Digital Manuscripts Without Borders

A Discovery Platform of Manuscripts and Rare Books

2019 IIIF Conference

University of Göttingen (June 26th, 2019)

Biblissima

@biblissima

Régis ROBINEAU

@regisrob

Search for interoperable digitized manuscripts and rare books...

This prototype application allows you to search across IIIF-compliant manuscripts and rare books dated before 1800 coming from many digital libraries in the world. It is a work in progress, the platform is updated and enriched on a regular basis. Read more...

Browse collections

Bibliothèque nationale de France (Gallica)

44154 manifests

Digitized manuscripts from Gallica, the digital library of the National Library of France (BnF)

Browse collection

Digital.Bodleian

5838 manifests

Bodleian Libraries' digital collections (Oxford)

Browse collection

BVMM

4213 manifests

Digital library of medieval manuscripts (IRHT-CNRS)

Browse collection

Biblissima IIIF-Collections platform

➜ cross-collections search and discovery prototype for manuscripts and rare books (prior to 1800 only)

➜ aggregates metadata from **10** IIIF digital libraries

➜ complementary to the main Biblissima portal at biblissima.fr

Current scope: 10 datasets (June 2019)

- Gallica (BnF)

- Digital.Bodleian (Oxford)

- BVMM

- e-codices

- British Library
  (BL-BnF Polonsky project)

- Europeana Regia

- Parker Library On the Web

- Bibliothèque Mazarine

- Ghent UL

- Cambridge UL

- Durham University and
  Cathedral Library

~65,000 IIIF Manifests
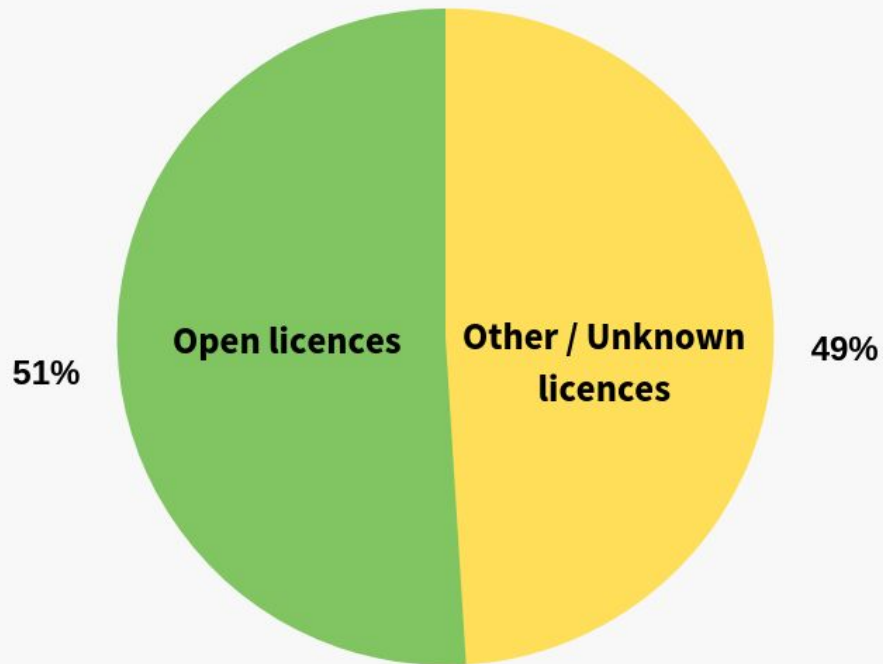
Targeted IIIF datasets

➜ **57** manuscripts repositories identified in the IIIF world (mostly from UK, France, USA, Germany)

