



# Content re-use Data re-use Digitisation funding under the PSI Directive

## Perspectives from the National Library of Portugal

Helena Simões Patrício



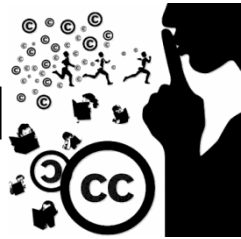
# PUBLIC SECTOR INFORMATION DIRECTIVE

## ❖ DIGITAL CONTENT RE-USE

- ✓ NLP digital collection
- ✓ Access statistics
- ✓ Rights clearance processes and procedures

## ❖ DATA RE-USE

## ❖ DIGITISATION FUNDING



# DIGITAL CONTENT RE-USE: NLP DIGITAL COLLECTION

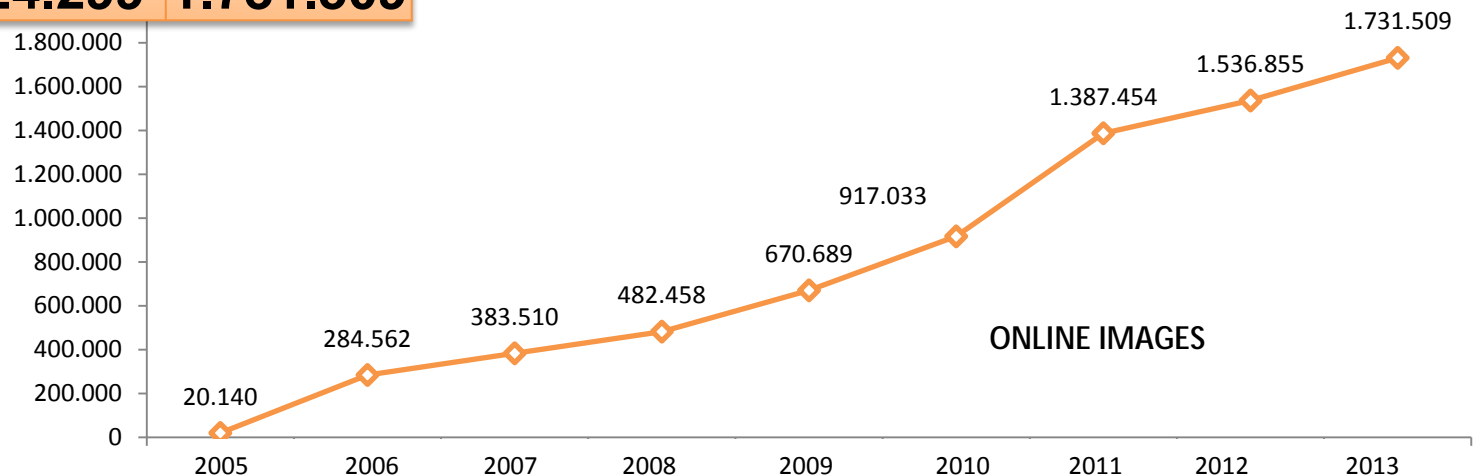
PSI DIRECTIVE STATES THAT...

*“RE-USE OF EXISTING DOCUMENTS HELD BY LIBRARIES  
FOR COMMERCIAL OR NON-COMMERCIAL PURPOSES;  
THROUGH ELECTRONIC MEANS”*

