The Share Family:
One Project, Many Tenants

LD4 conference 2022

Share Family Tenant Infrastructure

Core Principles

- Redundancy of data is complex to manage

- Linking entities is easier than duplicate data

- Cooperate and maintain autonomy at the same time

- Homogeneity of datasets and possible services to be shared

- Centralize core data through a lightweight method

- Distribute the technologic load to achieve long-term sustainability

- Profile levels of cooperation among systems and initiatives

Solution in the Share Family architecture

Creation of more branches in the Share Family, named tenants

Consistent groups of institutions gathered by similar scope or from the same domain:

Share-VDE

Share-Catalogue

Kubikat-LOD

PCC data pool

Parsifal project (network of ecclesiastical university libraries in Rome)

National Bibliographies Group

What is a tenant

Definition of tenant from Wikipedia:

- "The term software multitenancy refers to a software architecture in which a single instance of software runs on a server and serves multiple tenants".

- "A tenant is a group of users who share a common access with specific privileges to the software instance. With a multitenant architecture, a software application is designed to provide every tenant a dedicated share of the instance - including its data, configuration, user management, tenant individual functionality" etc.

Benefits

More efficient data management

Technological sustainability

Dedicated applications and services tailored to the institutions members of the various branches

From the users perspective this enables richer and specialized sets of resources to be consulted

Share
FAMILY

Main components of the Share Family tenants

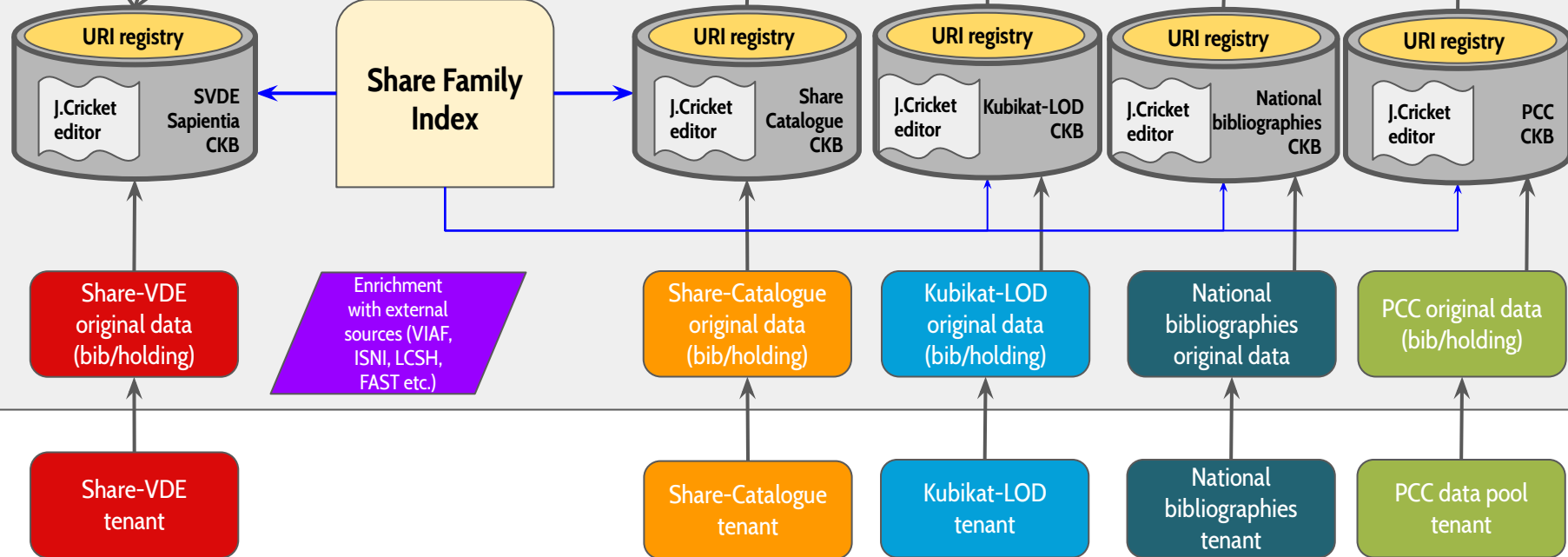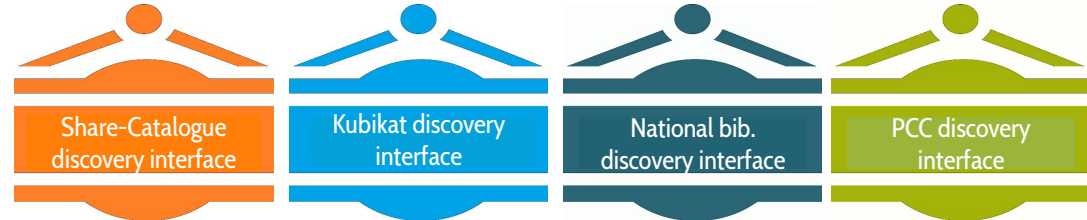Each tenant of the Share Family will have its own components and data will live in autonomous storages.

Each tenant has:

- its own CKB (e.g. Sapientia CKB, PCC CKB, Kubikat-LOD CKB, URBE CKB etc.)
- its own Cluster (URI) Registry with specific namespace
- its own J.Cricket CKB editor (= entity editor to manually update linked data entities)
- its own datastores (i.e. RDBMS, Search Engine and RDF Store)
- its own discovery portal with one or more skins (skin = filter on a part of tenant whole data)
  - e.g. general SVDE discovery portal skin https://svde.org + other ad hoc skins such as https://penn.svde.org/)

Share Family tenant infrastructure

Share-VDE User Interface

- Share-VDE skin portal
- Penn skin portal
- Other skin portals

- Share-Catalogue discovery interface
- Kubikat discovery interface
- National bib. discovery interface
- PCC discovery interface

Share Family Index

URI registry — J.Cricket editor — SVDE Sapientia CKB

URI registry — J.Cricket editor — Share Catalogue CKB

URI registry — J.Cricket editor — Kubikat-LOD CKB

URI registry — J.Cricket editor — National bibliographies CKB

URI registry — J.Cricket editor — PCC CKB

Enrichment with external sources (VIAF, ISNI, LCSH, FAST etc.)

