BIBFRAME Implementation in Canada

•••

The Canadian Cohort of LD4P and Share VDE

Ian Bigelow, Abigail Sparling, and Michele Casalini

BIBFRAME in context of a wider vision for moving forward with linked data at the University of Alberta Library

Cataloguing staff are not alone in their work in linked data

UAL Linked Data Projects:

- Jupiter
- CanLink
- University of Alberta Libraries' Linked Data Enrichment tool

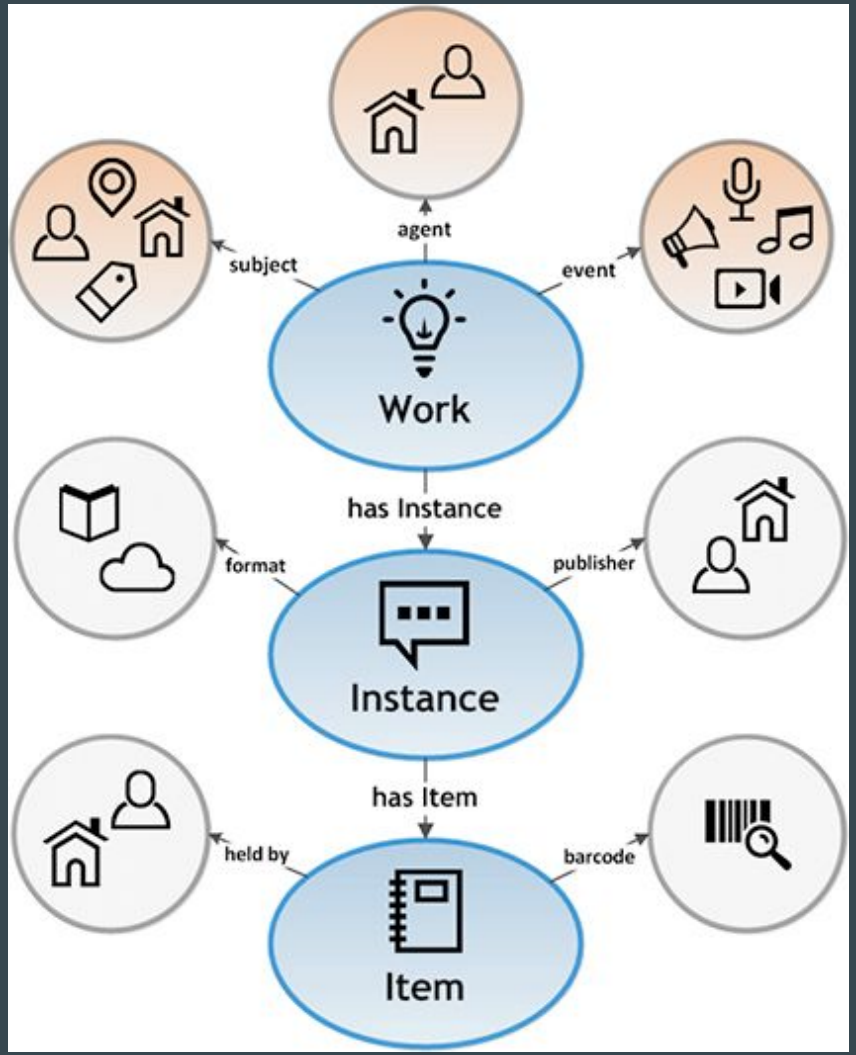
What is BIBFRAME?

...

What does it look like in practice?

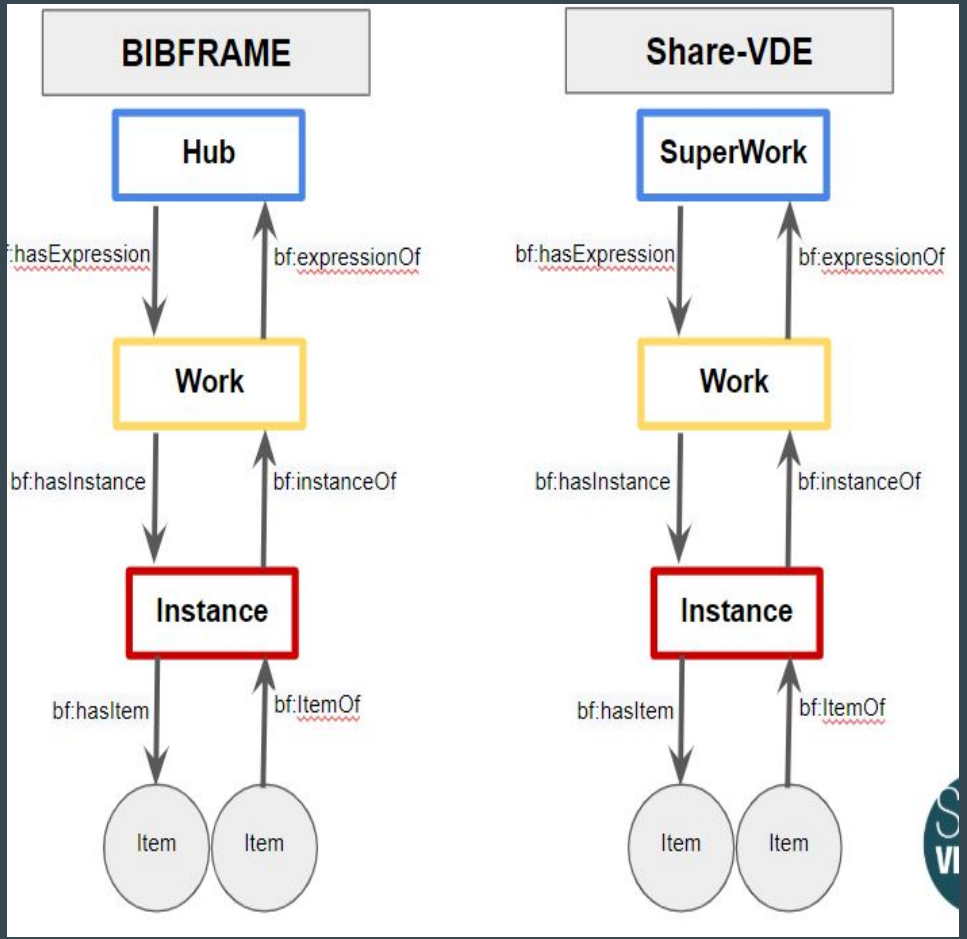
BIBFRAME (as you know it)

- Three core levels of abstraction
  - Work
  - Instance
  - Item
- Additional key concepts
  - Agents
  - Subjects
  - Events
- Consists of RDF classes and properties
  - members of a class share certain characteristics and may have subclasses
  - properties describe characteristics of resources as well as relationships among resources



BIBFRAME today

- Four core levels of abstraction
  - *Opus (Hub, SuperWork)*
  - Work
  - *Master* Instance
  - Item
- Additional key concepts
  - Agents
  - Subjects
  - Events
- Consists of RDF classes and properties
  - members of a class share certain characteristics and may have subclasses
  - properties describe characteristics of resources as well as relationships among resources



Slide courtesy Anna Lionetti and Share VDE

UNIVERSITY OF ALBERTA LIBRARIES

BIBFRAME in Canada

...

Canadian Linked Data Initiative (2015-2017):
Strategic Intentions

1. Create a linked data learning culture at our institutions and in Canada
2. Sustain an environment of experimentation with linked data
3. Communicate effectively within CLDI and to the larger community

Working Groups: Metadata, Digital Projects, Education and Training, Grants, Groupe de Travail Francophone, IT, Planning, User Experience

Canadian Federation of Library Associations (CFLA)

CFLA Cataloguing and Metadata Standards Committee

- **Canadian Committee on Cataloguing**
  - "The Canadian Committee on Cataloguing (CCC) is a national advisory committee on matters of cataloguing and bibliographic control. It also represents Canada on the North American RDA Committee (NARDAC). NARDAC represents North America on the RDA Steering Committee (RSC)."[1]
- **Canadian Committee on Metadata Exchange**
  - "The Canadian Committee on Metadata Exchange (CCM) is the national advisory committee on MARC 21 formats and related national and international standards for the representation in machine-readable form of bibliographic information."[2]
- **Canadian BIBFRAME Readiness Task Force**

1. Canadian Committee on Cataloguing (2018). Introduction. Retrieved from: https://www.bac-lac.gc.ca/eng/services/cataloguing-metadata/Pages/canadian-committee-cataloguing.aspx
2. Canadian Committee on Metadata Exchange (2017). General Information. Retrieved from: http://www.marc21.ca/040010-203-e.html

Building an International BIBFRAME Community:

...

Partnerships and Collaborations

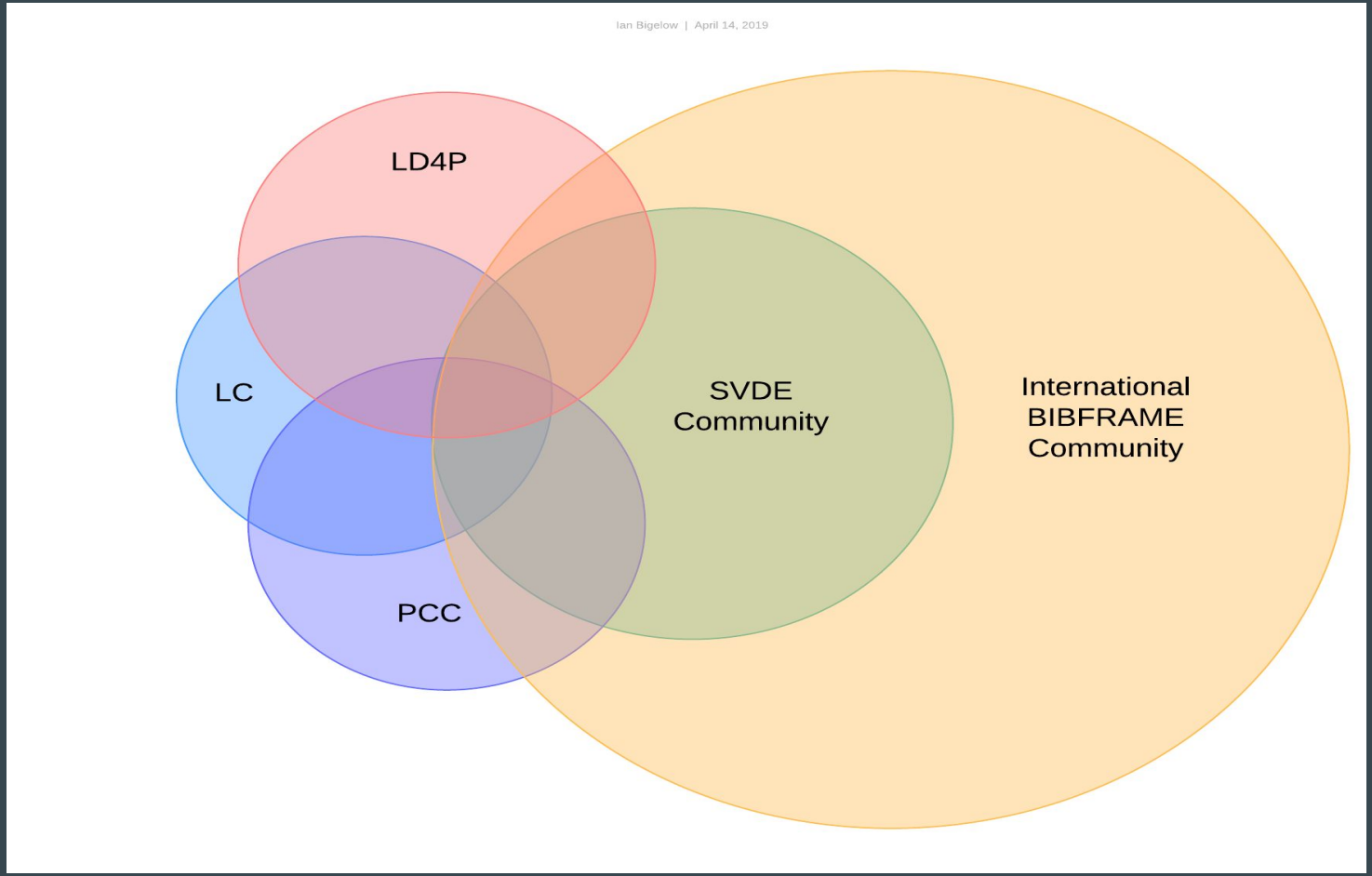
Program for Cooperative Cataloguing (PCC)