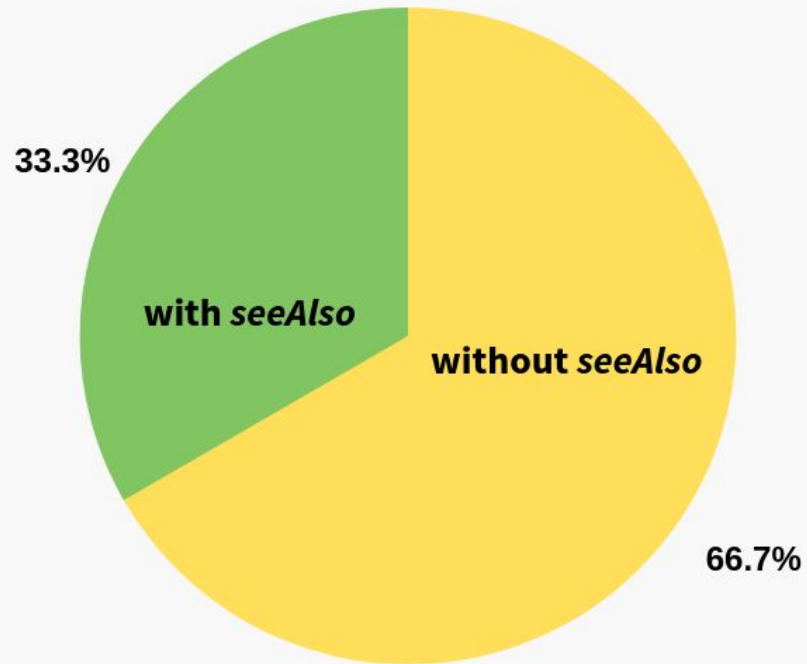frama.link/iiif-mss-registry

➜ A few observations on:

- ◆ open licences
- ◆ use of the *seeAlso* property
- ◆ metadata schemas linked from the *seeAlso*
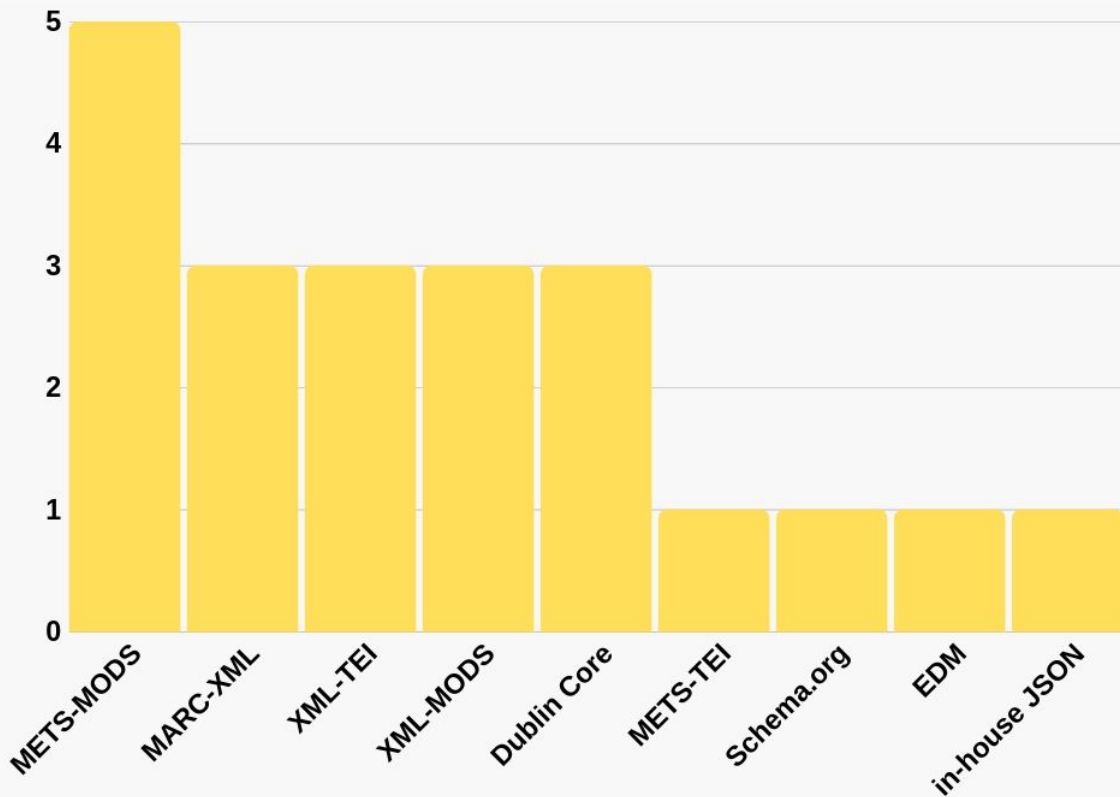- ◆ presence of a crawlable IIIF endpoint