- Share-VDE original data (bib/holding)
- Share-Catalogue original data (bib/holding)
- Kubikat-LOD original data (bib/holding)
- National bibliographies original data
- PCC original data (bib/holding)

- Share-VDE tenant
- Share-Catalogue tenant
- Kubikat-LOD tenant
- National bibliographies tenant
- PCC data pool tenant

PCC Tenant

PCC data pool

- Developed as part of the LD4P3 grant
- BIBCO, CONSER and NACO MARC records supplied by OCLC
- Transformed and clustered by SVDE clusterization tools
- Available as linked open data
- https://pcc-lod.org/

THE **PCC** CATALOG BETA

Search for people, original works and publications

e.g. *Jules Verne* or *Around the world in 80 days*

Advanced search

Share-VDE Community

Share-VDE members		Share Family members	
Share-VDE full members	**LD4P cohort members**	**Share-Catalogue Institutions**	**Kubikat-LOD project**
Duke University	Cornell University	Università Degli Studi di Napoli "Federico II"	Kunsthistorisches Institut in Florenz
New York University	Frick Art Reference Library	Università degli Studi della Basilicata	Biblioteca Hertziana, Rome
Stanford University	Harry Ransom Center Texas A&M	Università Degli Studi di Napoli L'Orientale	Central Institute of Art History, Munich
University of Alberta - NEOS consortium	Harvard University	Università degli Studi di Napoli Parthenope	Deutsches Forum für Kunstgeschichte, Paris
University of Chicago	National Library of Medicine	Università del Salento	
University of Michigan at Ann Arbor	Northwestern University	Università degli Studi di Salerno	
University of Pennsylvania	Princeton University	Università degli Studi del Sannio RCost	
Yale University	University of California Davis	Università degli Studi della Campania "Luigi Vanvitelli"	
	University of California San Diego		
National libraries	University Colorado at Boulder		**Share-Music project**
Library of Congress	University of Minnesota		Bayerische Staatsbibliothek
National Library of Norway	University of Texas A&M		Library of Congress
National Library of Finland	University of Washington		Stanford University
Smithsonian Institution			
The British Library			

https://wiki.share-vde.org/wiki/ShareFamily:Main_Page
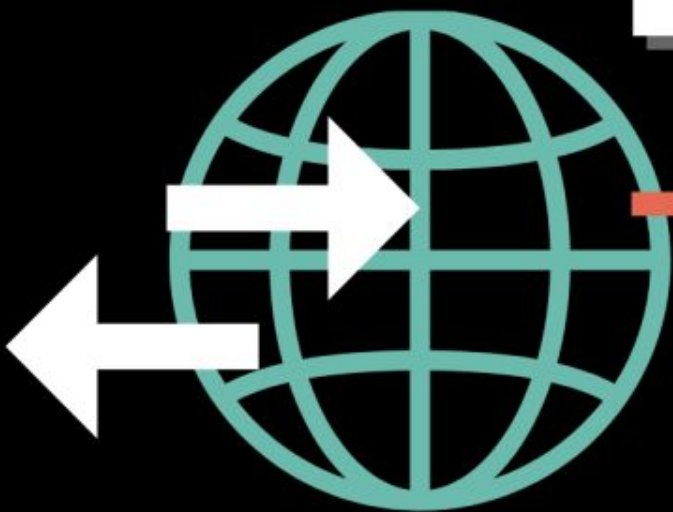
Linked Data Development Trajectory at UAL

- Experimentation and exploration (xxxx-2017)

- Moving forward with linked data (2017-2019)

- Assessment and planning (2019-2022)

- Linked Data Implementation Plan (2022-2027)

The Plan's Key Objectives and Areas

- Review the current state of linked data initiatives and practices at UAL
- Identify strategic priorities and provide context for them
- Define plans for short-term (6 – 18 months), medium-term (18 months – 3 years), and long-term (3 – 5 years) goals
- Create milestones to allow regular checking of progress
- Identify resources required to achieve short-term goals

Thematic Areas

- Data integration and workflows between systems
- Shared vocabularies and authority files and supporting technologies
- Authority control/entity management
- Decolonizing description
- Data enhancement with URI enrichment

