

Closing the circle: using linked data to enrich authority records and make them linkable on the web

Roxana Popistaşu

National Library of Luxembourg

IT department

Stefanie Zutter

National Library of Luxembourg

Cataloguing & Indexing



Bibliothèque nationale de Luxembourg

Nationalbibliothéik

IGeLU Conference

Oxford, September 2014

National Library of Luxembourg (BnL)

Legal missions

Heritage library

- collect, catalogue, preserve, **enrich** the national heritage

Scientific and research library

- collect, catalogue, preserve, **enrich** non-Luxemburgensia collections

Provide widespread access to all collections

- loans and remote consultation

Manage the network of Luxembourgish libraries

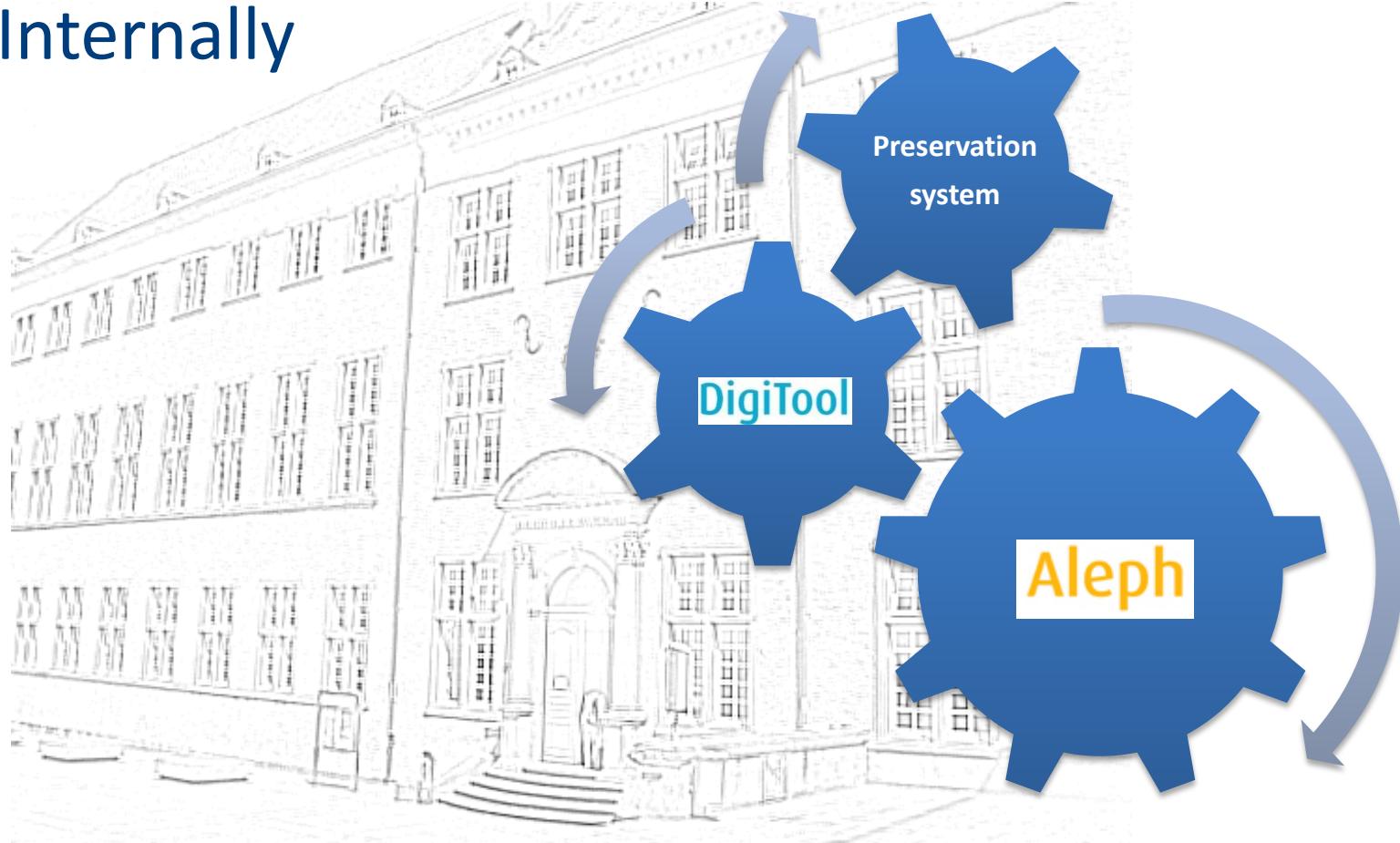
- manage the Union catalogue and the library network