NATIONAL DIGITAL LIBRARY - DIGITAL COLLECTION GROWTH

<http://bnd.bn.pt>

Online content	Items	Images
	24.299	1.731.509



# DIGITAL CONTENT RE-USE: NLP DIGITAL COLLECTION

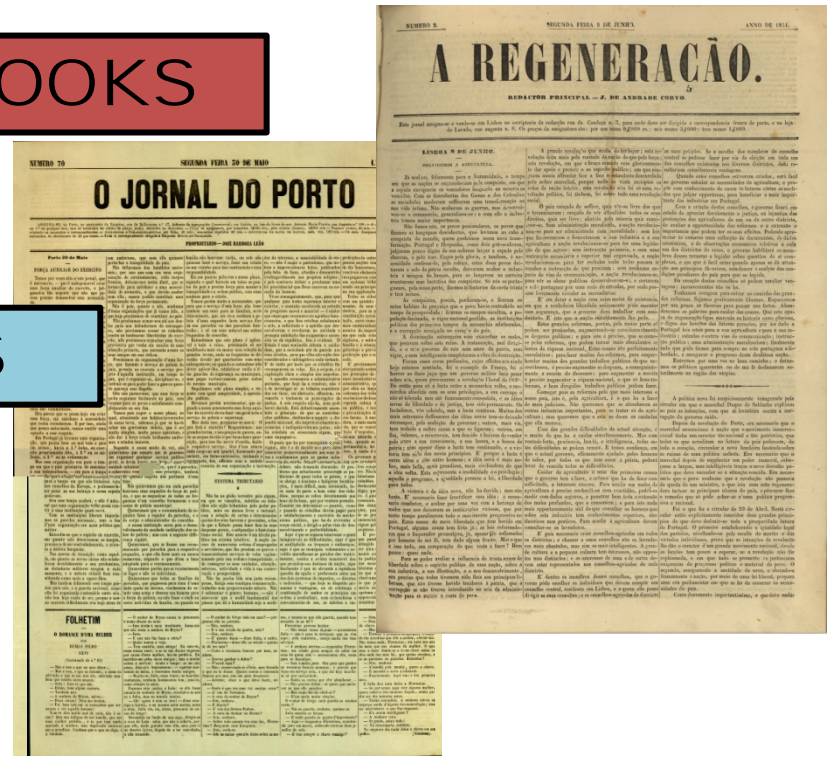


## DIGITAL COLLECTION: TYPE OF RESOURCES (ONLINE IMAGES)

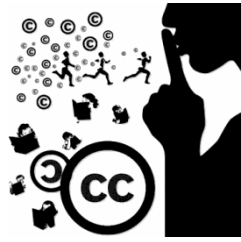
78% BOOKS

19% NEWSPAPERS

1% / 1% 1%





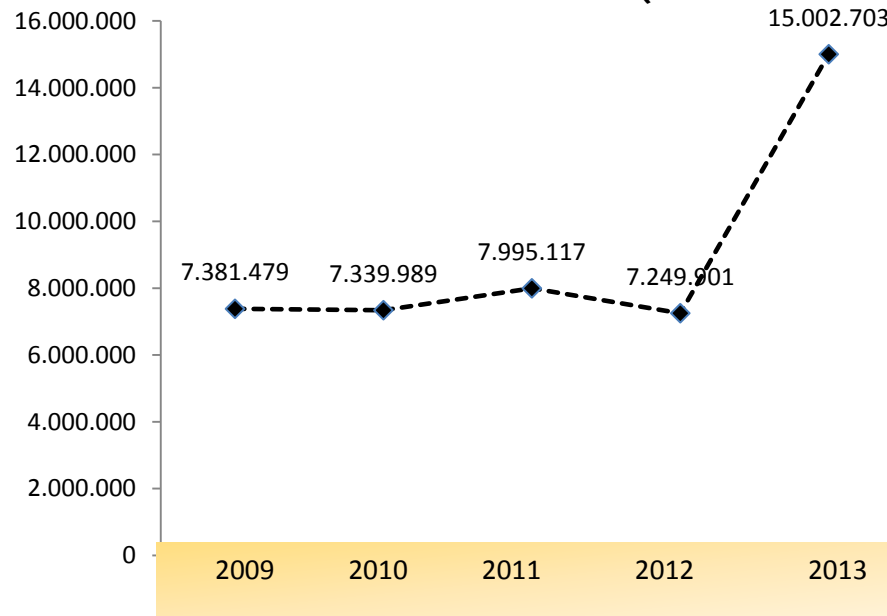


# DIGITAL CONTENT RE-USE: ACCESS STATISTICS

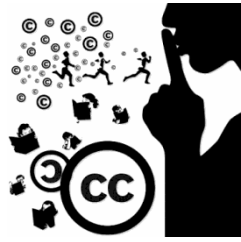
PSI DIRECTIVE STATES THAT...

*“RE-USE OF EXISTING DOCUMENTS HELD BY LIBRARIES FOR COMMERCIAL OR NON-COMMERCIAL PURPOSES; THROUGH ELECTRONIC MEANS”*

N° OF VISITS PER DIGITAL OBJECT (1 VISIT = 20 MINUTES)



# DIGITAL CONTENT RE-USE EXAMPLES





# DIGITAL CONTENT RE-USE: RIGHTS CLEARANCE

PSI DIRECTIVE STATES THAT...

