OXLOD Pilot
Oxford Linked Data

4 October 2018 - OeRC

Background

- What did we set out to achieve and why is this important?

- What have we delivered?

- Purpose of today's session
  - Pilot findings (Gabriel Hanganu)
  - Introduction to Linked Data (Kevin Page)
  - Integration principles (Athanasios Velios)
  - OXLOD workflow (Athanasios Velios, Andrew Morrison, Sarah Joomun and Chris Powell)
  - OXLOD demo (Athanasios Velios)

OXLOD Pilot findings

OXLOD pilot

- Limited in scope, time, resources
- Aim: explore if Linked Data can bring significant benefits to GLAM

Delivered

- Workflow
- Server and software
- Processed datasets
- User interface
- Use cases
- Evaluation framework
- Metrics
- Community

OXLOD in numbers

- **13,280,376** records (triples)
- **40,869** words of documentation text
- **325** minutes of workshop videos
- **63** workshop participants
- **26** datasets
- **10** workshops

GLAM benefits

- Better data
  - Structured data
  - Enriched data

- Better research
  - Cross-search
  - Discovery

- More effective collection data management

- Stronger GLAM community
  - Open workshops, knowledge base
  - Knowledge transfer, capacity building

Representing knowledge on the Web using Linked Data
a 10 minute introduction

(The Semantic Web and Why You Should Care)

Kevin Page
University of Oxford e-Research Centre
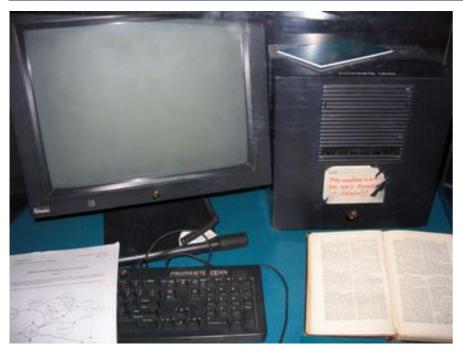
OXLOD final workshop, October 2018

Why we should care

When knowledge has been generated, we should capitalise on its value by

- capturing it
- publishing it
- using it
- linking it
- re-using and building upon it ("unintentionally"?)
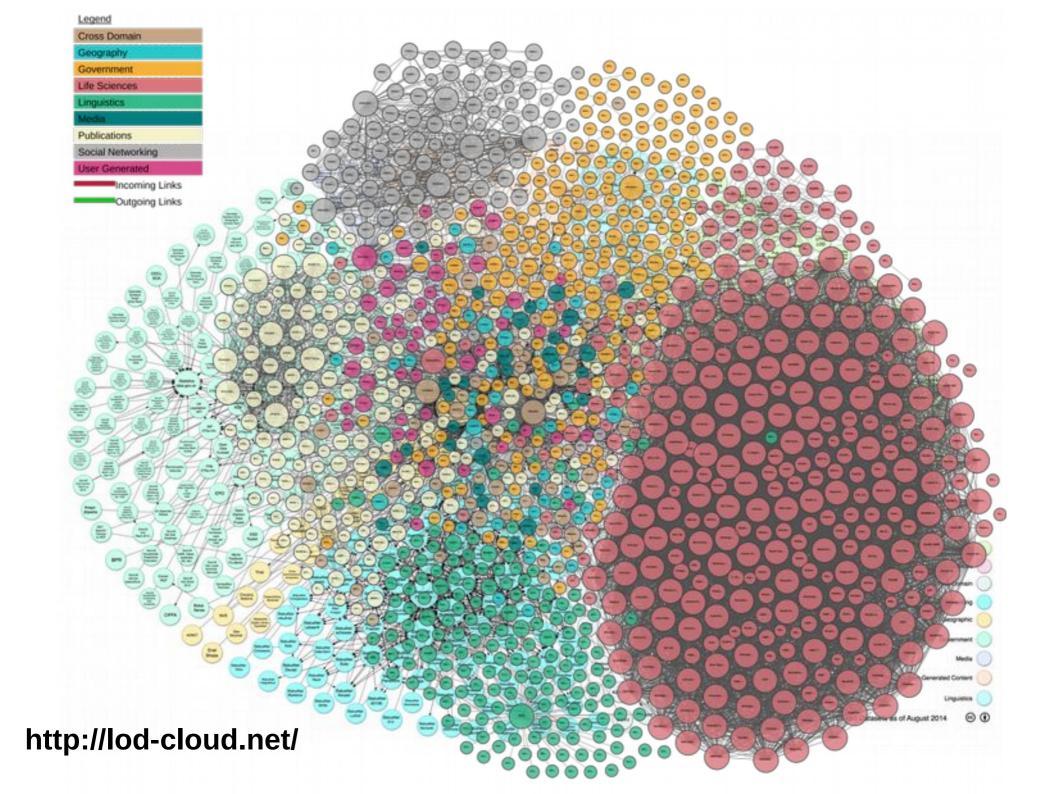
What is the *Semantic* Web?