PCC Strategic Directions January 2018-December 2021

"It is time to move beyond knowledge and skills related to linked data at a theoretical level and into implementation. Building on the PCC's strong tradition of providing training for metadata creators, active experimentation and piloting of linked data practices will help inform policy decisions, training, and operationalizing such practices. As we move to a culture of greater data sharing, it is crucial to extend our community, both by engaging a more diverse range of members in the work of the PCC and by collaborating with vendors, open source communities, and others." (Program for Cooperative Cataloguing, 2018)

Program for Cooperative Cataloguing (2018). PCC (Program for Cooperative Cataloging) Strategic Directions January 2018-December 2021. Retrieved from https://www.loc.gov/aba/pcc/about/PCC-Strategic-Directions-2018-2021.pdf

BIBFRAME Project Overlap



European BIBFRAME Workshop -> BIBFRAME Workshop in Europe -> International BIBFRAME Community

From Experimentation to Implementation

In her article on the development of BIBFRAME, McCallum (2017) reflects on the current state of the transition of standards in libraries:

"In the 1960's and 1970's the AACR cataloguing rules and MARC format for bibliographic data were developed. Forty years later we are in the transition to new cataloguing rules and also a new carrier environment, with RDA and BIBFRAME." (p. 84)

How far away is this transition for libraries? We can look to the work of the library community:

- LC and other National Libraries
- PCC
- LD4
- SHARE VDE
- Work in Canada

Work through to 2018 resulted in a much wider community of practice for BIBFRAME and we were faced by a tipping point where experimentation needed to transition to implementation.

McCallum, S. s. (2017). BIBFRAME Development. JLIS.It, Italian Journal Of Library, Archives & Information Science, 8(3), 71-85.

At the Annual Meeting of the American Library Association (ALA) in June 2018, LC confirmed that BIBFRAME will be their replacement for MARC

Alea iacta est

Linked Data at UAL:

...

Vision and Strategic Priorities

Moving Forward with Linked Data at UAL

Linked data implementation as a strategic priority

"In order to reap the benefits of full participation in the linked open data environment, UAL should continue to take steps towards complete conversion of existing library data to linked open data. This would involve a full transition of workflows for resource description/metadata creation to linked open data, transitioning all library systems for resource discovery so they work with linked open data formats, and developing new workflows, both internal and with associated vendors and partners, to support these steps."[1]

"Work collaboratively with national and international partners to provide leadership and advance a transition toward open linked data for libraries." [2]

1. Moving Forward with Linked Data at UAL
2. UAL Strategic Priorities 2019-20

Linked Data for UAL

...

Work at UAL

- Strategic priorities
- Unit level updates
- Training
- Technical infrastructure
- Experimentation and analysis
- Engagement with community

Building support and capacity for BIBFRAME at UAL

Institutional support

Vision: Moving Forward with Linked Data at UAL & Strategic priorities 2019-2020

Cataloguing and Metadata Strategies Unit

Staffing

- Updates to expectations and job fact sheets
  - “Position must continue building expertise in current formats and standards (MARC, AACR, RDA), while developing expertise for description in RDF and associated query languages such as SPARQL”
  - “Works on ongoing provision of metadata to meet current needs, with an eye to new specifications such as BIBFRAME and associated transitioning of workflows”
- New Monographs Cataloguing Specialist tied to the LD4P Cohort project
- Linked Data Librarian Resident

Training

- Ongoing review of webinar options
- Common training through LibraryJuice Certificate in XML & RDF Based Systems
- Linked data lab time sessions to collaboratively work through concepts & ideas
- LC provided training sessions for LD4P Cohort members, with further sessions pending for Sinopia

Infrastructure

- Testing of NEOS data in a test triplestore (GraphDB) with support from ComputeCanada
- Testing of triplestore database work via the Stardog triplestore for Share VDE
- Share VDE provided discovery tool for testing
- Sinopia provided cataloguing module for working with BIBFRAME
- ITS is working on the setup of a local “production” triplestore for NEOS data

Experimentation and Analysis

- MARC to BIBFRAME conversions
- University of Alberta Libraries' Linked Data Enrichment tool

Engagement with Community

- Participation and presentations at relevant conferences and meetings
- PCC URI in MARC committee
- Canadian BIBFRAME Readiness Task Force
- Canadian Linked Data Initiative
- Share Virtual Discovery Environment (Share VDE)
- Linked Data for Production Phase 2 (LD4P2)

LD4P Cohort

- What is LD4P
- Data creation vs conversion
- Overview of UAL proposal for the LD4P Cohort
- Outline of work to date
- Our experiences - Working groups, affinity groups, et cetera

Linked Data for Production (LD4P)

For the past two years, Linked Data for Production has been focusing on:

- Developing standards, guidelines, and infrastructure to communally produce metadata as linked open data
- Developing end-to-end workflows to create linked open data in a technical services production environment
- Extending the BIBFRAME ontology to describe library resources in specialized domains and formats
- Engaging the broader library community to ensure a sustainable and extensible environment

LD4P Phase 2 and the LD4P Cohort

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

1. The creation of a continuously fed pool of linked data expressed in BIBFRAME-based application profiles.
2. The development of an expanded cohort of libraries (the LD4P Cohort) capable of the creation and reuse of linked data through the creation of a cloud-based sandbox editing environment.
3. 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.
4. The development of policies, techniques, and workflows for the creation and reuse of linked data and its supporting identifiers as libraries' core metadata.
5. Better integration of library metadata and identifiers with the Web through collaboration with Wikidata.
6. The enhancement of a widely-adopted library discovery environment (Blacklight) with linked-data based discovery techniques.
7. The orchestration of continued community collaboration through the development of an organizational framework called LD4.

LD4P2 Cohort Membership

University of Alberta
University of California, Davis
University of California, San Diego
Casalini Libri
University of Chicago
University of Colorado
Cornell University
Duke University
Frick Art Reference Library
Harry Ransom Center
Harvard University
University of Iowa
Library of Congress
University of Michigan
University of Minnesota
National Library of Medicine
Northwestern University
PCC
University of Pennsylvania
Princeton University
Stanford University
Texas A&M University
University of Washington
Yale University

Share VDE's Current Role in LD4P: Native BIBFRAME Data Creation vs. MARC Record Conversion

- Focused on creating BIBFRAME data natively using the Sinopia linked data editor
- Functionality to import external RDF data, and export BIBFRAME descriptions
- Investigating copy cataloguing workflows utilizing BIBFRAME data converted from MARC by SVDE

- Conversion of cohort institutions' MARC records to BIFFRAME
- Entity clustering and management within the Sapientia Knowledge Base for:
  - Agents
  - Places
  - Subjects
  - Bibliographic entities (Superwork, work, etc.)

UAL LD4P Cohort Project Summary

1. Enhancement of conversion, reconciliation and enrichment processes for MARC to BIBFRAME
2. Exploration of new forms of authority control based on URIs by utilizing MARC and BIBFRAME data enriched with URIs
3. Conversion of Monographs Team Operations
4. Community building:
    a. To help foster a wider community of linked data experimentation and implementation in Canada, UAL will work with other Canadian participants to liaise with the cataloguing community and standards organizations in Canada (CFLA, CCC, CCM, CLDI)
    b. Engage NEOS Consortium members in aspects of this work to transition towards linked data, so that we can move forward together.

SINOPIA

Announcing Sinopia Version 1.0

- New in this release
  - Search for descriptions created in Sinopia
  - Edit previously created descriptions
  - Look up and refer to Sinopia-created entities and to entities from Share-VDE datasets (as they become available)
  - Better messaging when a resource template is not configured as expected
- For complete 1.0.1 release notes see the Sinopia help site.
- Version 1.0.1: adds support for lookups to Share-VDE data and fixes bug where validation messages were not properly cleared