Contribute to the development of library science nationally and internationally



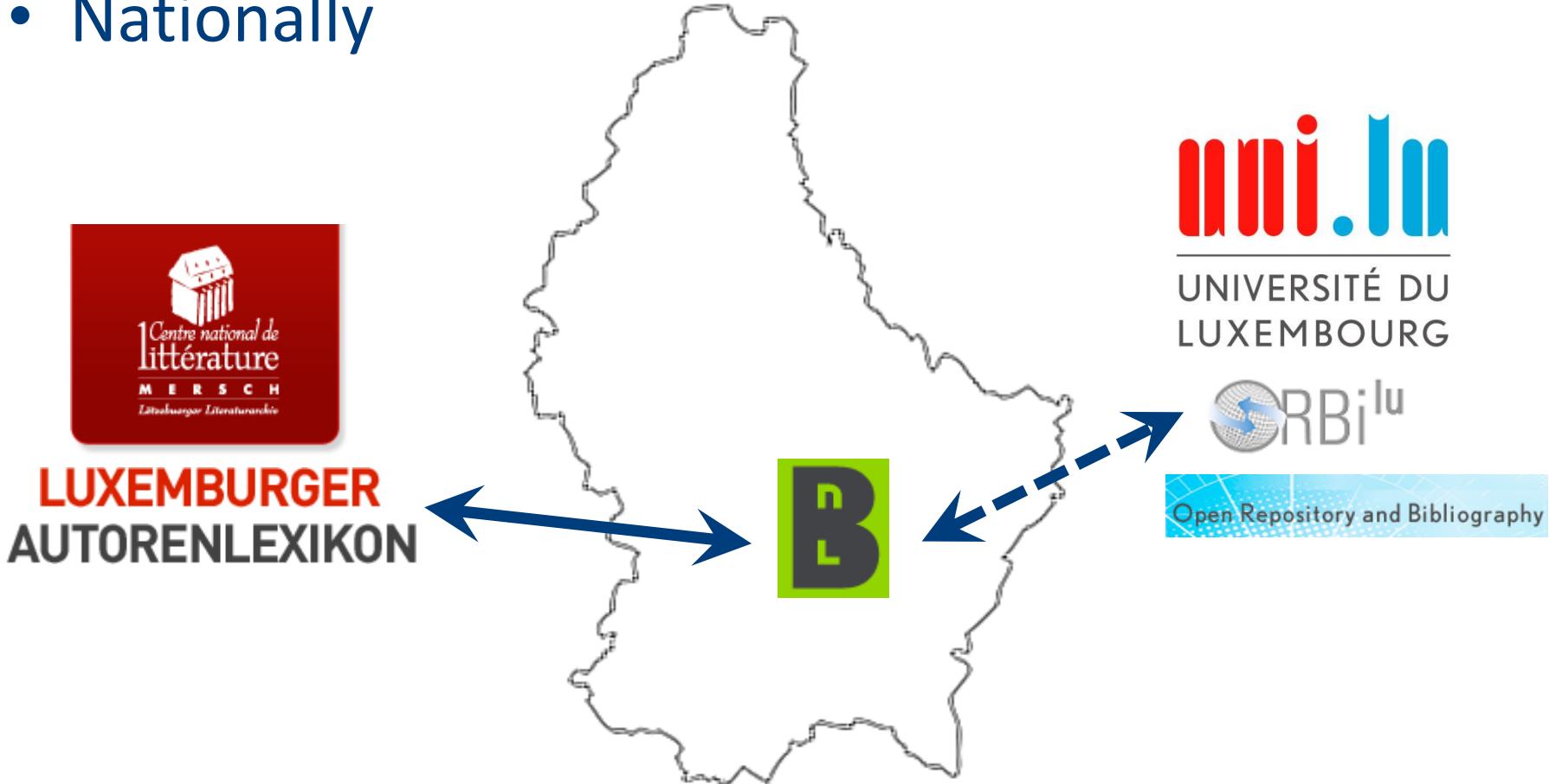
Linking BnL Data...