"This machine is a server – DO NOT POWER IT DOWN!!"

- The web is the largest and most successful distributed system ever constructed

- But it is a Web of Documents

The Semantic Web is the effort to create the equivalent Web of Data

- Semantic Web activities have been baking since the late 1990s

- The infamous layer cake has evolved

- Linked Data is a more recent movement...

There are two words in

Semantic Web

Both are important!

Legend
- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated
- Incoming Links
- Outgoing Links

As of August 2014

http://lod-cloud.net/

SW vs. LD vs. LOD vs. OD

- The World Wide Web Consortium (W3C) **Semantic Web** activities and *standards* define underlying technologies

- **Data** can be **Open**, without being **Linked** – Open Data

- **Data** can be **Linked**, without being **Open** – Linked Data

- **Data** can be both **Open** and **Linked – ** Linked Open Data

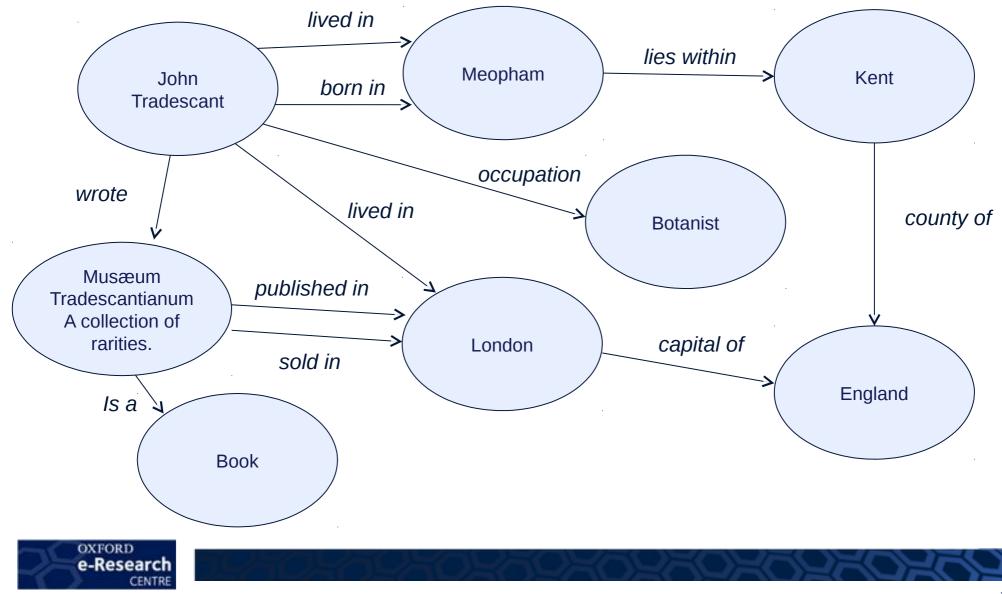- Underlying technologies can support any combination – *which* combination is a matter of policy

The fundamental building block of Linked Data is RDF

Subject —Predicate→ Object

That's it.