*“THESE OBLIGATIONS SHALL APPLY ONLY INsofar AS THEY ARE COMPATIBLE WITH PROTECTION OF INTELLECTUAL PROPERTY RIGHTS”*

RIGHTS CLEARANCE PROCESSES AND PROCEDURES  
CAN BE EXTREMELY COMPLEX



❖ IT TAKES A LOT OF **TIME** TO DETERMINE THE COPYRIGHT STATUS OF A GIVEN WORK

❖ IT IS OFTEN NOT **CLEAR** WHETHER THE DOCUMENT IS COPYRIGHT FREE OR NOT



IN SUCH CASES DIGITAL CONTENT IS AVAILABLE FROM THE LOCAL NETWORK ONLY

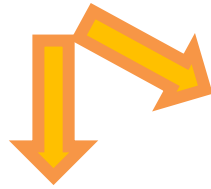




# DIGITAL CONTENT RE-USE: RIGHTS CLEARANCE

Some use cases

IT TAKES A LOT OF **TIME** TO DETERMINE THE COPYRIGHT STATUS OF A GIVEN WORK



Lack of information about death dates of less popular authors or people with very common names

Authored newspapers articles: Public domain 70 years after author's death



Check death dates for all authors of signed articles?  
Make available each article according to copyright?



# DIGITAL CONTENT RE-USE: RIGHTS CLEARANCE

Some use cases

IT TAKES A LOT OF **TIME** TO DETERMINE THE COPYRIGHT STATUS OF A GIVEN WORK



Orphan works: No identified or located rightholder



Make a diligent search, consulting “all relevant sources”?

Register all the diligent searches?

Check the rightsholders of illustrations embedded in books, newspapers, magazines, etc?



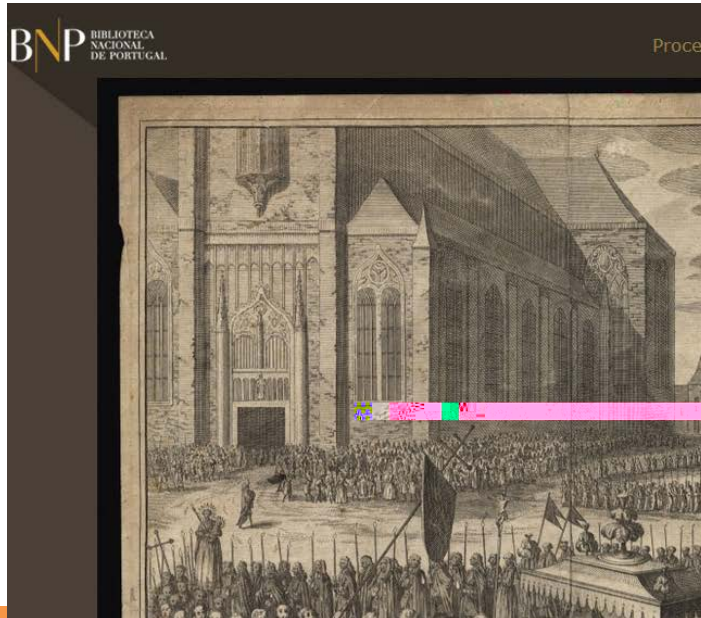
# DIGITAL CONTENT RE-USE: RIGHTS CLEARANCE

Some use cases

IT IS OFTEN NOT **CLEAR** WHETHER THE DOCUMENT IS COPYRIGHT FREE OR NOT

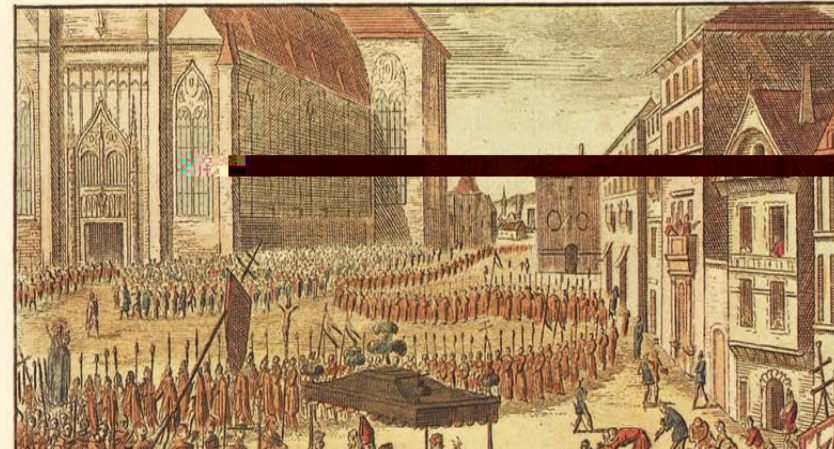


## CRITICAL EDITIONS OF PUBLIC DOMAIN WORKS



BNP BIBLIOTECA NACIONAL DE PORTUGAL

Procession des disciplinans (cota E-1759-V)







# PUBLIC SECTOR INFORMATION DIRECTIVE

## ❖ DIGITAL CONTENT RE-USE

## ❖ DATA RE-USE

- ✓ NLP bibliographic metadata
- ✓ NLP Open Data experiences

## ❖ DIGITISATION FUNDING



# DATA RE-USE: NLP BIBLIOGRAPHIC METADATA

PSI DIRECTIVE STATES THAT...