Licences

Use of the seeAlso property

Metadata schemas (from seeAlso property)

Bar chart (y-axis 0 to 5):
- METS-MODS: 5
- MARC-XML: 3
- XML-TEI: 3
- XML-MODS: 3
- Dublin Core: 3
- METS-TEI: 1
- Schema.org: 1
- EDM: 1
- in-house JSON: 1

+ different flavours of RDF (mix of vocabularies / various serialization formats)

Crawlable IIIF endpoint?

33.3%

with crawlable IIIF Collection

without crawlable IIIF Collection

66.7%
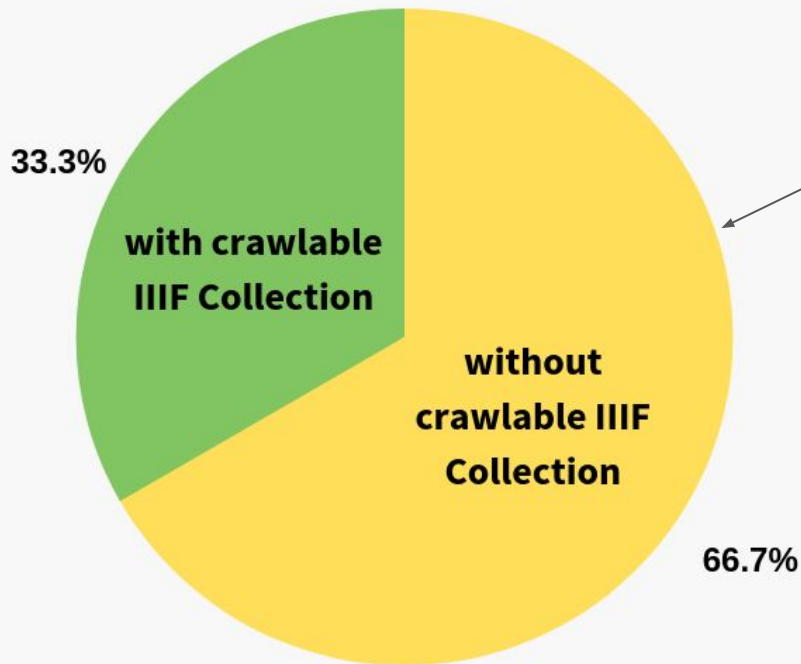
some of them have an OAI-PMH endpoint, as an alternative...

Workflow

1. Crawl and harvest the repositories' endpoints

2. Extract and process the data:
   - ◆ filter by date, normalise the dates formats
   - ◆ normalise strings of languages and roles
   - ◆ **reconcile and cluster named entities** (agents, places…)

3. Transform and ingest the data into ElasticSearch

Demo

iiif.biblissima.fr/collections

Search tips

- **Exact Phrases** " " : e.g.: "Rudimenta grammatices" finds *Rudimenta grammatices*
- **Wildcards** *? : e.g.: Arnold* Ro?e finds *Arnoldus*, *Rome*
- **Fuzzyness** ~ : e.g.: Arnold~ finds *Arnold*, *Arnoud*, *Arnoldus*, *Arnoldi*
- **Must / Must not** + / − : e.g.: +Arnoldus +"Rudimenta grammatices" -Calfurnius finds documents with both *Arnoldus* and *Rudimenta grammatices*, but without *Calfurnius*
- **Boosting** ^n : e.g.: "Rudimenta grammatices" Arnoldus^3 gives a better score for documents with the word *Arnoldus*