An RDF "triple"
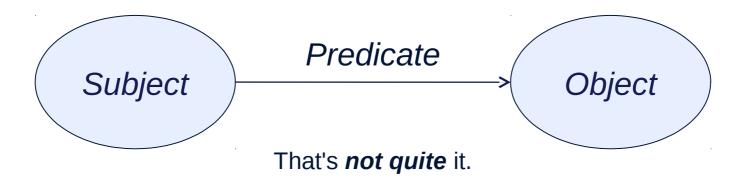
(RDF is the Resource Description Framework)

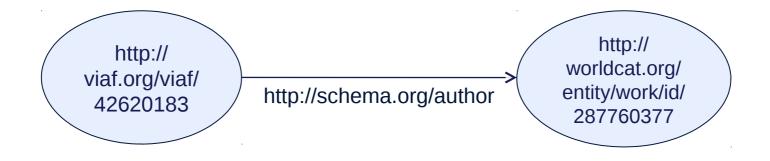# "Triples" of information, linked and expanding to the scale of the Web

Universal Resource Identifiers

Subject → *Predicate* → Object

That's **not quite** it.

Universal Resource Identifiers

Universal Resource Identifiers

Universal Resource Identifiers

That's *really* it.

Linked Data considerations
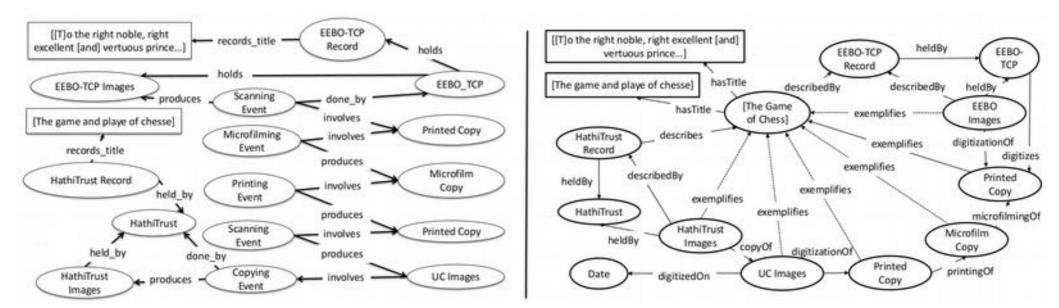
- The previous example was *very* simplistic

    - The *combination* of many specialised fields of knowledge will form a large and complex web

- Linked Data is about making associations between **information with common meaning** (semantics) where that is agreed upon

    - Not (necessarily) about forcing common meanings across information

- Linked Data is a framework for scalable semantic interoperability of digital information. **It does not mandate use of specific semantics,** only standardises the technical medium through which to link

Ontologies and simultaneous *contexts*

We encode and share our specialised meanings through **ontologies**.