- Internally



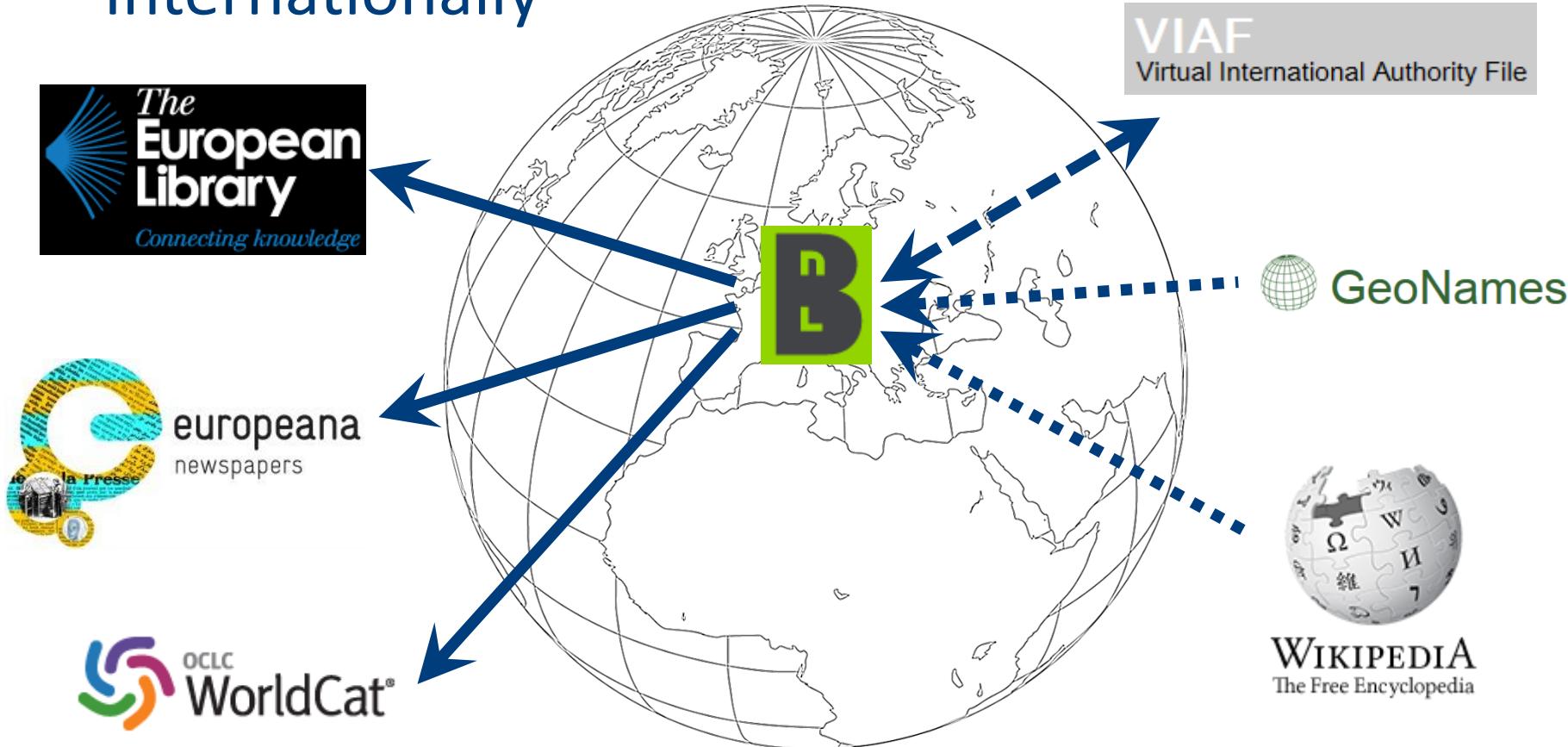
Linking BnL Data...

- Nationally



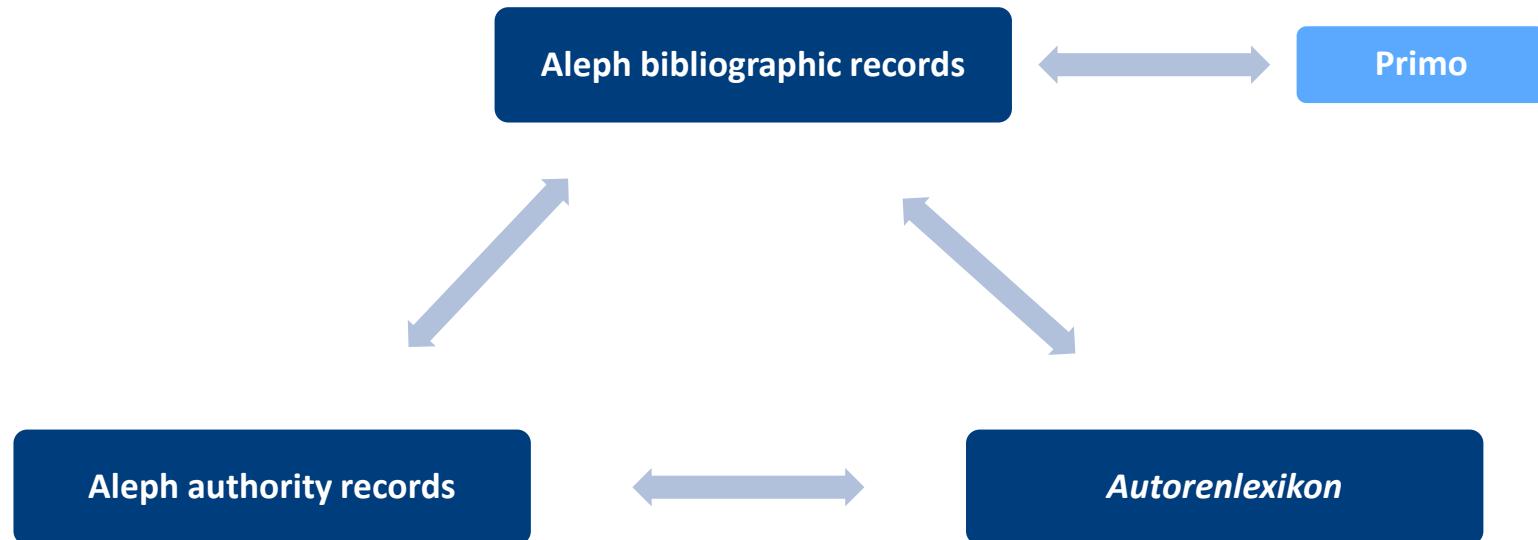
Linking BnL Data...