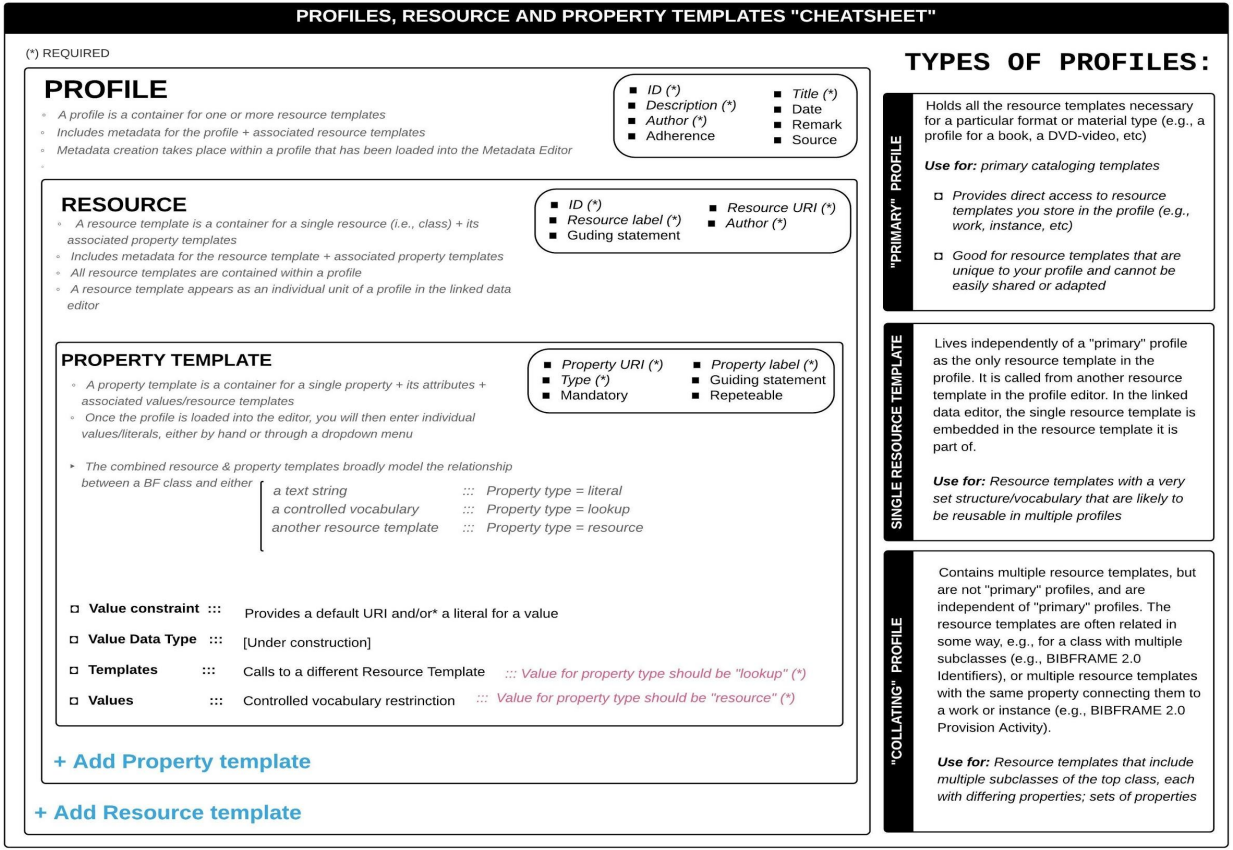
The underdrawing for the new world of linked data in libraries

Sinopia is a linked data creation environment where libraries can:

- create metadata in a linked data environment without having to set up and maintain tools
- learn best practices related to linked data creation
- explore the idea of cooperative cataloging (linking to shared descriptions and identifiers) in a linked data environment
- contribute feedback and expertise to iterative development of tools for working in a linked data environment

Sinopia is developed by the Linked Data for Production: Pathway to Implementation (LD4P2)project, a collaboration among Cornell University, Harvard University, the Library of Congress, Stanford University, the School of Library and Information Science at the University of Iowa, and the Program for Cooperative Cataloging (PCC).

Profile Building



PROFILES, RESOURCE AND PROPERTY TEMPLATES "CHEATSHEET"

(*) REQUIRED

PROFILE

- A profile is a container for one or more resource templates
- Includes metadata for the profile + associated resource templates
- Metadata creation takes place within a profile that has been loaded into the Metadata Editor

- ID (*)
- Description (*)
- Author (*)
- Adherence
- Title (*)
- Date
- Remark
- Source

RESOURCE

- A resource template is a container for a single resource (i.e., class) + its associated property templates
- Includes metadata for the resource template + associated property templates
- All resource templates are contained within a profile
- A resource template appears as an individual unit of a profile in the linked data editor

- ID (*)
- Resource label (*)
- Guding statement
- Resource URI (*)
- Author (*)

PROPERTY TEMPLATE

- A property template is a container for a single property + its attributes + associated values/resource templates
- Once the profile is loaded into the editor, you will then enter individual values/literals, either by hand or through a dropdown menu

- Property URI (*)
- Type (*)
- Mandatory
- Property label (*)
- Guiding statement
- Repeteable

- The combined resource & property templates broadly model the relationship between a BF class and either
  - a text string ::: Property type = literal
  - a controlled vocabulary ::: Property type = lookup
  - another resource template ::: Property type = resource

- **Value constraint** ::: Provides a default URI and/or* a literal for a value
- **Value Data Type** ::: [Under construction]
- **Templates** ::: Calls to a different Resource Template ::: Value for property type should be "lookup" (*)
- **Values** ::: Controlled vocabulary restrinction ::: Value for property type should be "resource" (*)

+ Add Property template

+ Add Resource template

TYPES OF PROFILES:

"PRIMARY" PROFILE

Holds all the resource templates necessary for a particular format or material type (e.g., a profile for a book, a DVD-video, etc)

Use for: *primary cataloging templates*

- *Provides direct access to resource templates you store in the profile (e.g., work, instance, etc)*
- *Good for resource templates that are unique to your profile and cannot be easily shared or adapted*

SINGLE RESOURCE TEMPLATE

Lives independently of a "primary" profile as the only resource template in the profile. It is called from another resource template in the profile editor. In the linked data editor, the single resource template is embedded in the resource template it is part of.

Use for: *Resource templates with a very set structure/vocabulary that are likely to be reusable in multiple profiles*

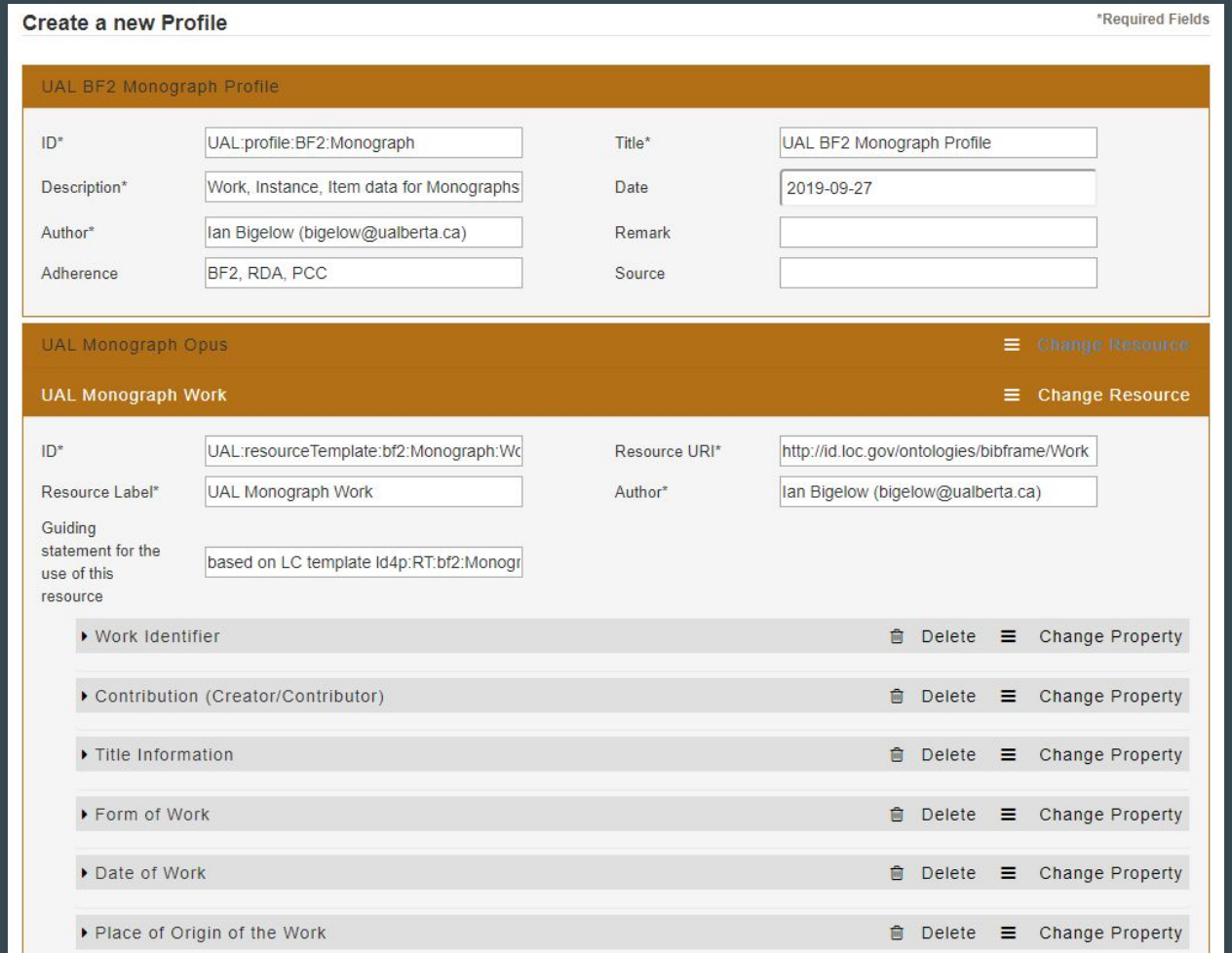
"COLLATING" PROFILE

Contains multiple resource templates, but are not "primary" profiles, and are independent of "primary" profiles. The resource templates are often related in some way, e.g., for a class with multiple subclasses (e.g., BIBFRAME 2.0 Identifiers), or multiple resource templates with the same property connecting them to a work or instance (e.g., BIBFRAME 2.0 Provision Activity).