Share FAMILY

Focus on Partnerships & New Infrastructure
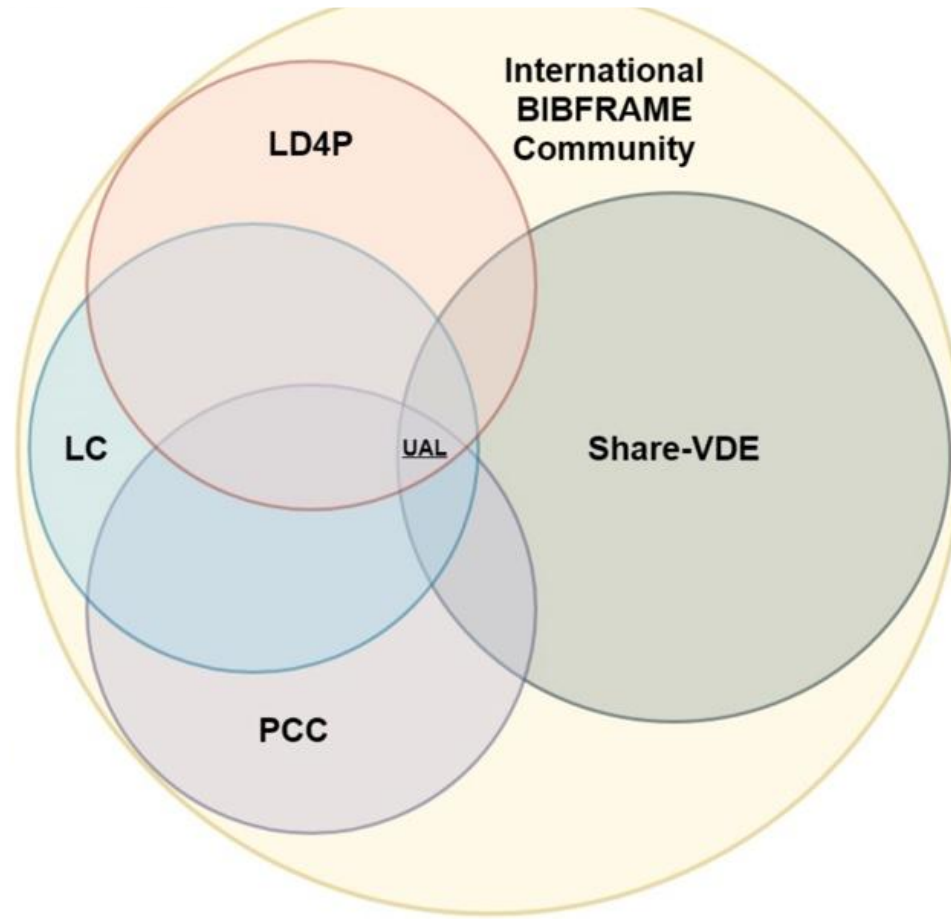
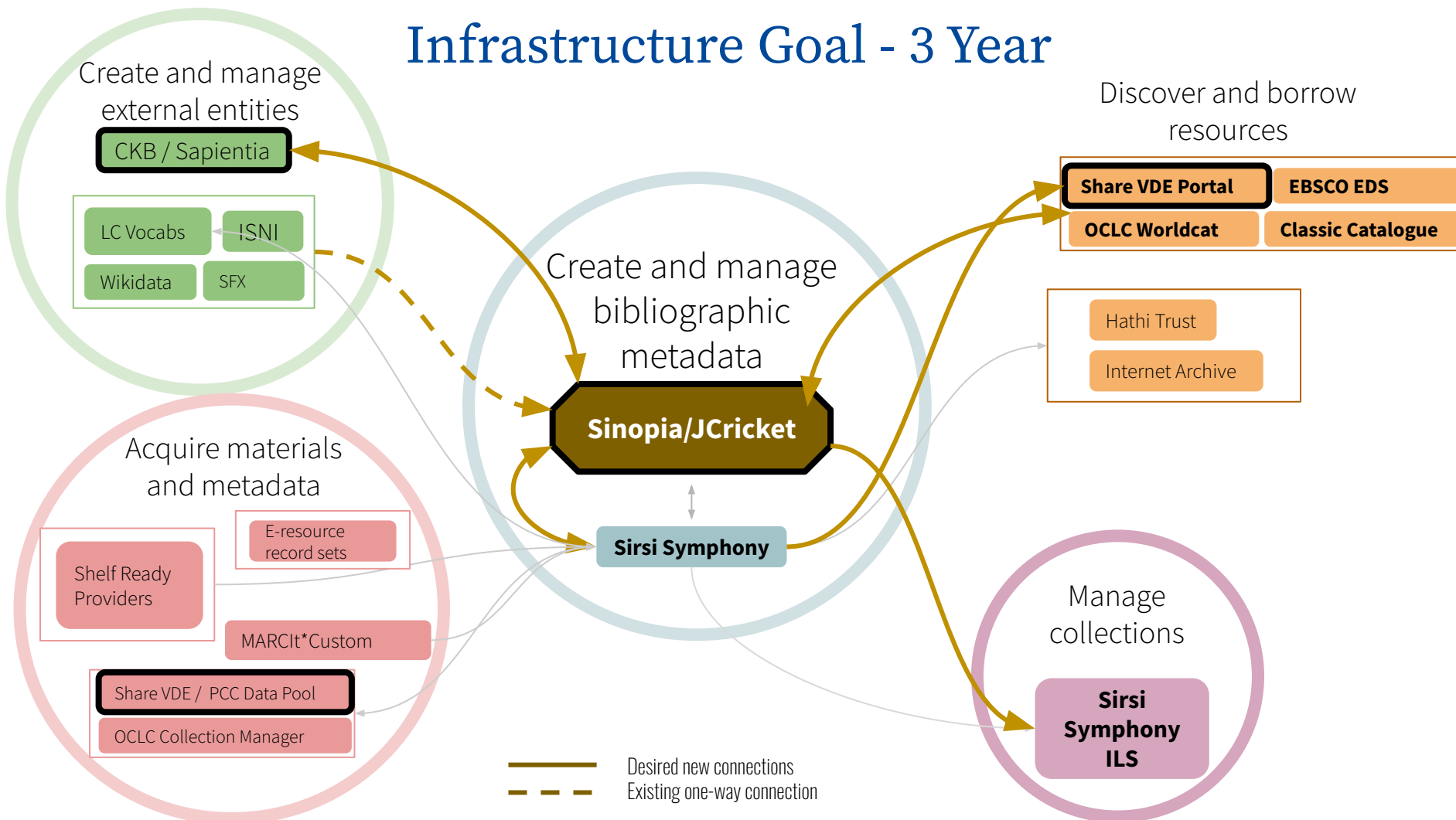# LDIP and Share-VDE