- Internationally



Linking authors (LiDa Project)

Create links between the authors in Aleph (bibliographic & authority records) and their corresponding authors in *Autorenlexikon*, an encyclopaedia of Luxembourgish literary authors



LiDa Project

Project partners



National Library of Luxembourg

- Manages ALEPH & PRIMO for the network of libraries in Luxembourg



Centre National de Littérature (National Centre for Literature)

- Manages the content of *Autorenlexikon*



Magic moving pixel

- IT management side for *Autorenlexikon*

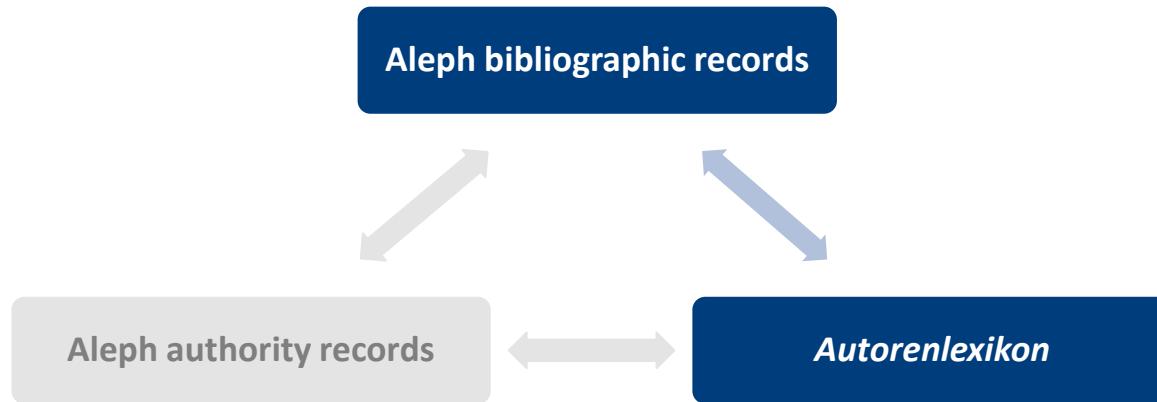


Bibliothèque nationale de Luxembourg
Nationalbibliothéik

IGeLU Conference
Oxford, September 2014

LiDa Project - Phase 1

- Create links between the authors in Aleph and their corresponding authors in *Autorenlexikon* using ***contextual string matching*** (author names + associated bibliographic information)
- Not possible (yet) to use authority records



LiDa Project - Phase 1

More info:

- Presentation at IGeLU 2013 in Berlin:
[*Linked Data: From ALEPH/PRIMO to the Dictionary of Luxembourgish Authors and back*](#)
- ELCommons Linked Open Data Use Cases and Scenario:
[*Two-way links between authors in Primo & Aleph and Autorenlexikon, an encyclopaedia of Luxembourgish literary authors*](#)



LiDa Project - Phase 2

- Create links between the authors in the Aleph authority file and their corresponding authors in *Autorenlexikon* using existing links
- **Part of a bigger project – *Authority file:***
 - “clean up” existing authority file
 - enrich existing authority records
 - create new authority records

Authority file project

Project areas

➤ *Organisational*

- Establish a collaboration framework between BnL and experts from other Luxembourgish organisations

➤ *Cataloguing*

- Create new cataloguing template in line with international standards and suitable for a VIAF export

➤ *Technical*

- Develop tools for validating and enriching authority records



The librarian's view

Build initiatives for an enhanced catalog and discovery layer

- Encourage patron engagement with collection
- Promote Luxembourgish content
- Leader in digital content and user experience



Authority data strategy (1)

Harvest data from a range of processes

- legal deposit
- subject experts
- external databases



Authority data strategy (2)

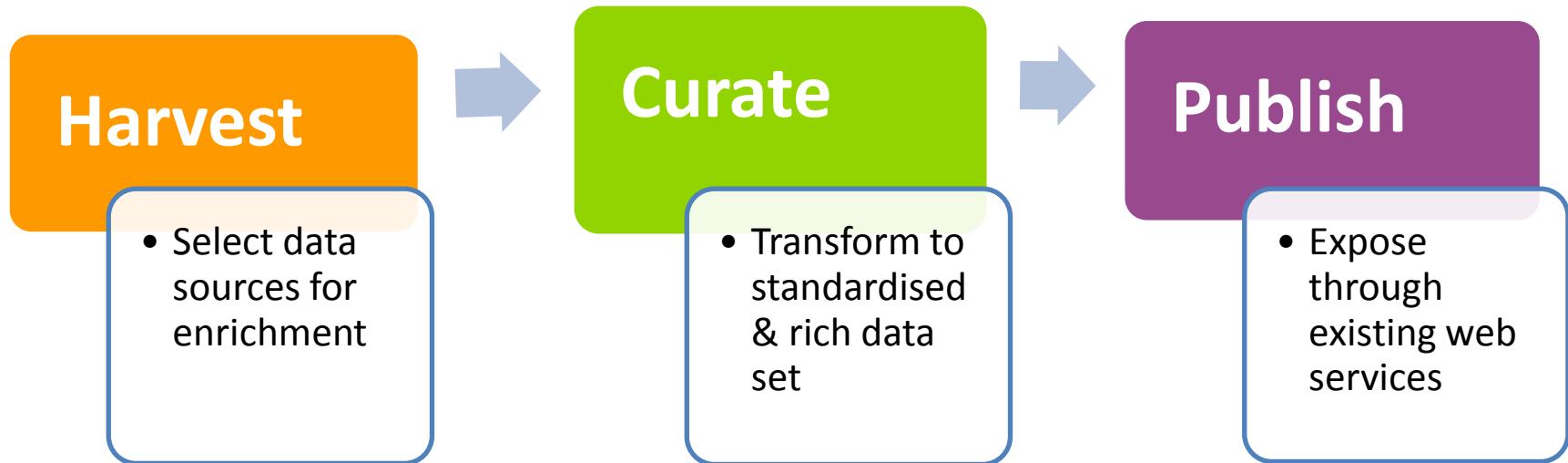
- Make data fit for purpose
- Make data interoperable
 - IDSMARC to MARC21
- Enable semantic web by FRBR-ising authority data
 - AACR2 to RDA