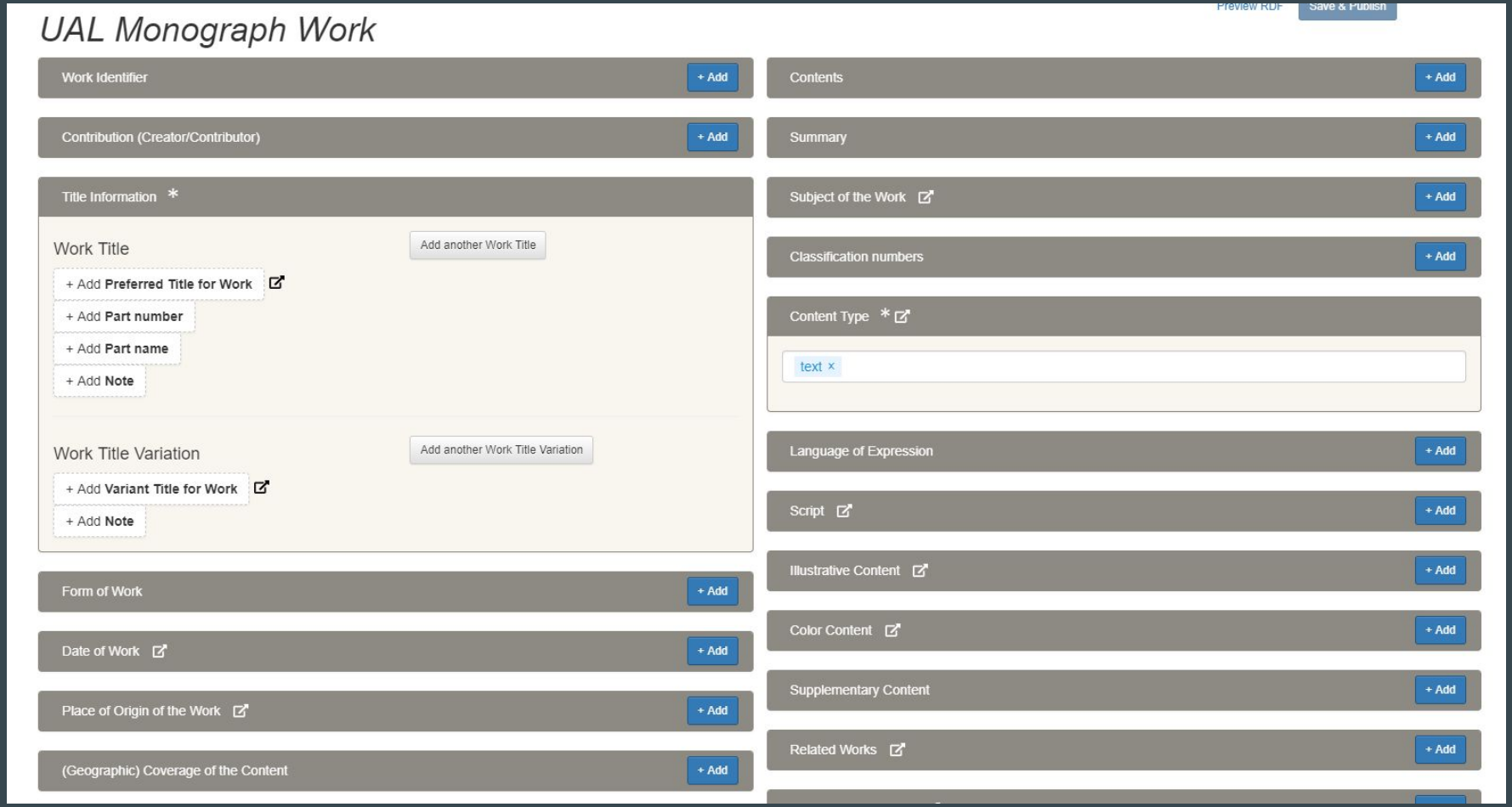
Use for: *Resource templates that include multiple subclasses of the top class, each with differing properties; sets of properties*

Credit: *Corralling Bibframe Profiles: Using the BIBFRAME Profile Editor in a Shared Environment*, Nancy Lorimer and Paloma Graciana Picardo, LD4 Conference, Boston, May 2019

Profile Editor



Linked Data Editor



Authority Lookup Service

Linked Data for Libraries

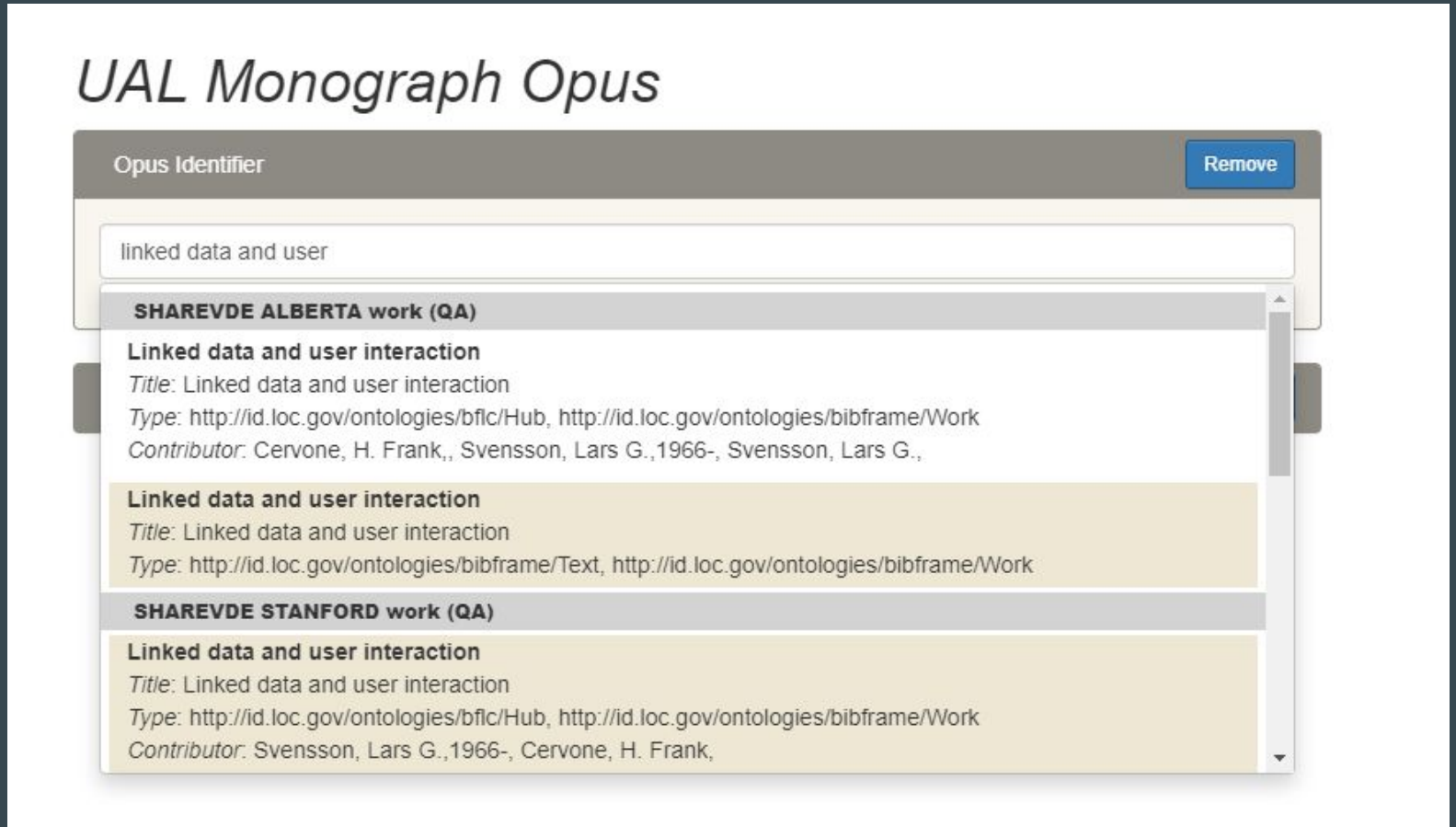
Authority List

See Check Status to test whether an authority is online.

Authority/Subauthority	Service	Action	Sample URL
AGROVOC_LD4L_CACHE			
	ld4l_cache	term	/authorities/fetch/linked_data/agrovoc_ld4l_cache?uri=http%3A%2F%2Faims%2Efao%2Eorg%2Faos%2Fagrovoc%2Fc_9513
	ld4l_cache	search	/authorities/search/linked_data/agrovoc_ld4l_cache?q=milk&maxRecords=4
DBPEDIA_DIRECT			
	direct	term	/authorities/fetch/linked_data/dbpedia_direct?uri=http%3A%2F%2Fdbpedia%2Eorg%2Fresource%2FBarack_Obama
DBPEDIA_LD4L_CACHE			
	ld4l_cache	term	/authorities/fetch/linked_data/dbpedia_ld4l_cache?uri=http%3A%2F%2Fdbpedia%2Eorg%2Fresource%2FBarack_Obama
	ld4l_cache	search	/authorities/search/linked_data/dbpedia_ld4l_cache?q=Barack Obama&maxRecords=4
GEONAMES_DIRECT			
	direct	term	/authorities/fetch/linked_data/geonames_direct?uri=http%3A%2F%2Fsws%2Egeonames%2Eorg%2F261707%2F
	direct	search	/authorities/search/linked_data/geonames_direct?q=Ithaca&maxRecords=4
GEONAMES_LD4L_CACHE			

Example of Questioning Authority (QA) Lookups

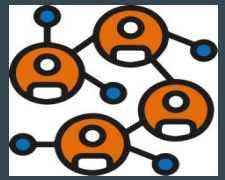
Universal Work Identifiers in Practice!



Affinity and Working Groups

- Profiles Working Group
- Discovery Affinity Group
- Non-Latin Script Materials Affinity Group
- Rare Materials Affinity Group
- LD4-Wikidata Affinity Group
- Ethics in Linked Data Affinity Group
- Serials Affinity Group
- Sinopia Training Task Group

What is Share-VDE?



Share-VDE is a library-driven initiative to establish an effective working environment for the use of linked data by libraries within a global context.

Library data are enriched with additional information and relationships, and bibliographic and authority data are reconciled and converted into linked data.

A virtual discovery platform with the structure of the BIBFRAME data model is created to simplify the way in which that data is consumed.

The network of resources created is the basis for the Share-VDE Sapientia Cluster Knowledge Base, the common authoritative source of clusters accessible in RDF, open to the entire Share-VDE community.

Who is responsible for it?

Share-VDE is a collaborative endeavour based on the needs of libraries, developed by:

the joint effort of the Share-VDE Advisory Council and of the Working Groups;

Casalini Libri, provider of bibliographic and authority data as member of the Program for Cooperative Cataloguing;

@Cult, provider of ILS, Discovery tools and Semantic web solutions for the cultural heritage sector;

influenced by the vision of the LD4P initiative;

with input and active participation from an international group of research libraries.

Share-VDE Advisory Council

The Share-VDE Advisory Council's role is to provide insight and analysis of the MARC to BIBFRAME transformation to make recommendations for improvements based on member library data analysis, and project documentation. The AC also provides overall guidance to the activities of Share-VDE initiative.

There are 4 sub-committees focusing on specific areas:

- Work Identification Working Group
- Authority/Identifier Management Services Working Group
- Cluster Knowledge Base Editor Working Group
- User experience/User Interface Working Group

Work Identification WG

- Reviews the Share-VDE work clustering processes and submits feedback on potential improvements or optimizations;
- reviews the use of primary resource identifiers in the Share-VDE dataset and provides feedback as appropriate;
- engages with the PCC to identify and/or develop best practices for use of these identifiers in BIBFRAME and MARC data.

Authority/Identifier Management Services WG

- Defines guidelines and best practices for Authority/Identifier management in the linked data environment;
- defines scope and data-flow for creation and implementation of Service based on preliminary documentation;
- proposes additional use cases identified as essential for effective knowledge base management.