*“PUBLIC SECTOR SHALL MAKE THEIR DOCUMENTS  
AVAILABLE  
TOGETHER WITH THEIR METADATA.  
METADATA SHALL COMPLY WITH OPEN STANDARDS”*

GRAND ARCHIPEL D'ASIE - ILE TIMOR [ VISUAL GRÁFICO ] : PORTRAITS / A. PELLION PINXT ; CHOUBARD SCULPT

AUTOR(ES): Choubard, ca 1810-1830; Pellion, A., ca 1800, ant. bibliog.  
PUBLICAÇÃO: [S.l. : s.n., ca 1820]  
DESCR. FÍSICA: 1 gravura : pontado, aguarelado  
NOTAS: Data baseada no período de actividade do gravador  
Dim. da composição sem letra: 18,7x26,2 cm  
CDU: 39(=1.594.75)(084.1)  
762(=1.44)\*18 (084.1)

Ver formato UNIMARC | Exportar registo em ISO2709 | Exportar registo em MarcXchange

Exemplares

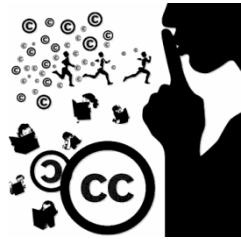
Tipo Acesso	Estado
Sem Acesso	Digitalizado

Pesquisar no Catálogo  
por Autor:  
• Choubard, ca 1810-1830

END. WWW: <http://purl.pt/12329>

Cota	Colecção
E. 3704 P.	Iconografia

[Ver formato UNIMARC](#) | [Exportar registo em ISO2709](#) | [Exportar registo em MarcXchange](#)



# DATA RE-USE: NLP BIBLIOGRAPHIC METADATA

PSI DIRECTIVE STATES THAT...

*“PUBLIC SECTOR SHALL MAKE THEIR DOCUMENTS  
AVAILABLE  
TOGETHER WITH THEIR METADATA.  
METADATA SHALL COMPLY WITH OPEN STANDARDS”*

## REPOX - NLP XML BIBLIOGRAPHIC METADATA REPOSITORY

<http://repop.bn.pt/>

- ✓ Imports data from NLP catalogue
- ✓ Transforms/maps data between schemas (Marcxchange; DC; MODS)
- ✓ Exposes results via OAI-PMH
- ✓ Exposes SEARCH results via HTTP (URN Services)