Associated Projects and Initiatives
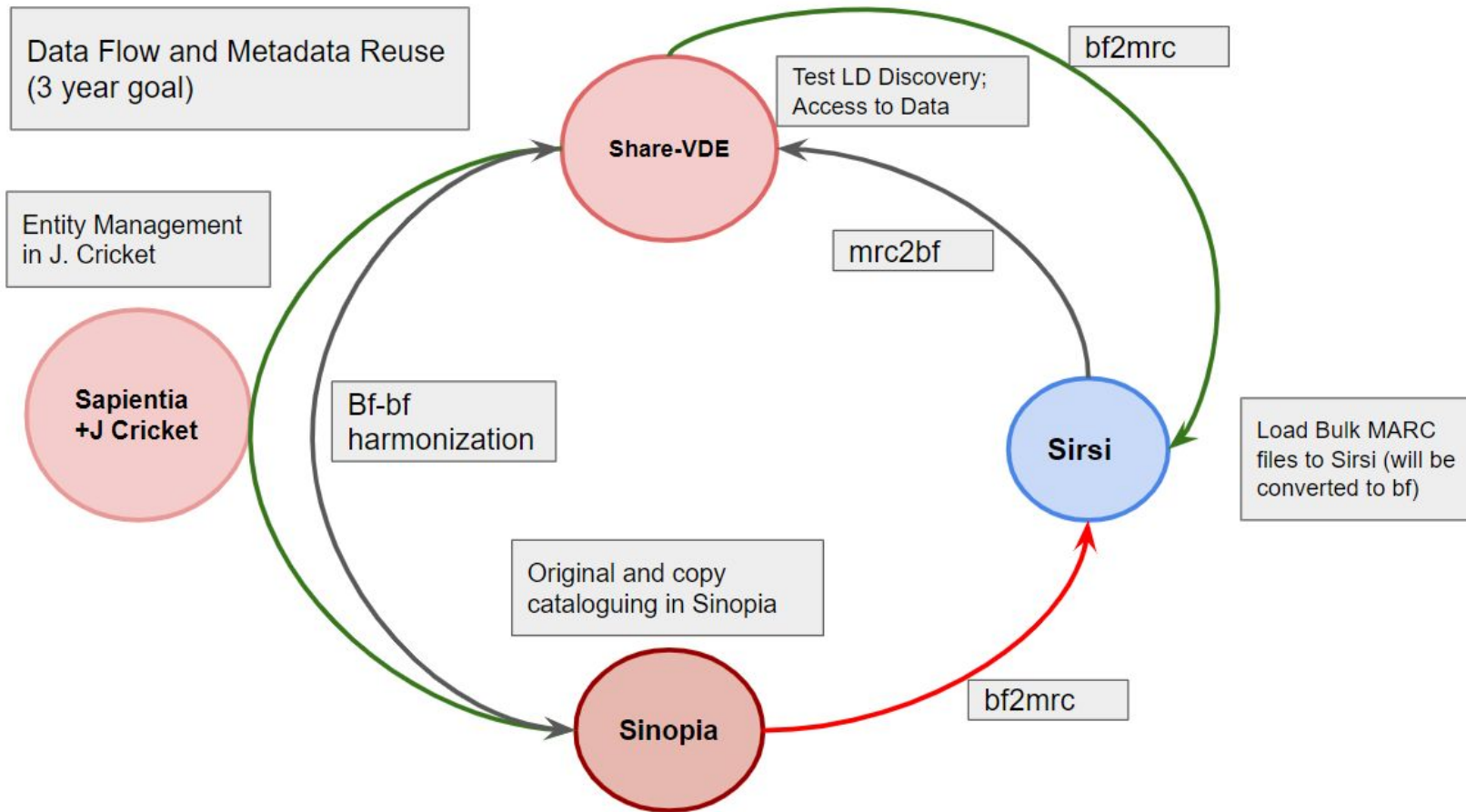
- UAL/NEOS Share-VDE Portal Instance
- MARC2BIBFRAME Conversion
  - Share-VDE Sapientia Cluster Knowledge Base (Stardog triplestore backend) and Associated Conversion and Enrichment Processes
- Authority Control/Entity Management Workflows
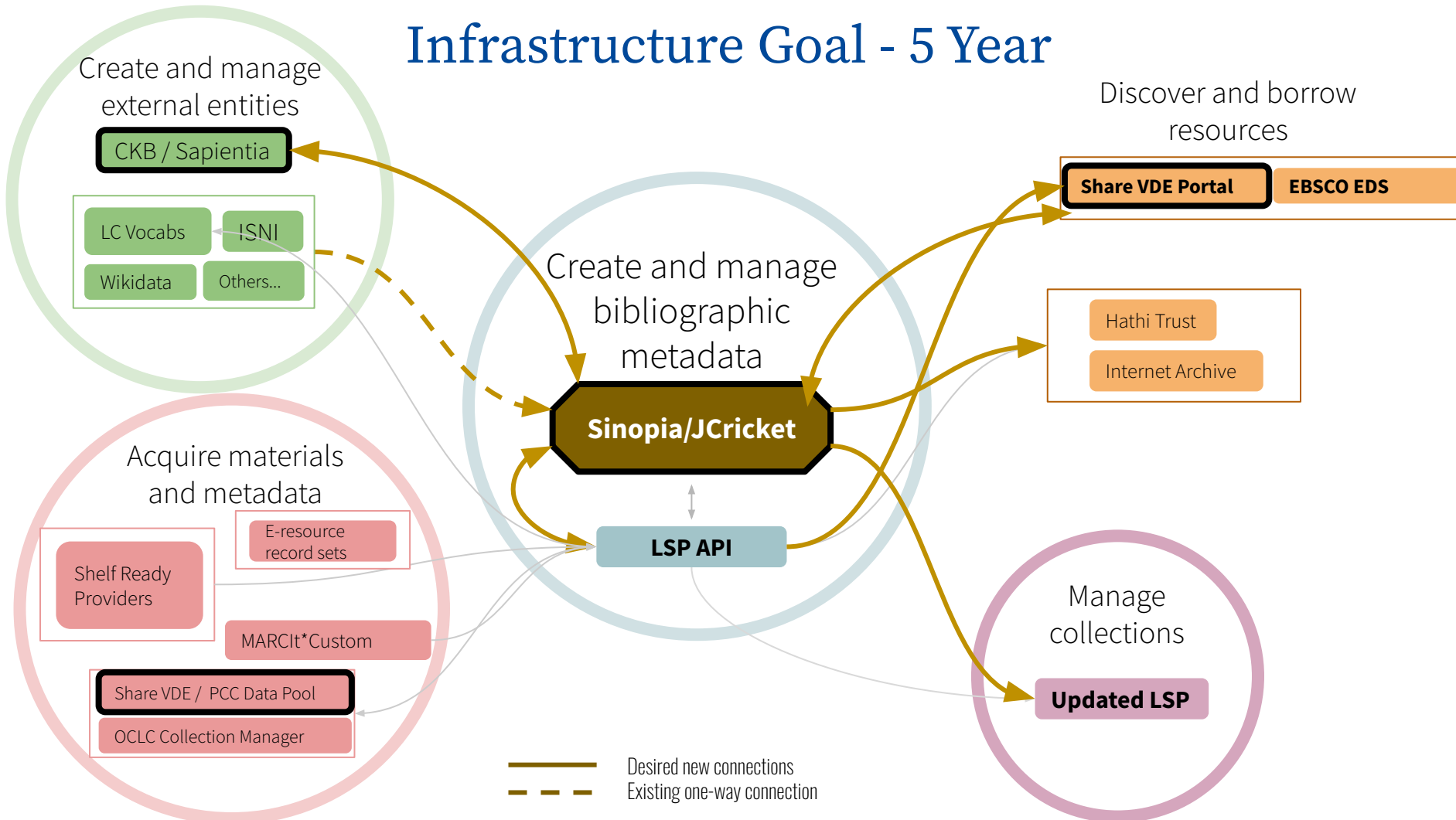 - Share-VDE J. Cricket Cluster Knowledge Base Editor
- PCC BIBFRAME Data Store