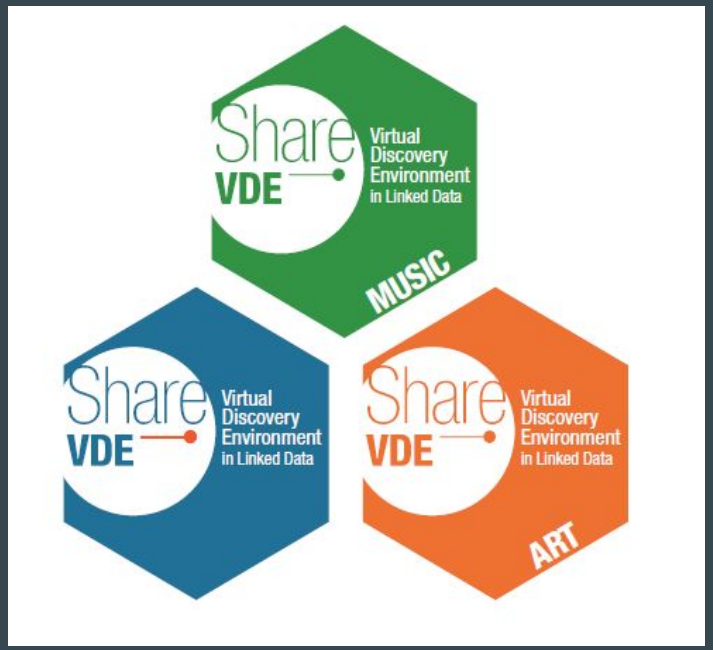
Cluster Knowledge Base Editor WG

- An essential part of the MARC to RDF conversion process is the maintenance of metadata that have been produced and registered on the Cluster Knowledge Base (CKB);
- the group analyses how libraries interact with the CKB and their use of the Editor for modifying (correcting / enriching), deleting, merging and separating clusters;
- the same approach will be applied on the data originally created in BIBFRAME.

User Experience/User Interface WG

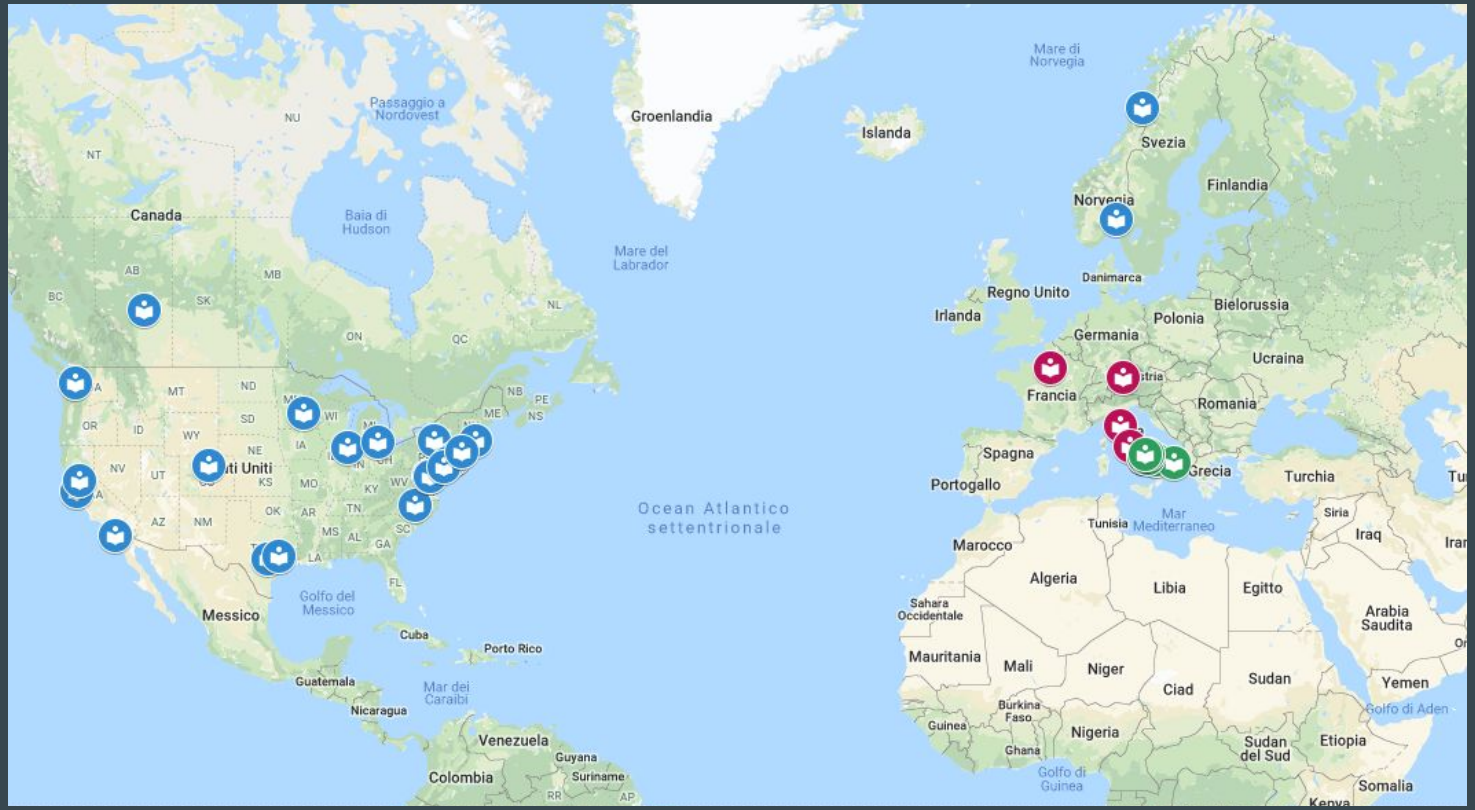
- Re-design Share-VDE user interface to respond to both patrons and library staff requirements and expectations;
- reflect the components of the Share-VDE data model infrastructure;
- provide an intuitive user experience.

The Share family

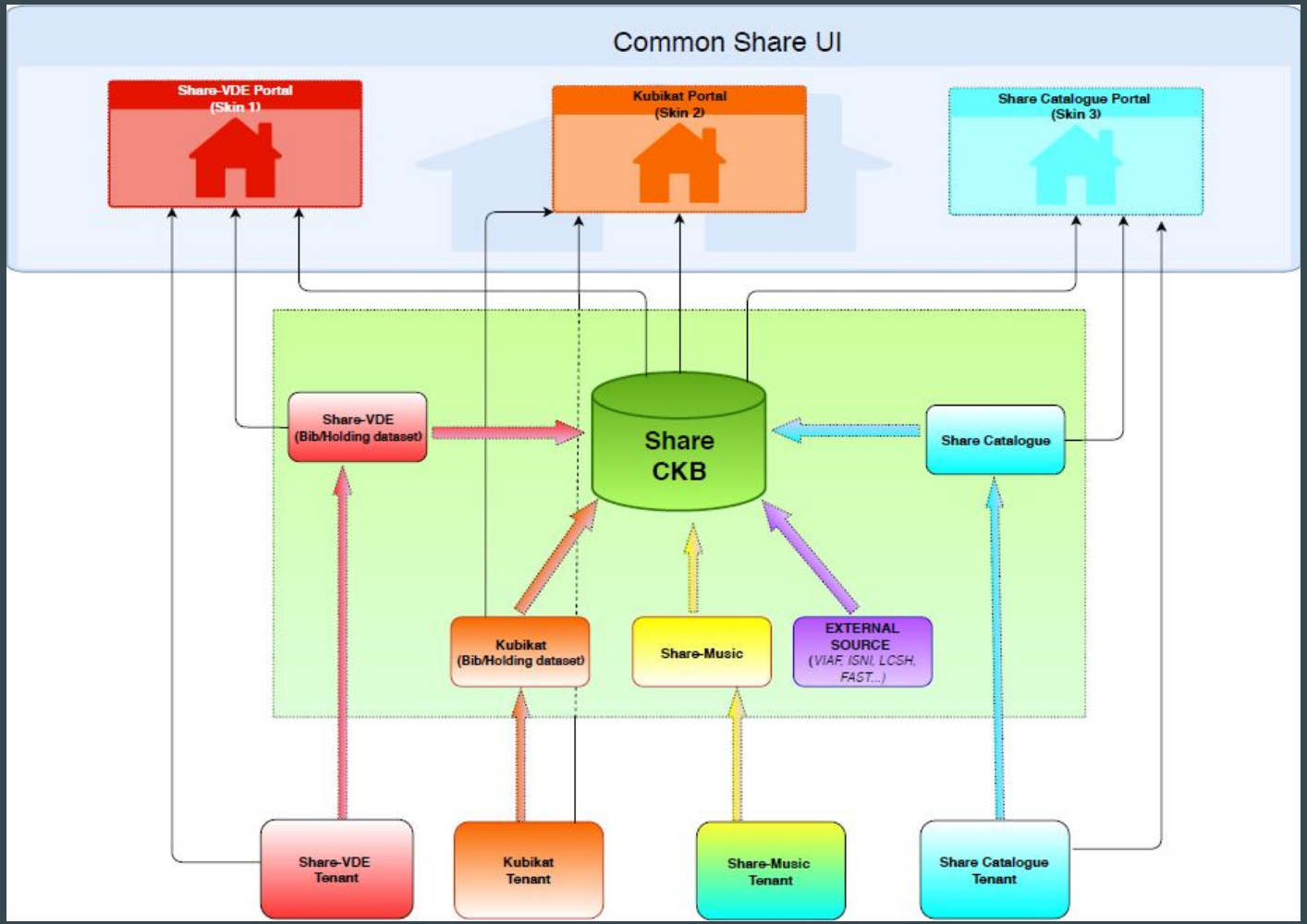


The Share family of initiatives based on linked data comprises Share-VDE (Virtual Discovery Environment), Share-ART (prompted by the Art History libraries of the Max Planck Institut), and Share-MUSIC (a pilot in the music domain). The different characteristics of each field are a useful asset that can be used to the advantage not only of the Share family as a whole, but for each single discipline.