Advanced search parameters :
- exact search
- wildcards
- fuzzyness
- inclusion/exclusion
- boosting

Cambridge. Corpus Christi College, Parker Library MS 6

Collection	Parker Library On the Web
Library	Cambridge. Corpus Christi College, Parker Library
Date	1400 - 1499
Language	lat

View entry

Cambridge. Corpus Christi College, Parker Library MS 21

Collection	Parker Library On the Web
Library	Cambridge. Corpus Christi College, Parker Library
Date	1375 - 1399
Language	lat

View entry

Stanford LIBRARIES

Cambridge. Corpus Christi College, Parker Library MS 335

cassiodorus

About

Library

Filter Library …

Paris. Bibliothèque nationale de France, Département des manuscrits	40
St. Gallen. Stiftsbibliothek	18
London. British Library	8
Cambridge. Corpus Christi College, Parker Library	6
Munich. Bayerische Staatsbibliothek	6
Bern. Burgerbibliothek	3
Durham. Cathedral Library	3

Filter by library

Language

Filter Language …

Latin	89
German	4
Greek	2
English	1
Middle French	1

Date

449	1700

Durham. Cathedral Library, MS. B.II.30

Collection	Durham University and Cathedral Library
Library	Durham. Cathedral Library
Date	second quarter, 8th century, 14th century
Language	Latin

View entry

- *Cassiodorus*
- *Cassiodorus*, Commentary on the Psalms, in an abbreviated form.

St. Gallen. Stiftsbibliothek, Cod. Sang. 200

Collection	Europeana Regia
Library	St. Gallen. Stiftsbibliothek
Date	3e quart du 9e siècle
Language	latin

View entry

Library

Filter Library ...

Paris. Bibliothèque nationale de France, Département des manuscrits	40
St. Gallen. Stiftsbibliothek	18
London. British Library	8
Cambridge. Corpus Christi College, Parker Library	6
Munich. Bayerische Staatsbibliothek	6
Bern. Burgerbibliothek	3
Durham. Cathedral Library	3

Language

Filter Language ...

Latin	89
German	4
Greek	2
English	1
Middle French	1

Filter by language

Date

449   1700

Durham. Cathedral Library, MS. B.II.30

Collection	Durham University and Cathedral Library
Library	Durham. Cathedral Library
Date	second quarter, 8th century, 14th century
Language	Latin

View entry

- *Cassiodorus*
- *Cassiodorus*, Commentary on the Psalms, in an abbreviated form.

St. Gallen. Stiftsbibliothek, Cod. Sang. 200

Collection	Europeana Regia
Library	St. Gallen. Stiftsbibliothek
Date	3e quart du 9e siècle
Language	latin

View entry

cassiodorus

About

- *Cassiodorus*
- *Cassiodorus*, Commentary on the Psalms, in an abbreviated form.

Language

Filter Language ...

Latin	89
German	4
Greek	2
English	1
Middle French	1

Date

Filter by date

449 | 1700

☐ Include results without date

Submit

St. Gallen. Stiftsbibliothek, Cod. Sang. 200

Collection	Europeana Regia
Library	St. Gallen. Stiftsbibliothek
Date	3e quart du 9e siècle
Language	latin

View entry

- CASSIODORUS

Munich. Bayerische Staatsbibliothek, Clm 6376

Collection	Europeana Regia
Library	Munich. Bayerische Staatsbibliothek
Date	1ere moitié du 9e siècle
Language	latin