Authority data strategy (3)

- Publish authority data widely
 - BnL discovery tool
 - international aggregators



Authority data management processes



Start small, evolve step by step



Harvest

Prerequisites

- Find the right data source
- Gain buy-in
- Establish framework for trans-institutional collaboration
- Proof Of Concept for documenting policies and procedures fit for implementation



Data sources

Selection criteria

- Copyright free through legal mechanisms such as contracts, waivers, and licenses (CC-0)
- Data quality
- Data privacy issues
- Freedom of Information Act?



Curate (1)

- Dealing with unstructured information
 - Use data profiling methods
- Nature and scale of the project
 - Document and stabilize evolutionary steps

Curate (2)

- Analyse current state of the database
- Define future requirements for authority data
- Design new authority data schema
- Write system specifications to meet new requirements
- Design governance model

Data governance model

- Provenance indicators
- Plan data maintenance lifecycles with creation/editing policies and procedures
- Keep licensing CC-0 to provision for reuse
- Protection of privacy & confidentiality
- Policies for production of derivatives
- Produce documentation on policies and practices used to create these records



AACR2 authority record

LDR	-----nz---22-----o--4500
008	070730---ad---ab--
040	LUX BINA
100	<u>a</u> Ansembourg, Gaston <u>c</u> comte d' <u>a</u> Gaston <u>c</u> comte de Marchant et d'Ansembourg
400	<u>a</u> Ansembourg, Gaston d' <u>a</u> Marchant d'Ansembourg, Gaston-Marie-Amaury-Nicol as de
400	<u>a</u> Gaston <u>c</u> comte d'Ansembourg
400	<u>a</u> Ansembourg, Gaston Amaury Nicolas Marie <u>c</u> comte de Marchant et d'Ansembourg
400	<u>a</u> D'Ansembourg, Gaston

Unique string of characters which represents a particular authority to be traced to bibliographic records



FRBR-isé data scheme

Record status

Access point of
the entity

Attributes of
the entity

Leader	LDR	-----nz-	22-----n--4500
Fix pos data	008	-----080326-----ad-----ab-----a-----	
Origine	040	a LUX BINA c LU-BNL	
Ved nom personn	100 1	a Prüm, André	
FR personne	400 1	a Prum, André	
Source	670	a Site Internet uni.lu b Curriculum vitae, publications u http://staff.uni.lu/andre.prum	
Naissance-mort	046	f 19610806	
Lieu associé	370	a Bridel c Luxembourg f Paris	
Activité	372	a Law	
Profession	374	a Dean of the Faculty of Law, Economics and Finance at the University of Luxembourg	
Genre	375	s 2005 a masculin	

MARC21 authority record

Leader	LDR	1----nz---22-----n--4500
Fix pos data	008	_ _ _ 131230--ad--ab--
Origin	040	__a LU-CELI
H-Personal name	100	__1 __a Gleit, Maria __d 1909-1981 __q Dictionnaire des auteurs luxembourgeois
SF perso name	400	__1 __a Gleitsmann, Herta
SF perso name	400	__1 __a Hofmann-Gleitsmann, Hertha
SF perso name	400	__1 __a Mgl.
Data source	670	__a Dictionnaire des auteurs luxembourgeois
Assoc dates	046	__f 10990228 __g 19810709 __v Dictionnaire des auteurs luxembourgeois
Assoc places	370	__a Crimmitschau __b Locarno __c Luxembourg __f New York __v Dictionnaire des auteurs luxembourgeois __f Berlin
Activity	372	__a Journalistin __s 1938 __t 1939 __v Dictionnaire des auteurs luxembourgeois
Occupation	374	__a Redakteurin
Gender	375	__a féminin

