms were applied to create an open pool of PCC-quality BIBFRAME data. The PCC data pool will be available as a trusted data source for catalogers creating natively-produced BIBFRAME using Sinopia, and also as linked open data available to any user.

Tools

In addition to other development work, the Share-VDE team is creating tools to work with the data. The Share-VDE data model anticipates that among the high volume of automated entity clusters, some clusters or relationship links will be inaccurate. The J.Cricket editor provides a means to apply direct user expertise to maintaining the Sapientia Cluster Knowledge Base. Member libraries, together with the development team, have also been exploring and recommending external data sources to be incorporated into authority dataflows that feed the Cluster Knowledge Base. Further, they are investigating new authority tools and services to align with BIBFRAME models and to expand the uses of authority data.

Discovery

Finally, Share-VDE brings a key element that is missing from many of the other BIBFRAME projects—Discovery. As the Share-VDE name indicates, Discovery has been a key focus of the project from the beginning. Share-VDE discovery is BIBFRAME based, using the entity model. Unlike record-based catalogs, Share-VDE focuses on the work and author elements. This new model avoids the long-standing problem of apparent record duplication in a MARC-based catalog, by focusing on the original work, rather than on the specific instances in an individual library. This approach goes beyond linked data enrichment (e.g. data cards) to a new way of approaching discovery. This is a huge addition to the BIBFRAME ecosystem and shows the power of linked data to improve the user experience.

Summary

Share-VDE is a BIBFRAME node, serving authoritative data that can be exchanged with libraries and with other BIBFRAME nodes, in an emerging bibliographic ecosystem. The data model and tools were developed by a strong, cooperative community. The Share-VDE initiative is a leader in this ecosystem and supports the final goal: to facilitate the reuse of enriched and structured data and to serve the research community with a new generation tool to access knowledge.