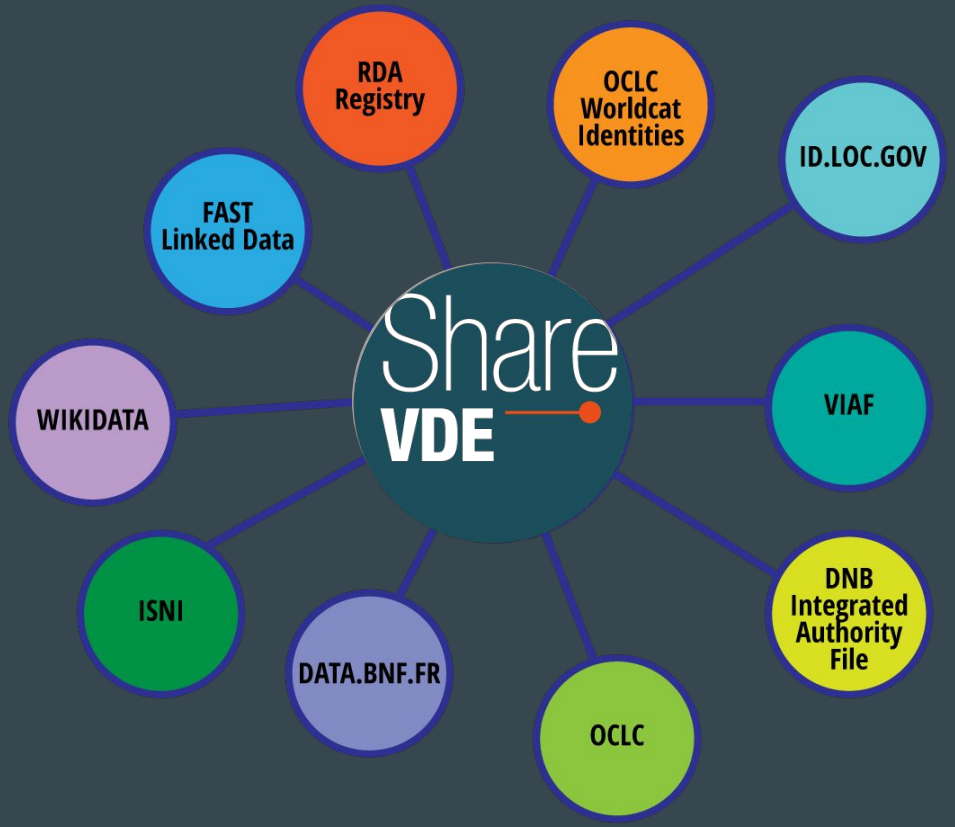
Share-VDE map



http://bit.ly/Share-VDE_map


Common Share UI
Share-VDE Portal (Skin 1)
Kubikat Portal (Skin 2)
Share Catalogue Portal (Skin 3)
Share-VDE (Bib/Holding dataset)
Share CKB
Share Catalogue
Kubikat (Bib/Holding dataset)
Share-Music
EXTERNAL SOURCE (VIAF, ISNI, LCSH, FAST...)
Share-VDE Tenant
Kubikat Tenant
Share-Music Tenant
Share Catalogue Tenant

Some of the external sources



Share-VDE deliverables overview (1)

Linked

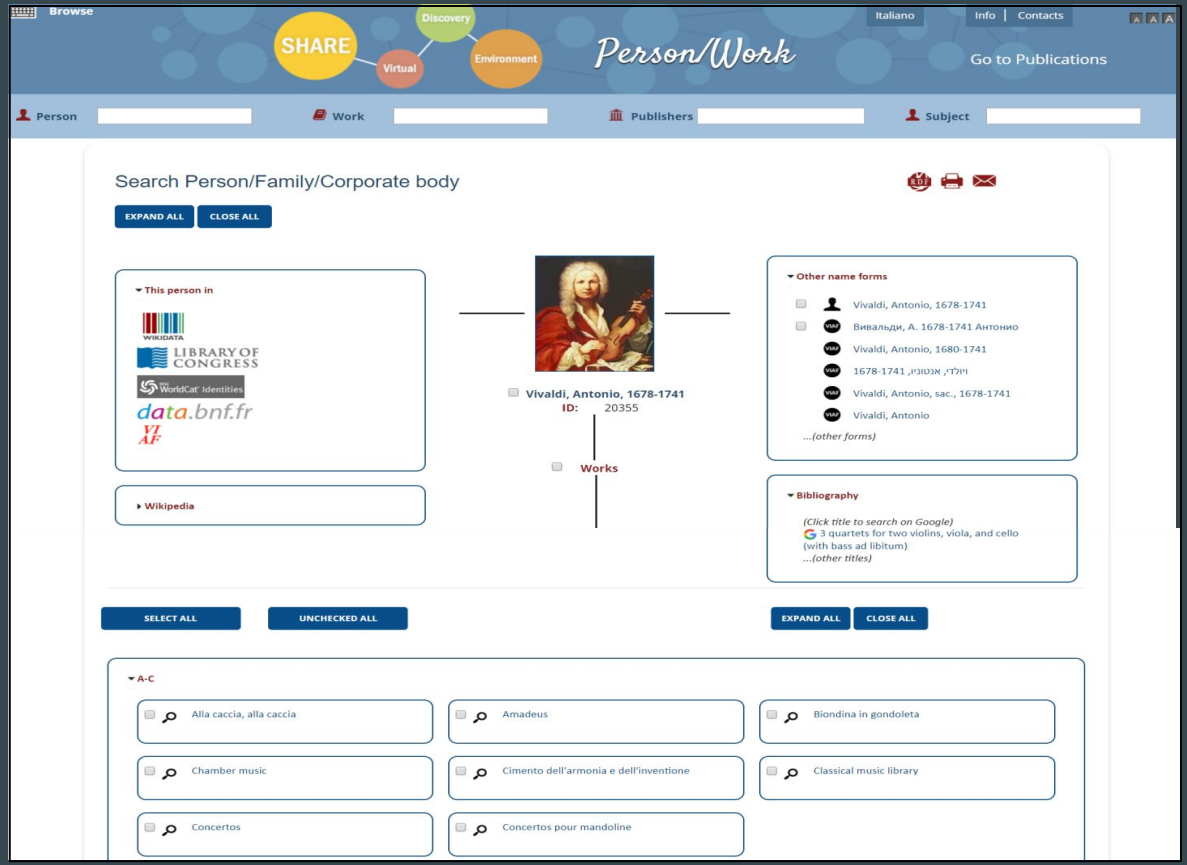
D1	The catalogue of each library converted into BIBFRAME 2.0* model	Entities are reconciled in the dataset and linked to Share-VDE project URIs of D2 for identification.
D2	The Cluster Knowledge Base in RDF format	Available at different levels for all institutions as it includes data from all of the participants. Entities in D2 are enriched with URIs from external sources. All variant forms are included.
D3	The dataset converted in BIBFRAME 2.0* with external URIs included	This dataset includes a certain number of relationships already present in the knowledge base. Works autonomously from D2.

* Including additional vocabularies and ontologies as needed

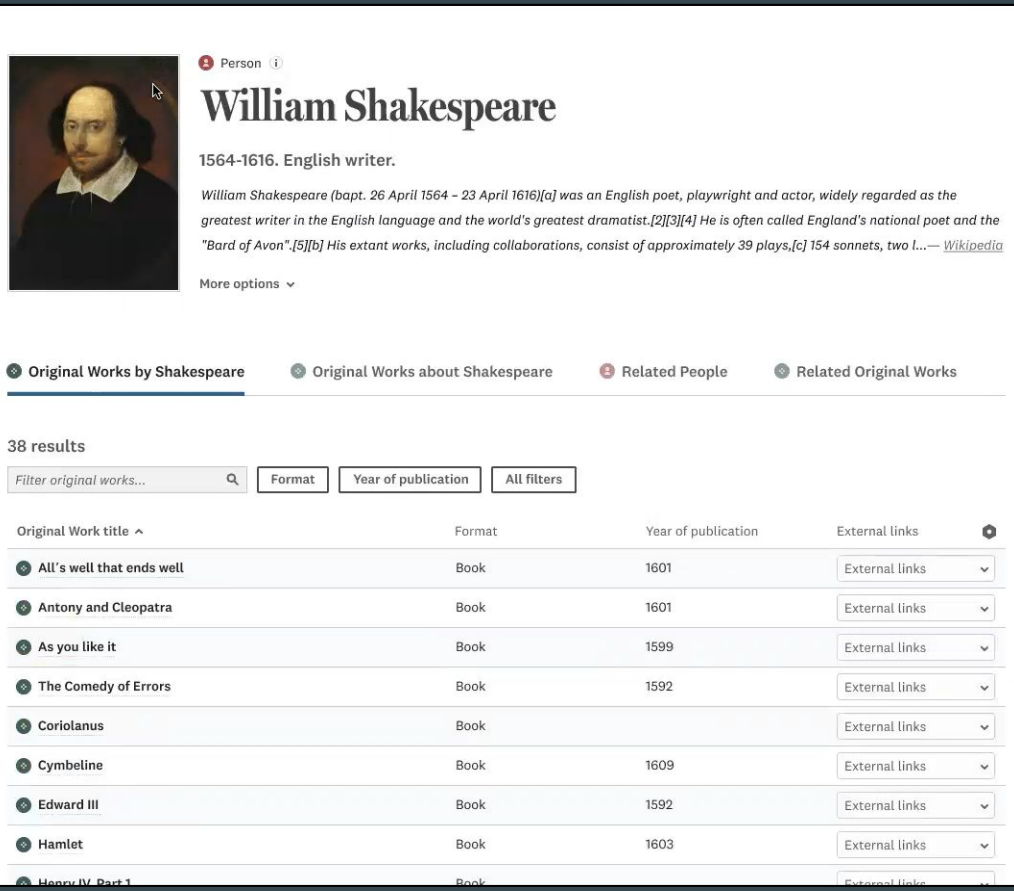
Share-VDE deliverables overview (2)

D4	The MARC21 version of D3	It includes all of the institution's records enriched with URIs.
D5	The Share-VDE bibliographical datasets in RDF made available on a triplestore	Constantly updated, queryable datasets with homogeneous data.
D6	The Share-VDE online platform as a common discovery system	Advanced discovery interface based on BIBFRAME, offering an easy and intuitive user experience and rich search results.