Status

Access points

Preferred name

Individualisation



Bibliothèque nationale de Luxembourg

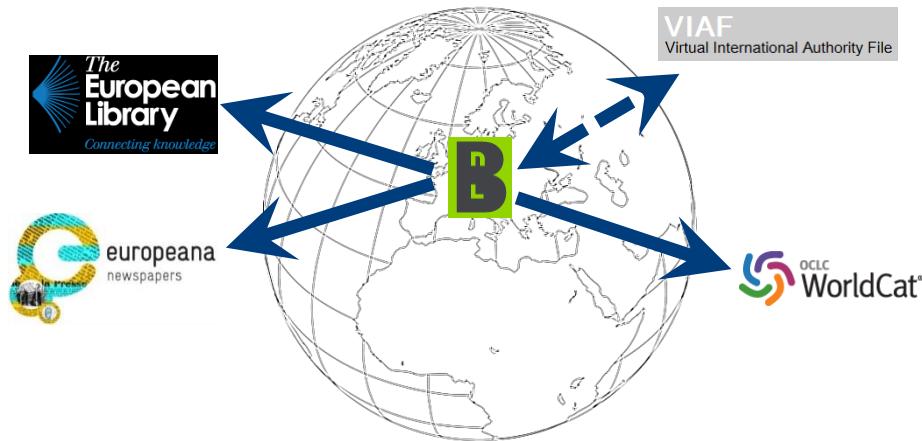
Nationalbibliothéik

IGeLU Conference

Oxford, September 2014

Publish

- Publish authority data widely
 - BnL discovery tool
 - International aggregators



Authority file project

Project areas

➤ *Organisational*

- Establish a collaboration framework between BnL and experts from other Luxembourgish organisations

➤ *Cataloguing*

- Create new cataloguing template in line with international standards and suitable for a VIAF export

➤ *Technical*

- Develop tools for validating and enriching authority records



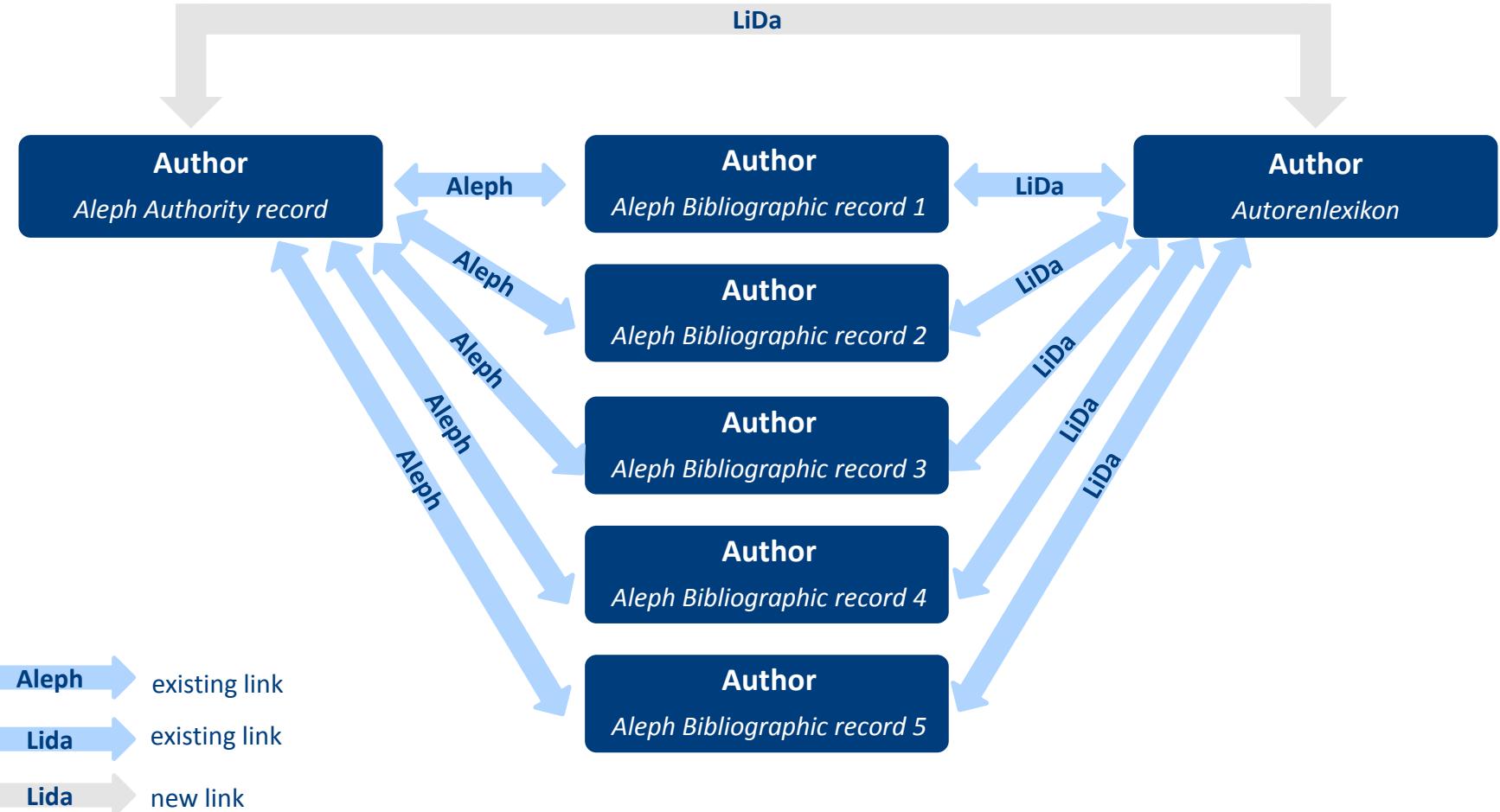
LiDa – Create new links

Match authors from the authority file in Aleph and *Autorenlexikon* using

- existing Aleph links between bibliographic and authority records (system numbers + author headings)
- links created by LiDa between authors bibliographic records and *Autorenlexikon* (author headings + *Autorenlexikon* ID)