Infrastructure Goal - 3 Year

Create and manage external entities
- CKB / Sapientia
- LC Vocabs
- ISNI
- Wikidata
- SFX

Acquire materials and metadata
- E-resource record sets
- Shelf Ready Providers
- MARCIt*Custom
- Share VDE / PCC Data Pool
- OCLC Collection Manager

Create and manage bibliographic metadata
- Sinopia/JCricket
- Sirsi Symphony

Discover and borrow resources
- Share VDE Portal
- EBSCO EDS
- OCLC Worldcat
- Classic Catalogue
- Hathi Trust
- Internet Archive

Manage collections
- Sirsi Symphony ILS

— Desired new connections
--- Existing one-way connection

Support Through Transition

Data Flow and Metadata Reuse (3 year goal)

Entity Management in J. Cricket

Sapientia +J Cricket

Share-VDE

Test LD Discovery; Access to Data

bf2mrc

mrc2bf

Bf-bf harmonization

Sirsi

Load Bulk MARC files to Sirsi (will be converted to bf)

Original and copy cataloguing in Sinopia

Sinopia

bf2mrc

Infrastructure Goal - 5 Year

Create and manage external entities

CKB / Sapientia

LC Vocabs · ISNI

Wikidata · Others...

Discover and borrow resources

Share VDE Portal · EBSCO EDS

Create and manage bibliographic metadata

Sinopia/JCricket

Hathi Trust

Internet Archive

Acquire materials and metadata

E-resource record sets

Shelf Ready Providers

MARCIt*Custom

Share VDE / PCC Data Pool

OCLC Collection Manager

LSP API

Manage collections

Updated LSP

— Desired new connections

- - - Existing one-way connection

Search for people, original works and publications

e.g. *Jules Verne* or *Around the world in 80 days*

Advanced search

Explore the library data of the 20+ Share-VDE member libraries

Learn more

National Bibliography Tenant:
British National Bibliography

Alan Danskin
Collection Metadata Standards Manager
alan.danskin@bl.uk

British National Bibliography (BNB)

The British National Bibliography lists:

- books and journals published or distributed in the United Kingdom and Ireland since 1950, including electronic publications (since 2013)

- Cataloguing-in-Publication (CIP) information for forthcoming book titles.

It supports:

- collection development

- research

- copy cataloguing
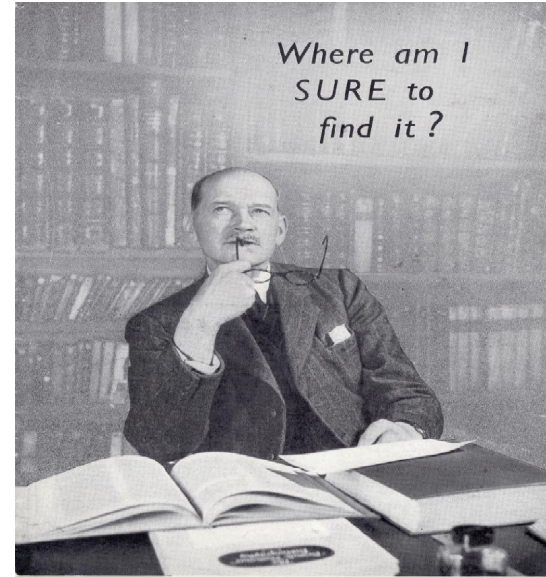
BNB: a timeline

Print

Card service (1956)

UKMARC Exchange Tapes (1969)

BLAISE (online) (1977)

BNB on CD-ROM (1989)

FTP (1998)

Web (2008) / Z39.50 (2009)

.pdf (2010)

Linked open Data (2011)

Share Family

British Library Data Model

New rec
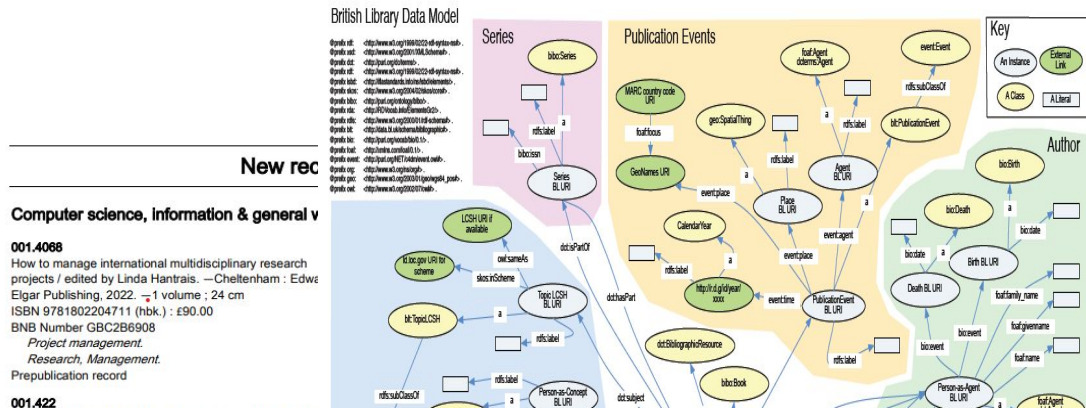
Computer science, information & general v