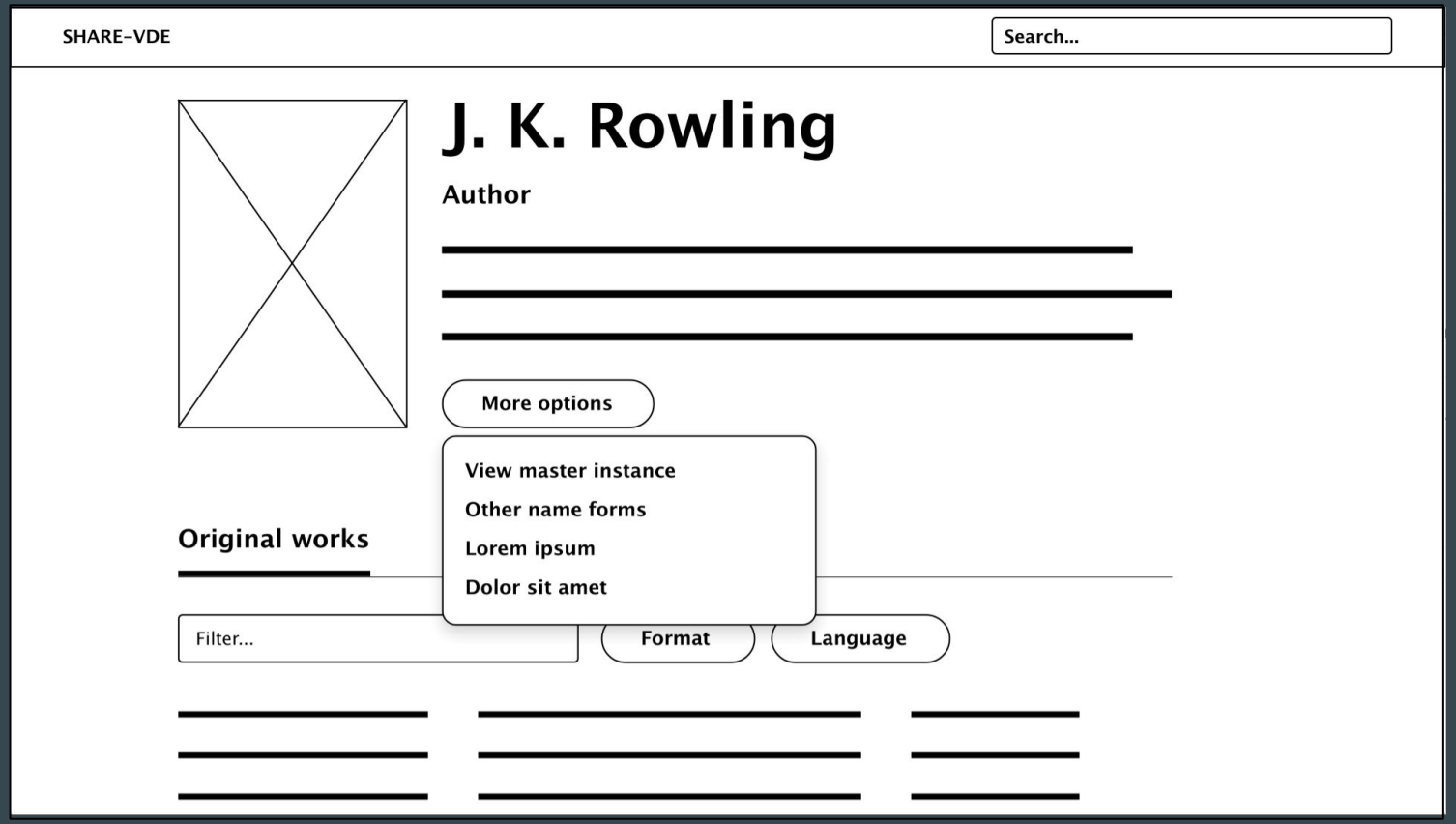
The Share-VDE portal



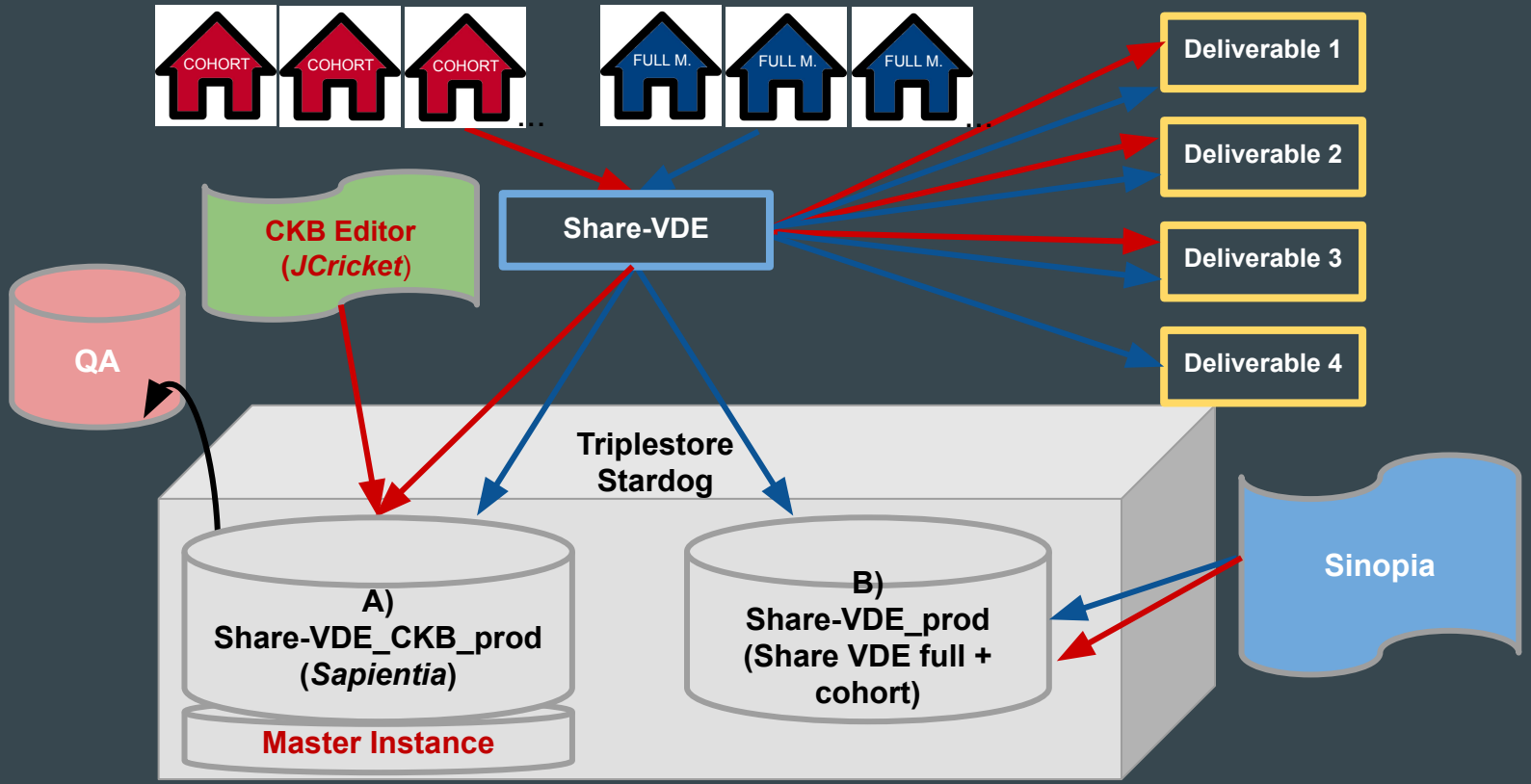
New User Interface design



JCricket – The CKB Editor



Progress of Share-VDE interactions



What comes next

- API layers for ILS, external applications and other LD systems (such as BF editors and triplestores);
- Authority Management and services;
- Reporting to serve library needs;
- Internationalization of the Share-VDE environment in relationship with new projects;
- Strategies to make the Share-VDE environment a trusted source of identifiers and to facilitate interaction with international initiatives as Wikidata, VIAF, ISNI etc.;
- Application for further Wikidata entity properties (Share-VDE has already https://www.wikidata.org/wiki/Property:P6329 assigned for authors).

Share VDE and UAL

...

BIBFRAME Implementation Scenarios

Share Participation

Share VDE Full Members
Duke University
New York University
Stanford University
University of Alberta – NEOS consortium
University of Chicago
University of Michigan at Ann Arbor
University of Pennsylvania
Yale University

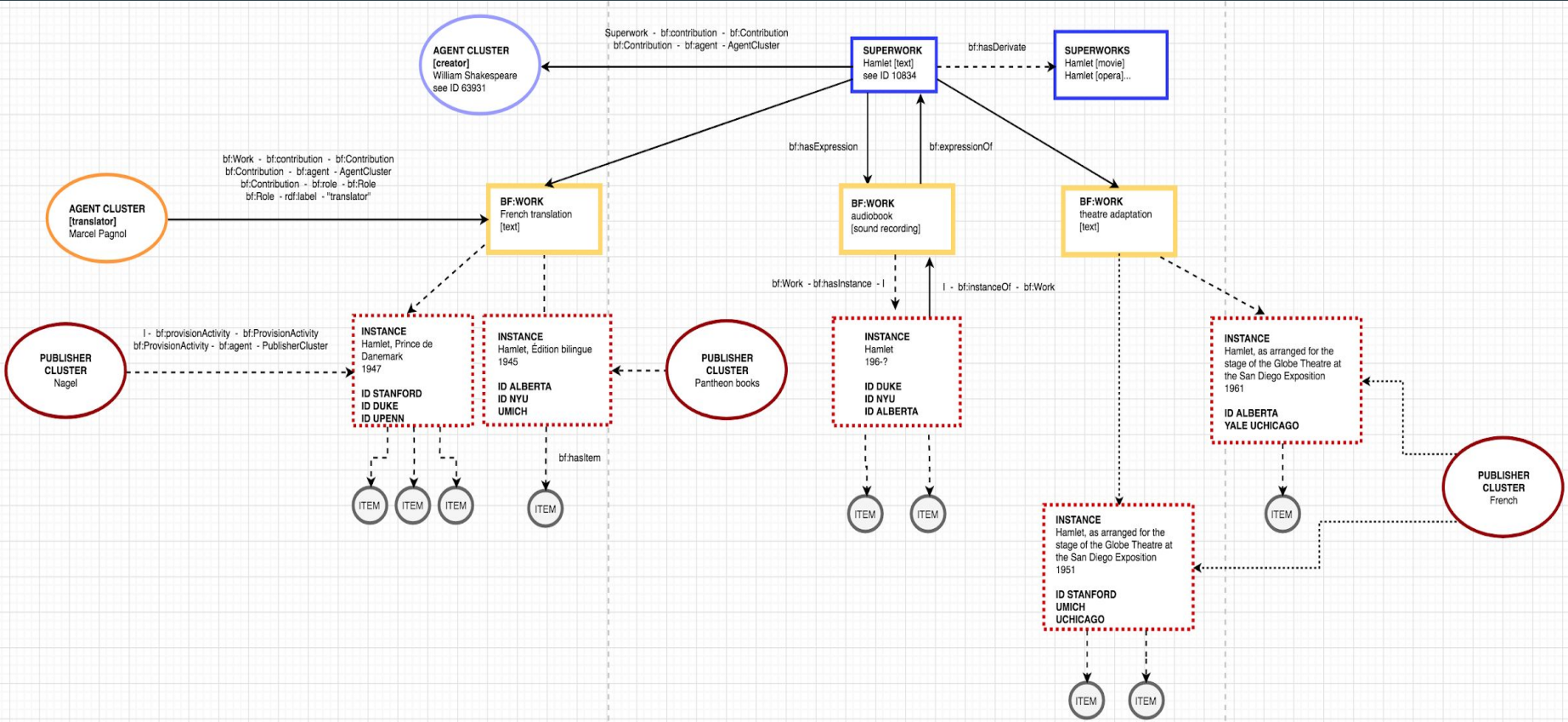
National Libraries
Library of Congress
National Library of Medicine
National Library of Norway
National Library of Finland

LD4P Cohort
Cornell University
Frick Art Reference Library
Harry Ransom Center
Harvard University
Northwestern University
Princeton University
UC Davis
UC San Diego
University Colorado at Boulder
University of Minnesota
University of Texas A&M
University of Washington