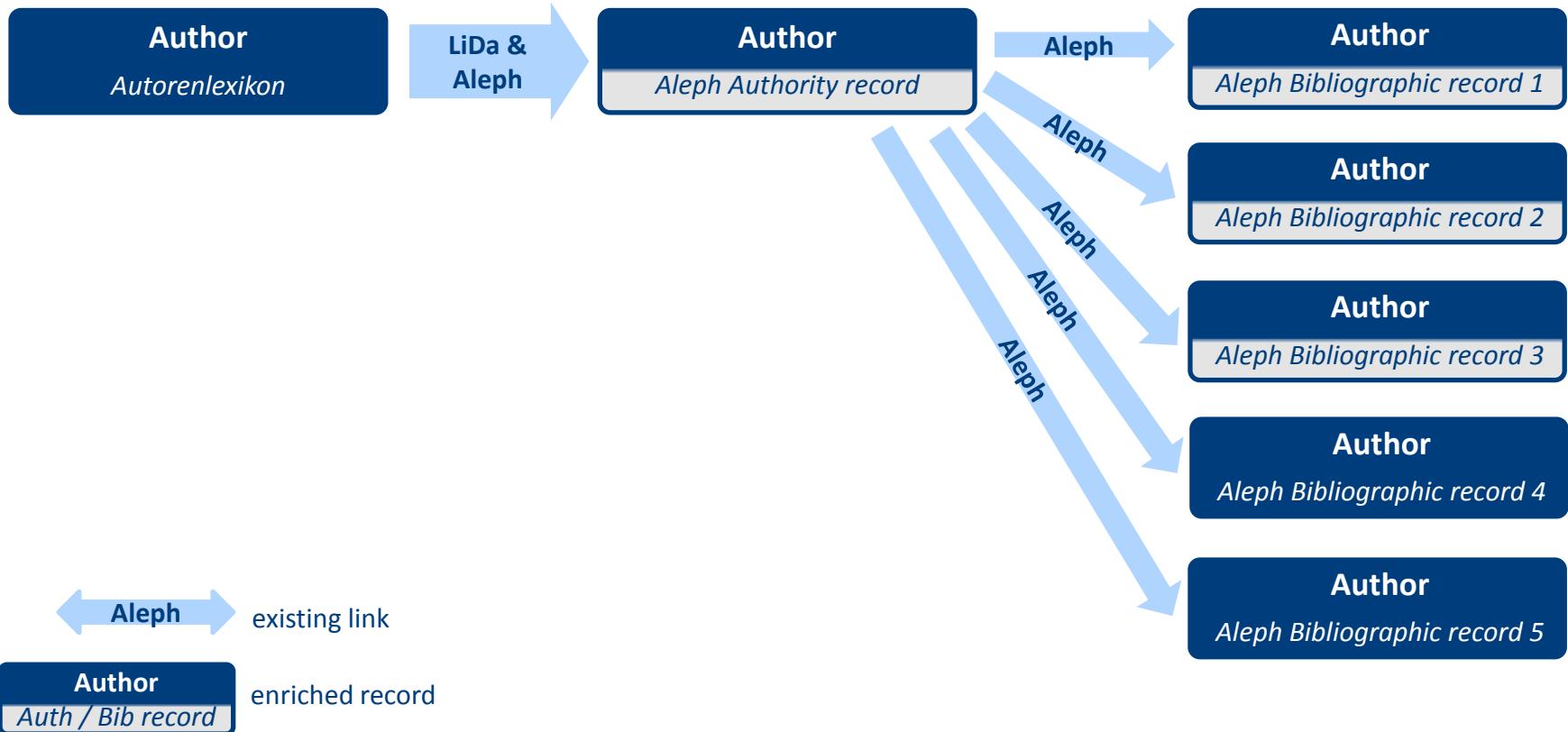
Create new links (transitivity)



Enrich existing authority records

- Add new information from *Autorenlexikon* to corresponding Aleph fields
- Make sure existing headings are kept → no links to bibliographic records are suppressed
- Correct (some) existing fields
- Validate enrichment of records
- Create Aleph sequential file & Load authority records in Aleph (manage_18, ue_08)