### Repositório XML da BNP

Consultar histórico de registo bibliográfico

NCB

Consultar

Consultar histórico de registo de autoridade

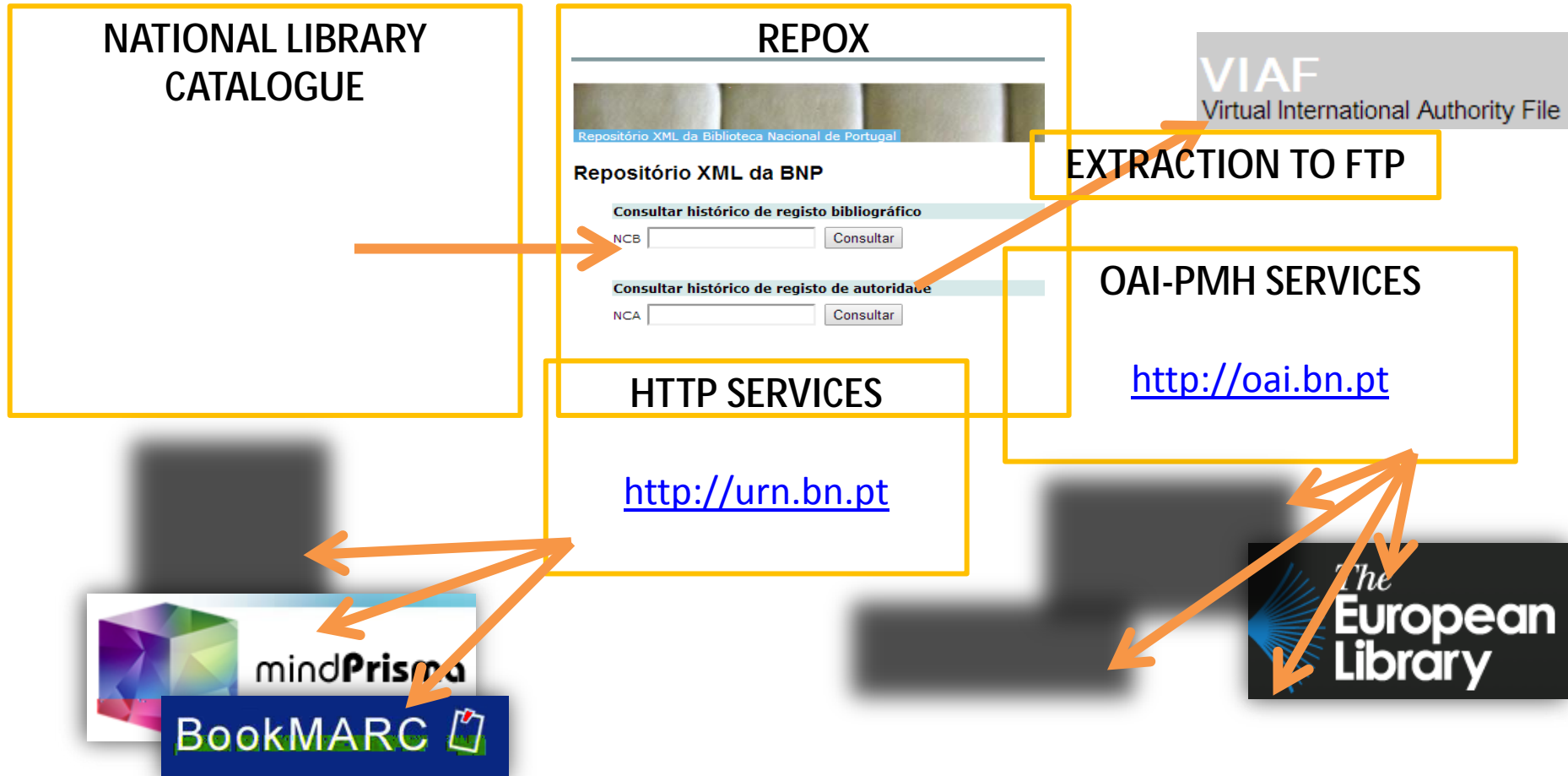
NCA

Consultar



# DATA RE-USE: NLP BIBLIOGRAPHIC METADATA

## REPOX - NLP XML BIBLIOGRAPHIC METADATA REPOSITORY







# DATA RE-USE: NLP BIBLIOGRAPHIC METADATA

## REPOX - NLP XML