001.4068
How to manage international multidisciplinary research
projects / edited by Linda Hantrais. —Cheltenham : Edwa
Elgar Publishing, 2022. — 1 volume ; 24 cm
ISBN 9781802204711 (hbk.) : £90.00
BNB Number GBC2B6908
    Project management.
    Research, Management.
Prepublication record

001.422

BRITISH NATIONAL BIBLIOGRAPHY BETA

Search for people, original works and publications

Publications   Original works   People and organizations

e.g. small island andrea levy

Language   Year   Type

Advanced search

Explore linked open data for books and journals published or distributed in the United Kingdom
and Ireland since 1950, and electronic publications since 2013

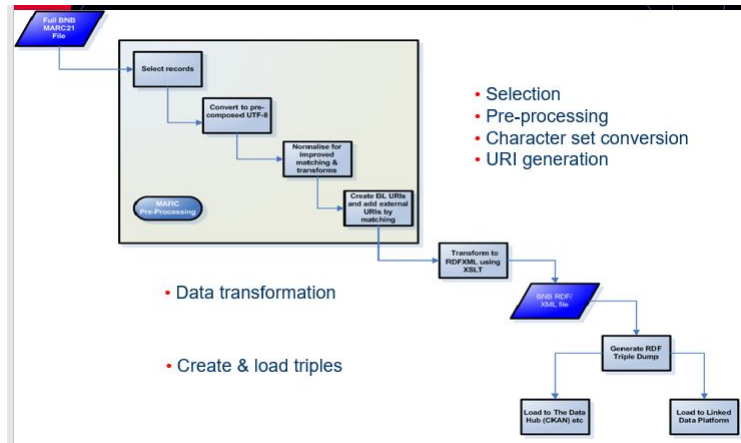← Back to British Library   Learn more   Options ⌄

BNB & Linked data

Linked data project began in 2010

- to make BNB available as linked open data

- to learn about linked open data

- to develop workflows and production processes from MARC 21 to RDF

Developed in partnership with TALIS

- Launched 2011

- Hosted by Talis then TSO

- No affordable path for future development
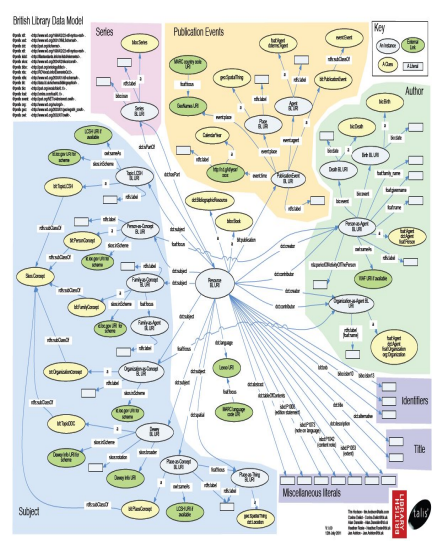
Why we chose Share-VDE?

Share- Family

Affordable

Collaborative development

Cultural Heritage focus

Entified data model

- Authorities

- Bibframe +LRM/RDA

Navigation

APIs

Linking

Development path

BRITISH NATIONAL BIBLIOGRAPHY

BETA

Search for people, original works and publications

Publications   Original works   People and organizations

e.g. small island andrea levy   ⌨   🔍

[Language]   [Year]   [Type]

Advanced search

Explore linked open data for books and journals published or distributed in the United Kingdom
and Ireland since 1950, and electronic publications since 2013

← Back to British Library   Learn more   Options ⌄

About

About the British National Bibliography

About the Share Family National Bibliographies portal

What's new in Share Family National Bibliographies

What is BIBFRAME?

Accessibility

Licenses

Contact

About the British National Bibliography

The British National Bibliography consists of:

- Books and journal titles published or distributed in the United Kingdom and Ireland since 1950, and electronic publications since 2013.
- Cataloguing-in-Publication (CIP) information for forthcoming book titles

Discover more about the British National Bibliography and UK Legal Deposit

Visit British Library Collection Metadata or contact us at metadata@bl.uk

Data summary

Total number of each entity type in the system:

- Person: 1,748,891
- Family: 3,770
- Organization: 415,869
- Conference: 73,983
- Original work: 3,728,084
- Work: 5,742,992
- Instance: 4,975,200
- Item: 0

👤 Person ⓘ

Andrea Levy

English author. Born in 1956 in London (England). Died in 2019.

Andrea Levy was an English author best known for the novels Small Island (2004) and The Long Song (2010). She was born in London to Jamaican parents, and her work explores topics related to British Jamaicans and how they negotiate racial, cultural and national identities. –
Wikipedia

More options ⌄

⚙ **Original works by**   🔵 **Publications by**   ⚙ **Original works about**

Filter original works...   🔍   Genre

9 results

No.	Title			Year	Genre	Creators

9 results

No.	Title	Year	Genre	Creators	⬡
1	⊕ Fruit of the lemon		Fiction	👤 Andrea Levy (author)	
2	⊕ Long song		Fiction/Novel	👤 Andrea Levy (author)	
3	⊕ Fruits of the lemon		Fiction	👤 Andrea Levy (author)	
4	⊕ Small island		Fiction/Novel	👤 Andrea Levy (author)	
5	⊕ Uriah's war		Fiction	👤 Andrea Levy (author)	
6	⊕ Never far from nowhere		Fiction	👤 Andrea Levy (author)	
7	⊕ Every light in the house burnin		Fiction	👤 Andrea Levy (author)	
8	⊕ Six stories and an essay		Short stories	👤 Andrea Levy (author)	
9	⊕ Six stories & an essay		Short stories	👤 Andrea Levy (author)	

Filter publications...

| Publication year | Publication place |

3 results

No.	Title	Language	Publication year	Format	Contributors	⬡
1	● Fruit of the lemon	English	2006	Volume	● Andrea Levy (author) ● Charnwood (publisher)	
2	● Fruit of the lemon	English	2000	Volume	● Andrea Levy (author) ● Review (publisher)	
3	● Fruit of the lemon	English	1999	Volume	● Andrea Levy (author) ● Review (publisher)	

End of list

● Publication ⓘ

Fruit of the lemon

Volume. Published in English in 2006 in Leicester by Charnwood.