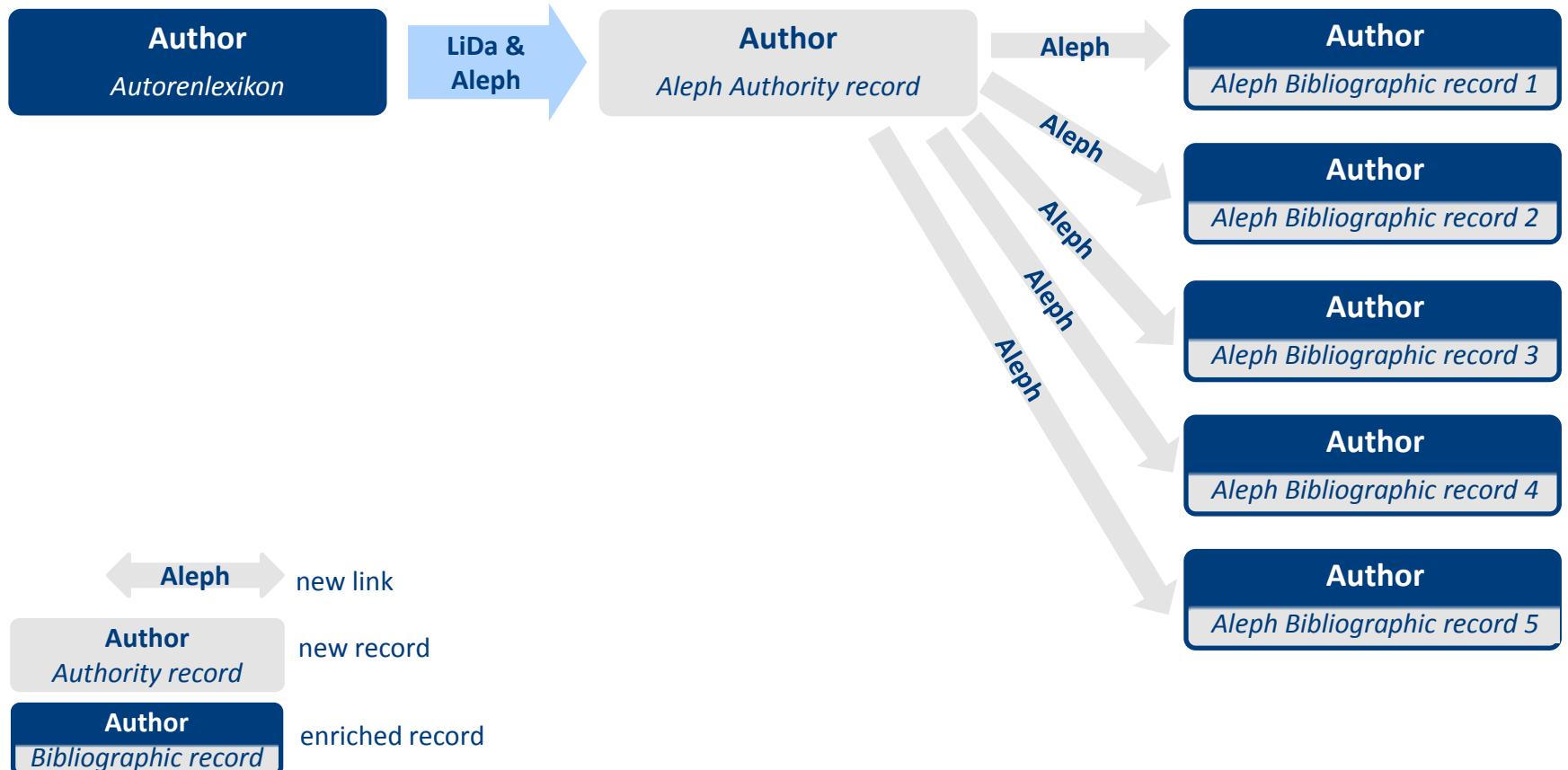
Enrich existing authority records



Load new authority records

- Map *Autorenlexikon* information to Marc21 authority record fields
- Create Aleph sequential file
- Load Authority Records in Aleph (manage_18)
- Create links between the new authority records and existing bibliographic records (ue_08 with “N” / manage_102, manage_02, manage_103)

Load new authority records



Live demo

- Links in Primo
- Links in Aleph
- Links in *Autorenlexikon*
- LiDa validation interface

Authority file – next steps

- Luxembourgish authorities
 - Enrich authority file in the production environment (after validation on the development server) – November 2014
 - Export authority file to VIAF – end 2014
 - Add VIAF IDs to Aleph and *Autorenlexikon* → matching based on IDs – end 2014
 - Collaborate with other Luxembourgish organisations – 2015
- Non Luxembourgish authorities
 - Correct & enrich records using VIAF feedback file

Linked data – new project ideas

- Link authority file authors to:
 - authors in ORBi^{lu} (digital repository of the University of Luxembourg)
 - authors in DigiTool (digitised Luxembourgish newspapers, manuscripts etc.)
 - authors in the new digital preservation system (electronic legal deposit, digital & digitized material)
- Link geographic names in Aleph, DigiTool etc. to GeoNames
- Create links to Wikipedia (DBPedia, WikiData)

Questions, remarks, suggestions?

Thank you!

Roxana Popistaşu

IT department

Roxana.Popistasu@bnl.etat.lu

Stefanie Zutter

Cataloguing & Indexing

Stefanie.Zutter@bnl.etat.lu



Bibliothèque nationale de Luxembourg
Nationalbibliothéik

IGeLU Conference

Oxford, September 2014