BIBLIOGRAPHIC METADATA REPOSITORY



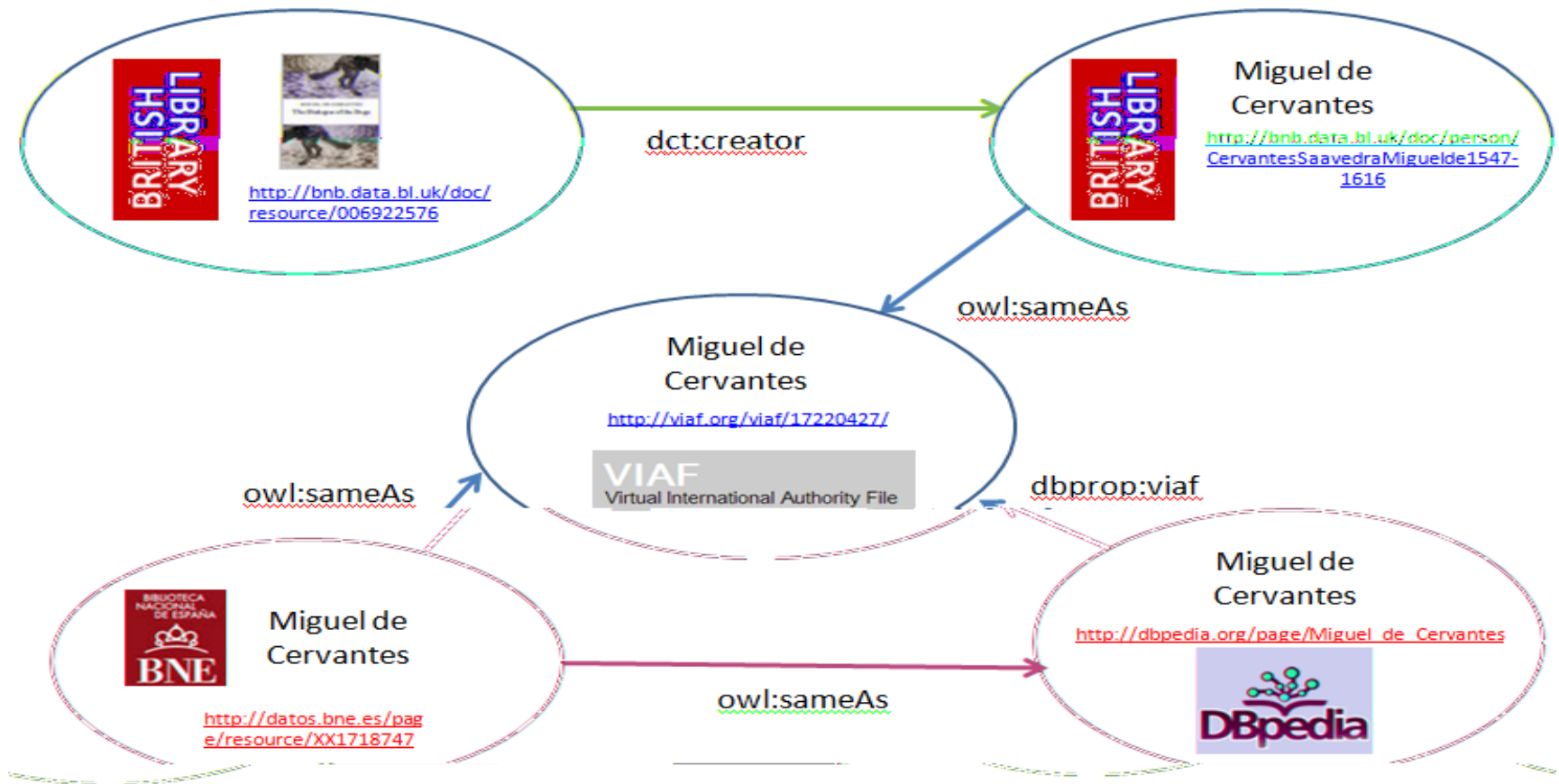


# DATA RE-USE: NLP OPEN DATA EXPERIENCES

## LINKED OPEN DATA

“Linked Open Data is a way of publishing structured data that allows metadata to be connected and enriched, so that different representations of the same content can be found, and links made between related resources”