ISBN: 1846171229
BNB: GBA586210
Identifier: p1891653381779206
Notes: ill. ;
Acquisition term: £18.99
Bibliographic level: Monograph
Cataloging source: StDuBDS
Descriptive cataloging form: AACR
Dimension: 24 cm.
Edition statement: Large print ed.
Encoding level: Full
Extent: 413 p. (large print) :
Record status: Increase in encoding level from prepub
Responsibility statement: Andrea Levy.

Explore the BNB at the British Library ↗

Copy URI for this page

https://natbib-lod.org/publications/1891653381779206   **Copy**

Download the page in these formats

.json

MARC (binary)

MARC (XML)

RIS

● **Related agents**   ● Subjects

Filter agents... 🔍

Benefits of being a family member

Collaborative development path

State of the art technology stack

Entifying and clustering

Dissemination of our data

Aggregation of data

SHARE Catalogue: LOD collective catalogue

Shared library services: mutual recognition of the respective institutional users as internal users to access services at universities

SHARE Discovery: federated search system to access information on physical and digital collections

SHARE Press (Books/Journals/Open Archive): common platforms for open access digital publishing

SHARE Catalogue: LOD collective catalogue

2017 → SHARE Catalogue Author ID

This person in
isni
WIKIDATA
LIBRARY OF CONGRESS
WorldCat Identities
data.bnf.fr
VIAF

☐ **Gibson, William**
ID: 149714

William Gibson (Q188987)

American-Canadian speculative fiction novelist and founder of the cyberpunk subgenre
William Ford Gibson

SHARE Catalogue author ID   149714
P3987

Why Wikidata?

[...] it is the richness of the data that makes Wikidata unique. Many statements come with provenance information or include additional context data, such as temporal validity [...]

Erxleben et al., *Introducing Wikidata to the Linked Data Web*, International Semantic Web Conference 2014, p. 50

SHARE Catalogue 1.0 → SHARE Catalogue 2.0

Basic tasks for the technical transition:

1. Updated UNIMARC-BIBFRAME mapping

2. Cluster Knowledge Base duplications analysis

3. Additional authority services analysis

1. UNIMARC → BIBFRAME

UNIMARC	UNIMARC Subfield	MARC21	MARC21 Subfield	Conversion 1	Conversion 2
500 - PREFERRED ACCESS POINT	$a - Preferred access point (NR)	130 - Main Entry-Uniform Title (NR)	$a - Uniform title (NR)	## - mainTitle - literal	
500 - PREFERRED ACCESS POINT	$b - General Material Designation (R)	130 - Main Entry-Uniform Title (NR)	$h - Medium (NR)	W - genreForm - GenreForm - rdfs:label- literal	
500 - PREFERRED ACCESS POINT	$h - Number of Section or Part (R)	130 - Main Entry-Uniform Title (NR)	$n - Number of part/section of a work (R)	## - partNumber - literal	
500 - PREFERRED ACCESS POINT	$i - Name of Section or Part (R)	130 - Main Entry-Uniform Title (NR)	$p - Name of part/section of a work (R)	## - partName - literal	
500 - PREFERRED ACCESS POINT	$k - Date of Publication (NR)	130 - Main Entry-Uniform Title (NR)	$f - Date of a work (NR)	W - originDate - literal	
500 - PREFERRED ACCESS POINT	$l - Form Subheading (R)	130 - Main Entry-Uniform Title (NR)	$k - Form subheading (R)	W - natureOfContent - literal	W - genreForm - GenreForm - rdfs:label- literal
500 - PREFERRED ACCESS POINT	$m - Language (when part of a heading) (NR)	130 - Main Entry-Uniform Title (NR)	$l - Language of a work (NR)	W - language - Language - rdfs:label - literal	W - translationOf
500 - PREFERRED ACCESS POINT	$n - Miscellaneous Information (R)	130 - Main Entry-Uniform Title (NR)	$g - Miscellaneous information (R)	W - genreForm - GenreForm - rdfs:label- literal	
500 - PREFERRED ACCESS POINT	$q - Version (or Date of Version) (NR)	130 - Main Entry-Uniform Title (NR)	$s - Version (R)	W - version - literal	
500 - PREFERRED ACCESS POINT	$r - Medium of Performance (for Music) (R)	130 - Main Entry-Uniform Title (NR)	$m - Medium of performance for music (R)	W -musicMedium - MusicMedium - rdfs:label - literal	

International Federation of Library Associations and Institutions

UNIMARC Bibliographic (3rd ed.) Updates

The Library of Congress > BIBFRAME > MARC 21 to BIBFRAME 2.0 Conversion Specifications

BIBFRAME

bf: BIBFRAME

- Home
- Frequently asked questions
- Webcasts & presentations
- Contact us

Model & Vocabulary
- BIBFRAME model
- BIBFRAME vocabulary
  - Category view
  - List view
  - RDF View
- Extension list view

🖨 Print 📶 Subscribe 🔗 Share/Save 💬 Give Feedback

MARC 21 to BIBFRAME 2.0 Conversion Specifications

The following specifications were developed to support the use of BIBFRAME 2.0 at the Library of Congress. They map the conversion of MARC Bibliographic record fields to BIBFRAME Hub, Work, Instance and Item.

MARC Bibliographic Conversion Specifications
- Fields 001-007 – Control, physical description (Excel, 40 KB, 06/22/2022)
- Fields 006, 008 - Fixed fields (Excel, 33 KB, 06/22/2022)
- Fields 010-048 – Identifiers, etc. (Excel, 24 KB, 06/22/2022)
- Field 048 - Instrument codes (Excel, 14 KB, 10/21/2021)
- Fields 050-088 - Class/call numbers, etc. (Excel, 14 KB, 10/21/2021)
- Fields 1XX, 7XX, 8XX – Names (Excel, 13 KB, 10/21/2021)
- Fields 200-24X, except 240 – Titles (Excel, 14KB, 10/21/2021)
- Fields 240, X30, etc. – Uniform titles (Excel, 13 KB, 10/21/2021)
- Fields 250-270 – Edition, imprint, etc. (Excel, 14 KB, 10/21/2021)
- Fields 3XX – Physical description, etc. (Excel, 27 KB, 06/22/2022)

Usage Notes

These specifications are written from the perspective of MARC so that each element in MARC is considered, even if not converted. The specifications are presented in Microsoft Excel files accompanied by a few explanatory specifications in Microsoft Word documents. If there is little or no use of an element in the Library of Congress records, the specifications usually say "nac" (no attempt to convert) in the conversion column. MARC fields and subfields that have been recently approved for the MARC21 format might be listed as "nac" initially, but the specifications may be modified as these fields begin to be used. Some elements in a MARC record are not relevant outside the MARC environment, and they are simply marked "ignore."