View entry

Bayerische
StaatsBibliothek
Information in erster Linie

Ghent. University Library, Hs. 76

Go to viewer ❯

SOURCE	Ghent University Library
LIBRARY	Ghent. University Library
SHELFMARK	BHSL.HS.0076
DATE	12e eeuw.
TITLE	Etymologiae. De anima. Cassiodorus, Isidore of Seville

AGENT		
	Preferred form	Isidore de Séville (saint, 0560?-0636)
	Original form	Isidorus Hispalensis, 560?-636
	Other form	Isidorus Hispalensis S. Isidorus Hispalensis ISIDORUS HISPALENSIS see more
	Biblissima portal	https://biblissima.fr/en/ark:/43093/pdatab39ccdae3eebcd2100fd6f8a04b2e94036efdcbb

Further info on this author in the Biblissima portal

	Preferred form	Cassiodore (0490?-0580?)
	Original form	Cassiodorus, Flavius Magnus Aurelius, senator, 485?-580?

Role	Former owner
Original form	Saint-Martial de Limoges
Other form	Abbaye Saint-Martial (Limoges ; 0848-1062)
	Abbaye Saint-Martial de Limoges
	Abbaye de Saint-Martial de Limoges
	see more
Biblissima portal	https://biblissima.fr/en/ark:/43093/cdata5b9a6479df162f04d4d2d8823893c9b3a3c3e205

PLACE	**Original form**	Abbaye de Saint-Denis

RIGHTS	Bibliothèque nationale de France, Département des manuscrits
LICENSE	https://gallica.bnf.fr/html/und/conditions-dutilisation-des-contenus-de-gallica
DIGITISATION	http://gallica.bnf.fr/ark:/12148/btv1b84262910
MANIFEST URL	iiif https://gallica.bnf.fr/iiif/ark:/12148/btv1b84262910/manifest.json
BIBLISSIMA PORTAL	https://biblissima.fr/en/ark:/43093/mdata4ff39440fa0bb9c432ce3a1d7ac06300aee66970

Further info on this manuscript in the Biblissima portal

Search across ~60,000 digitized manuscripts and rare books 🔍

Ghent. University Library, Hs. 439

Go to viewer ›

SOURCE	Ghent University Library
LIBRARY	Ghent. University Library
SHELFMARK	BHSL.HS.0439
DATE	eind 13e eeuw.
TITLE	Chronica mundi et chronicon S. Bavonis Gandavensis. Johannes de Thielrode

URI of the entity in the Biblissima authority file:
https://data.biblissima.fr/entity/Q32764

AGENT	**Preferred form**	Johannes de Thielrode (12..-12..?)
	Original form	De Thielrode, Jean, active 1295
	Biblissima authority file	https://data.biblissima.fr/entity/Q32764

Biblissima Authority data

➔ hub to manage and share our authority data:

◆ wiki-based technology (*Wikibase*): natively collaborative + versioning

◆ handles URIs identifiers

◆ natively produces RDF

◆ user-friendly forms to edit entries

◆ remote access for machines: Web API + SPARQL endpoint

data.biblissima.fr

Biblissima Authority data

Publication spread over 2019:

✓   Persons *(done, March 2019) ~26,500 entities*

✓   Geographical names *(done, April 2019) ~5,500 entities*

🚧   Organisations *(in progress)*

🚧   Shelfmarks of manuscripts and early printed books

➔   Textual works

➔   Iconographic descriptors

Feedback to data providers

➔ on the IIIF Manifests, on the metadata encodings…

➔ on the metadata itself (new identifications, corrections…)

➔ provision of data alignments/enrichments (URIs…)

	Feedback_Ghent_University_Library	moi	29 mai 2019 moi	–
	Feedback_Cambridge_Digital_Library	moi	20 mai 2019 moi	–
	Feedback_Durham_Library	moi	20 mai 2019 moi	–
	Feedback_Parker_library	moi	15:02 moi	–

Biblissima IIIF-Collections:

iiif.biblissima.fr/collections

Biblissima authority data:

data.biblissima.fr