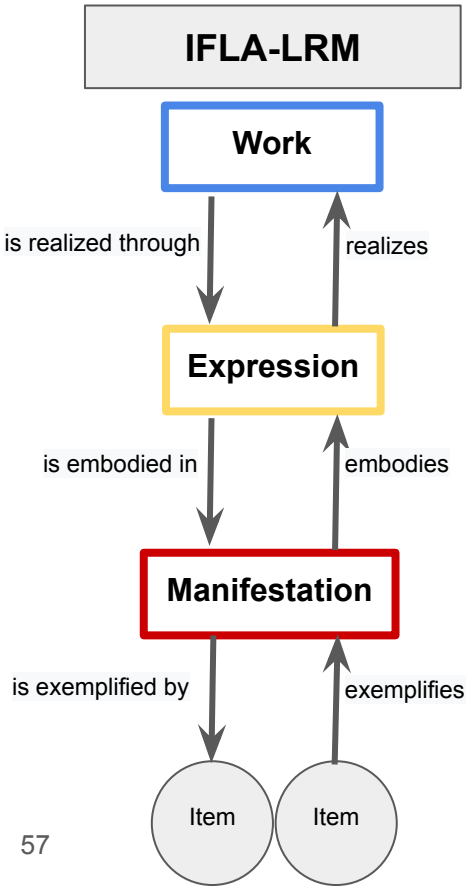
Share-Catalogue Institutions
Università Degli Studi di Napoli "Federico II"
Università degli Studi della Basilicata
Università Degli Studi di Napoli L'Orientale
Universita' degli Studi di Napoli Parthenope
Università del Salento
Università degli Studi di Salerno
Università degli Studi del Sannio RCost
Università degli Studi della Campania "Luigi Vanvitelli"

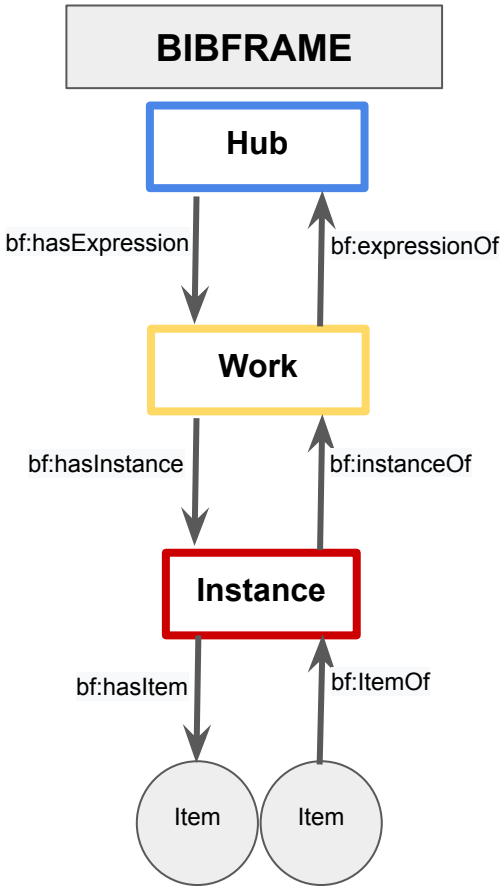
Share Art
Kunsthistorisches Institut in Florenz - Max-Planck-Institut
Central Institute of Art History
Deutsches Forum für Kunstgeschichte Paris / Centre allemand d'histoire de l'art Paris
Biblioteca Hertziana

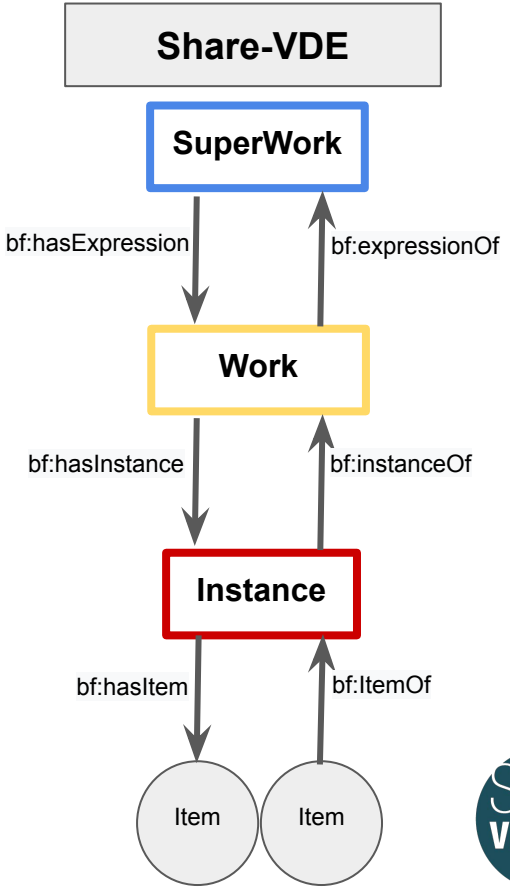
Share VDE Model Outline



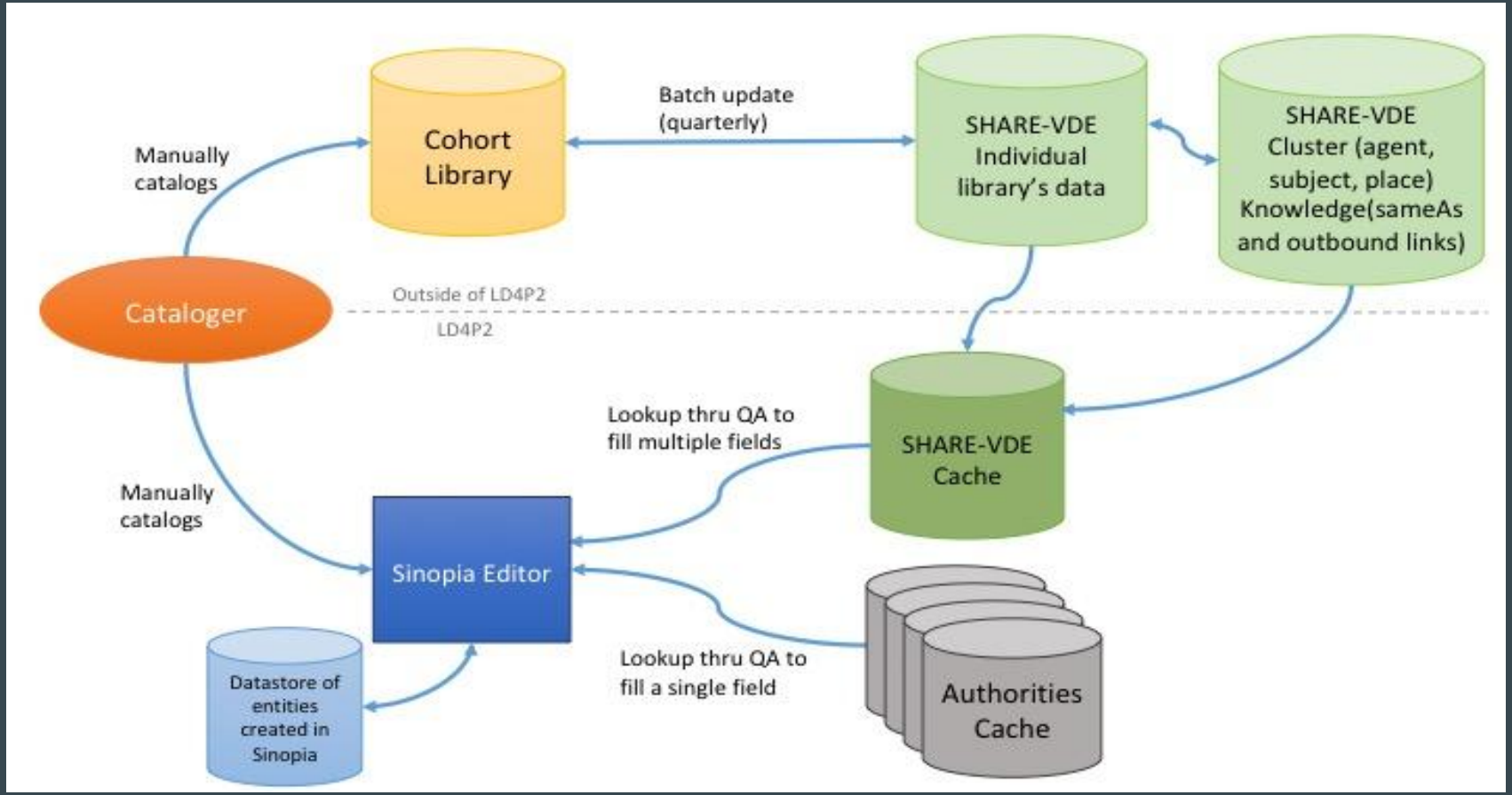
Ongoing BIBFRAME Developments







Bringing the projects together: LD4P/SHARE-VDE Data-flow diagram



Schreur, P. (2019). LD4P2 Data Flow

Next steps for UAL

- Infrastructure in general
- Refining an operational BIBFRAME Editor
- MARC to BIBFRAME conversion processes
- Ways to test discovery (Share VDE and LD4P)
- But regular work needs to continue, and we are part of a consortium
- Pain points
  - BIBFRAME to MARC
  - Batch processes
  - Testing workflows with vendors
  - Production environment that works with other key tools (ILL, Circ, ...)
  - Processes for entity/identifier support and maintenance
  - So far the focus has been on Monographs (for LD4P work)... what about things like serials
- Remember - Incremental change

Infrastructure change can seem messy!

Lessons Learned and Key Takeaways:

- Institutional strategic support is key (technical support, staff buy-in)
- Learning happens through training and on the ground, no one is an expert in everything (working groups, committee participation, emphasize that this community is growing and non-judgemental to new members)
- Acknowledge that during transition there won't be perfection (progress isn't synonymous with perfection!)
- Implementation through iteration (as above)
- Experimentation vs implementation: Much to be defined, refined, and developed, but we now have tools to move progressively into production

“Never let the future disturb you. You will meet it, if you have to, with the same weapons of reason which today arm you against the present.”

Marcus Aurelius, Emperor of Rome, 121-180. (2002). Meditations. London: The Folio Society.

Discussion: Building a Canadian Community of Practice

- **Work with some data:**
  - Editors
    - LC BIBFRAME Editor
    - Sinopia
  - Conversion
    - LC BIBFRAME converter (XSLT)
    - UAL active enrichment tool
  - Enrichment:
    - OpenRefine
    - MARCEdit

- **Training**
  - LC BIBFRAME training
  - Sinopia training videos
  - Watch for PCC training
- **Working Groups**
    - Canadian BIBFRAME Readiness
    - LD4P Affinity/Working Groups
    - Share VDE Working Groups
    - PCC task groups
- **Conferences**
  - LD4
  - Access
  - OLA