UNIMARC 4xx	Fields 76X-78X - Linking Entries - R0	Conversion 1
454	765 - ORIGINAL LANGUAGE ENTRY (R)	W - translationOf
453	767 - TRANSLATION ENTRY (R)	W - translation
421	770 - SUPPLEMENT/SPECIAL ISSUE ENTRY (R)	W - supplement
422	772 - SUPPLEMENT PARENT ENTRY (R)	W - supplementTo
461	773 - HOST ITEM ENTRY (R)	W - partOf
462	774 - CONSTITUENT UNIT ENTRY (R)	W - hasPart
451	775 - OTHER EDITION ENTRY (R)	W - otherEdition
452	776 - ADDITIONAL PHYSICAL FORM ENTRY (R)	I - otherPhysicalFormat
423	777 - ISSUED WITH ENTRY (R)	W - issuedWith
	780 - PRECEDING ENTRY (R)	If no valid value in ind2 then W - precededBy
	Indicators	
	Second - Type of relationship	
430	0 - Continues	W - continues
431	1 - Continues in part	W - continuesInPart
432	2 - Supersedes	W - precededBy
433	3 - Supersedes in part	W - precededBy
436	4 - Formed by the union of ... and ...	W - mergerOf

2. Wikidata/SHARE Catalogue ID Constraint violations

"Single value" violations [modifica]

Violations count: 2055

- Q38: 423419, 440885
- Q39: 606889, 361835
- Q272: 175093, 596216
- Q517: 405281, 754018, 688872, 688469, 466838, 560074
- Q859: 292329, 625353
- Q1001: 161878, 627256
- Q1048: 159728, 544911
- Q1065: 464492, 566139
- Q1398: 727867, 7260
- Q1405: 625431, 425322
- Q1407: 391602, 397764
- Q2038: 778213, 345598
- Q2098: 165201, 626783
- Q2494: 181463, 476795
- Q2677: 287847, 617926
- Q3044: 400582, 452688
- Q4214: 189651, 580637
- Q4289: 171423, 548840

"Unique value" violations [modifica]

Violations count: 24

- 123267: Q2891989, Q112669549
- 132973: Q356773, Q112624839
- 140698: Q18649413, Q112064518
- 143530: Q16105128, Q111693582
- 147207: Q62559671, Q64697727
- 169412: Q6178781, Q111077232
- 207175: Q59525874, Q111650804
- 219994: Q6327, Q111449951
- 222755: Q57265654, Q111677378
- 250913: Q669283, Q64878551
- 279511: Q50849713, Q102351007
- 304244: Q87066619, Q87066628
- 325056: Q19365657, Q109728558
- 330879: Q38792463, Q110597895
- 409684: Q522014, Q110678335
- 49016: Q975609, Q107546834
- 523500: Q3850989, Q112151783
- 535707: Q62862211, Q102086676
- 548667: Q15053360, Q18734860
- 57250: Q19586070, Q102123188

CKB Cleaning Process (through Wikidata)

3. More identifiers!

Label	Wikidata ID	Occupation	SHARE Cat ID	IRIS UNINA ID	VIAF ID	ORCID iD	SBN author ID	IdRef ID
Nicola Fusco	Q1477114	mathematician university teacher	9279	03887	14993937	0000-0001-8215-8351	CFIV333796	070707677
Marcello Gigante	Q1893332	papyrologist classical philologist university teacher hellenist byzantinist	36239	12069	109998667		CFIV014251	028370058
Alberto Izzo	Q2831772	architect	435931	11106	12310157		CFIV029247	026932261
Gabriele Frasca	Q3094072	writer poet	169992	23328	118224903		CFIV077975	029501717
Adriana Valerio	Q3605679	theologian	154848	16145	118613352		CFIV000465	050638513
Andrea Amatucci	Q3615474	jurist	113320	30502	165477041		BVEV009660	085227242
Attilio Stazio	Q3629300	numismatist	37574	22238	58934972 139146822010107380339		CFIV083256	193718847 069271003
Benedetto Conforti	Q3638066	magistrate politician	132949	05176	66557320		CFIV013212	033417458
Carlo Sbordone	Q3659842	mathematician university teacher	437300 333428	31166	112144648393318710774		CFIV036784	238107876
Corrado Beguinot	Q3694189	engineer urban planner	2172	24442	75327251		CFIV063270	115632239
Francesco Barbagallo	Q3749596	historian	308463	22888	109318158		CFIV043290	029319730
Franco Archibugi	Q3751120	economist planning theorist	32146 437786	12869	24774382		LO1V042841	053516893
Gianfranco Pasquino	Q3763131	political scientist politician university teacher	124397	24109	273668629		CFIV005136	028837509
Gianni Dal Maso	Q3763488	mathematician university teacher	42309	30757	24690067	0000-0002-1010-4968	MILV108896	033298920
Giuseppe Cacciatore	Q3770188	philosopher	5242	02934	114594052		CFIV014067	02999361X
Guido Trombetti	Q3779487	mathematician university teacher politician	6404 437319	21540	304973015		CFIV031723	

This list ⬆ **https://w.wiki/45RJ**   **SHARE Cat + IdRef** ➡ **https://w.wiki/45gA**

Work in progress...

Discussion/Questions

Useful resources

Wiki website to learn more https://wiki.svde.org/

About the Share Family

Share Family news and updates

Presentations and dissemination resources

The Working Groups

For more info contact info@svde.org