In Europeana Professional LOD Page

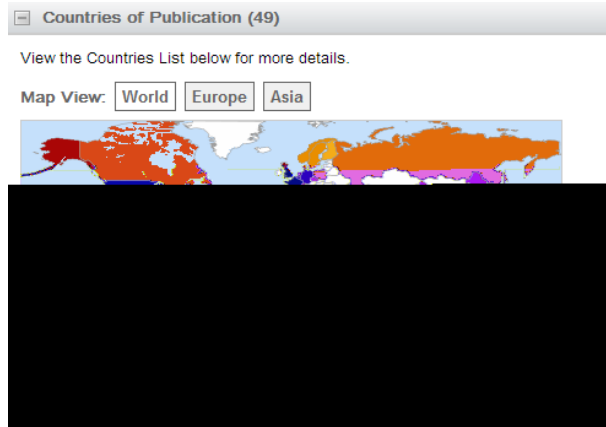




# DATA RE-USE: NLP OPEN DATA EXPERIENCES

## Libraries with LOD links to VIAF

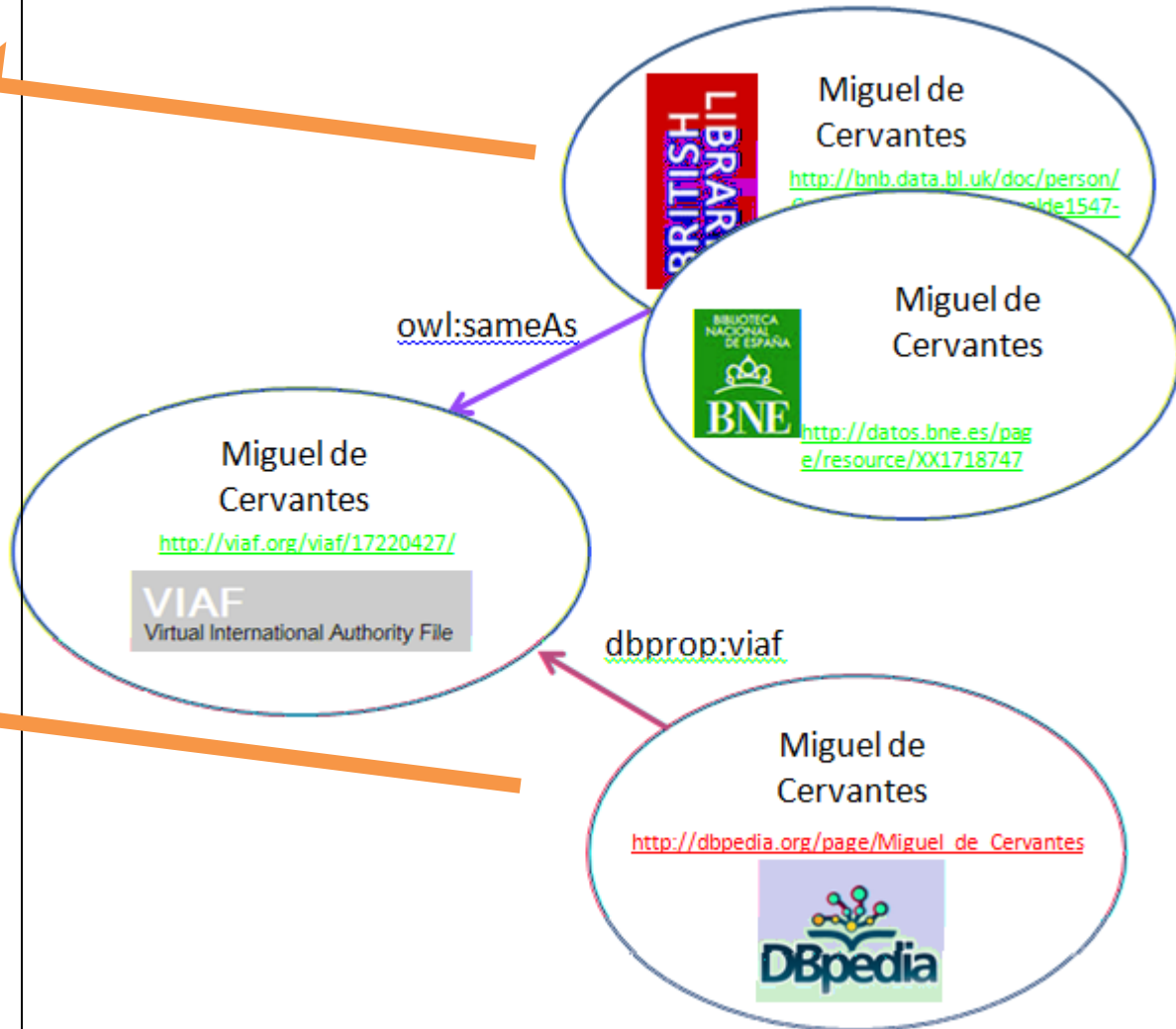
All works written by Cervantes

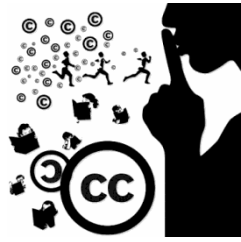


Works written by authors influenced by Cervantes

Works of authors born in Alcalá de Henares

Works of authors that lived in Europe in the XVI Century



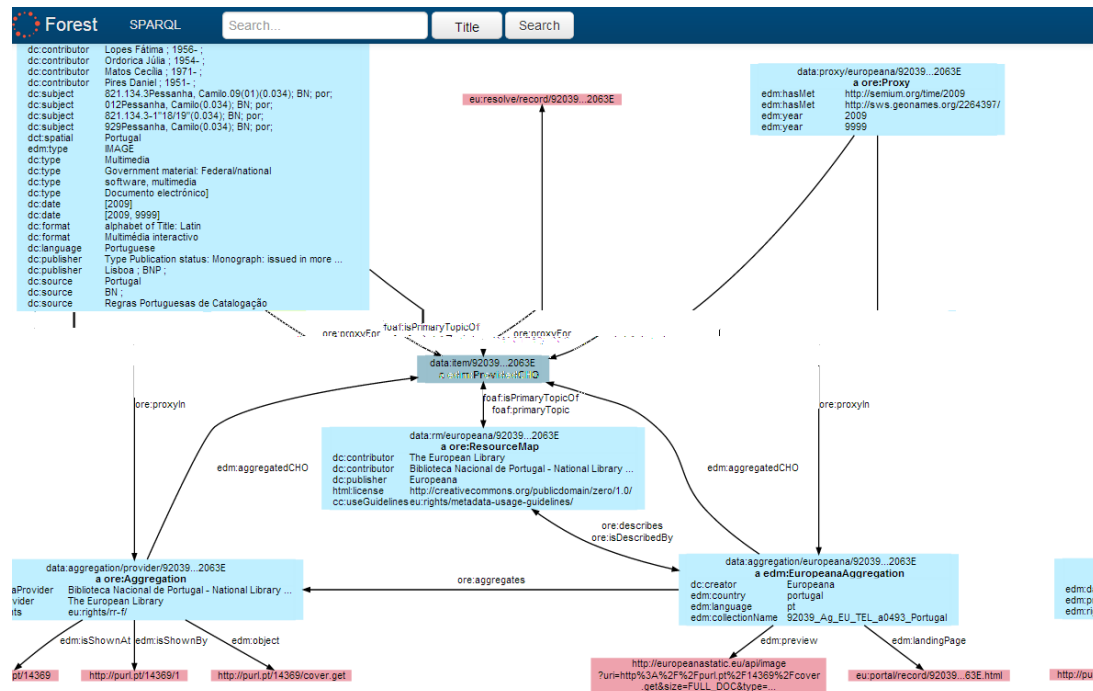


# DATA RE-USE: NLP OPEN DATA EXPERIENCES

## EUROPEANA LINKED OPEN DATA

“The metadata for all the objects in the Europeana portal is **open**, in that it is all licensed under the [CC0 Public Domain Dedication](#) under the terms of the [the Data Exchange Agreement\(DEA\)](#), and can be freely downloaded via the API.”

“In October 2012, a large subset of this data was transformed into **linked data** and made available from [data.europeana.eu](#). The data is represented in the Europeana Data Model (EDM).”





# DATA RE-USE: NLP OPEN DATA EXPERIENCES

## VIAF LINKED OPEN DATA

**datahub**  
The easy way to get, use and share data

Datasets Organizations About Blog Help Search

Home / Organizations / Linking Open Data Cloud / VIAF: The Virtual ...

<http://viaf.org/viaf/data/>

**International Authority File**

Followers: 1

**Organization**

**VIAF: The Virtual International Authority File**

VIAF (Virtual International Authority File) is an OCLC dataset and service -- but national libraries and other partners -- that virtually combines multiple LAM (Library of Congress) name authority files into a single name authority service. Put simply it is a large organizations that occur in library catalogs.

VIAF is a joint project of 20 national libraries, implemented and hosted by OCLC to lower the cost and increase the utility of library authority files by matching and national libraries, and then making that information available on the Web.

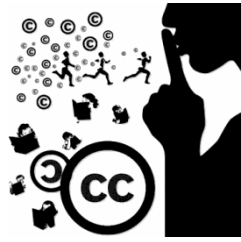
**Openness**

The data is released under Open Data Commons Attribution license. Attribution is required as follows:

Adherence to ODC Attribution Instructions for the correct assertion of attribution preferred form of attribution for VIAF is:

"This [title of report or article or dataset] contains information from VIAF (Virtual International Authority File) which is made available under the ODC Attribution License."

Special cases: In circumstances where providing the full attribution statement is not feasible, the use of canonical VIAF URIs is adequate to satisfy Section 4.3.1 of the License.



# PUBLIC SECTOR INFORMATION DIRECTIVE

❖ DIGITAL CONTENT RE-USE

❖ DATA RE-USE

❖ DIGITISATION FUNDING



# DATA REUSE: DIGITISATION FUNDING

## DIGITISATION COSTS COMPARISON

(The Cost of Digitising Europe's Cultural Heritage – 2010)

Source: European Commission, DG Culture, 2010. Digitisation costs comparison (per page) for different digitisation methods and funding sources.

Digitisation Method	Low cost (per page)	High cost (per page)	Medium cost (per page)	Funding Source
Project-based Digitisation	7.32	16.20	11.76	In-house
Centralised, project-based Digitisation	4.32	9.83	7.08	Outsource
Private Partnership Digitisation	3.25	6.32	4.79	Public-Private

COSTS



# DATA RE-USE: DIGITISATION FUNDING

DIRECTIVE 2003/98/CE and 2013/37/EU (PSI – PUBLIC SECTOR INFORMATION) – art. 11°

- ❖ Exclusive rights related to digitisation of cultural resources shall not exceed 10 years;
- ❖ In case where that period exceeds 10 years, its duration shall be subject to review during the 11th year and every 7 years thereafter;
- ❖ Arrangements existing on July 2013, shall be terminated at the end of the contract or not later than 2043.

GOOGLE LIBRARIES ARRANGEMENTS

Library of Congress 3rd party digitisation agreements:  
Paid access to enriched services (enhanced  
information search, retrieval or presentation)

BNF/PRO-QUEST  
70.000 books (1470-1700);  
3.700 immediately available at  
BNF Gallica;  
All the others available within 10  
years



# Thank you!

Helena Simões Patrício  
[hpatricio@bnportugal.pt](mailto:hpatricio@bnportugal.pt)