RDF, as a standard, is adaptable, and *can accommodate multiple*

simultaneous contexts (perspectives, specialisms, community practices)
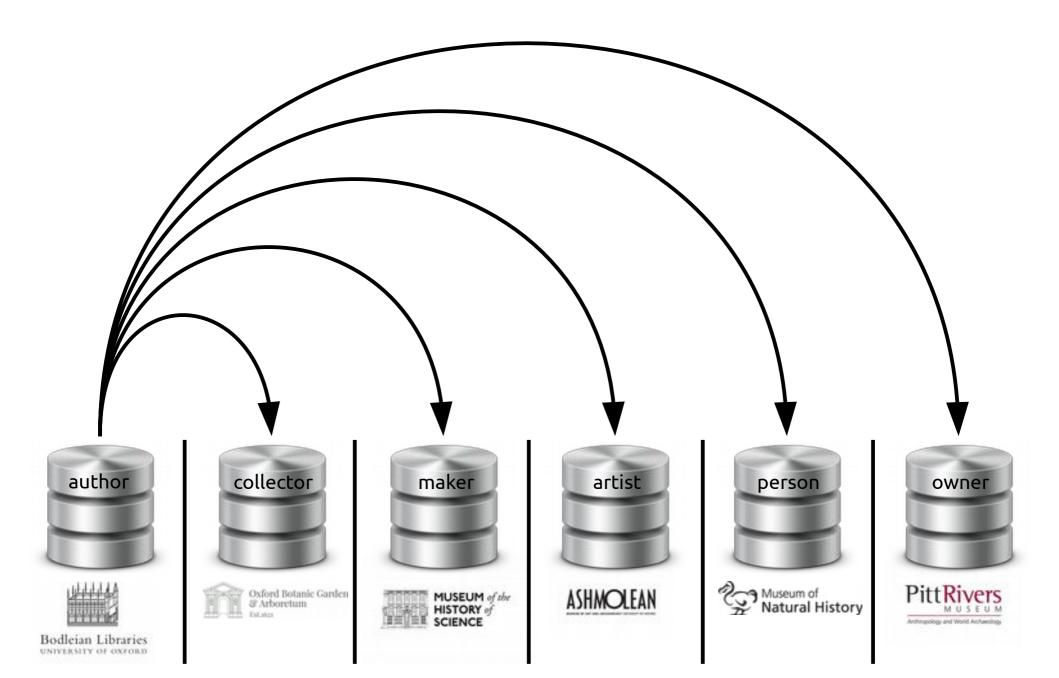

Nurmikko-Fuller, T., Page, K. et al. "Building Complex Research Collections in Digital Libraries: A Survey of Ontology Implications"
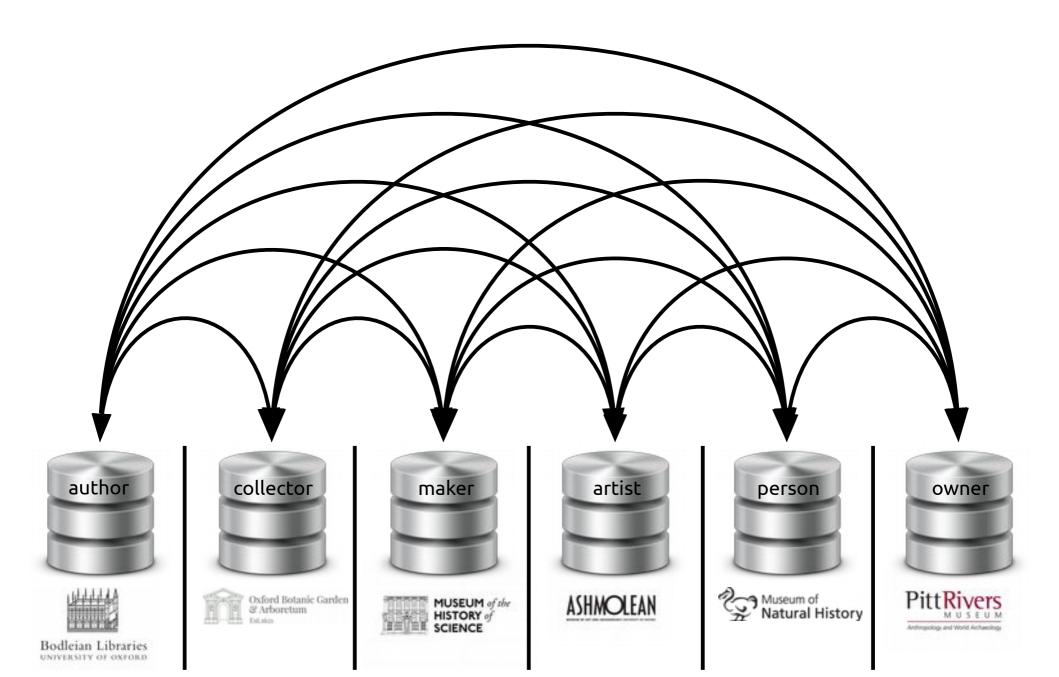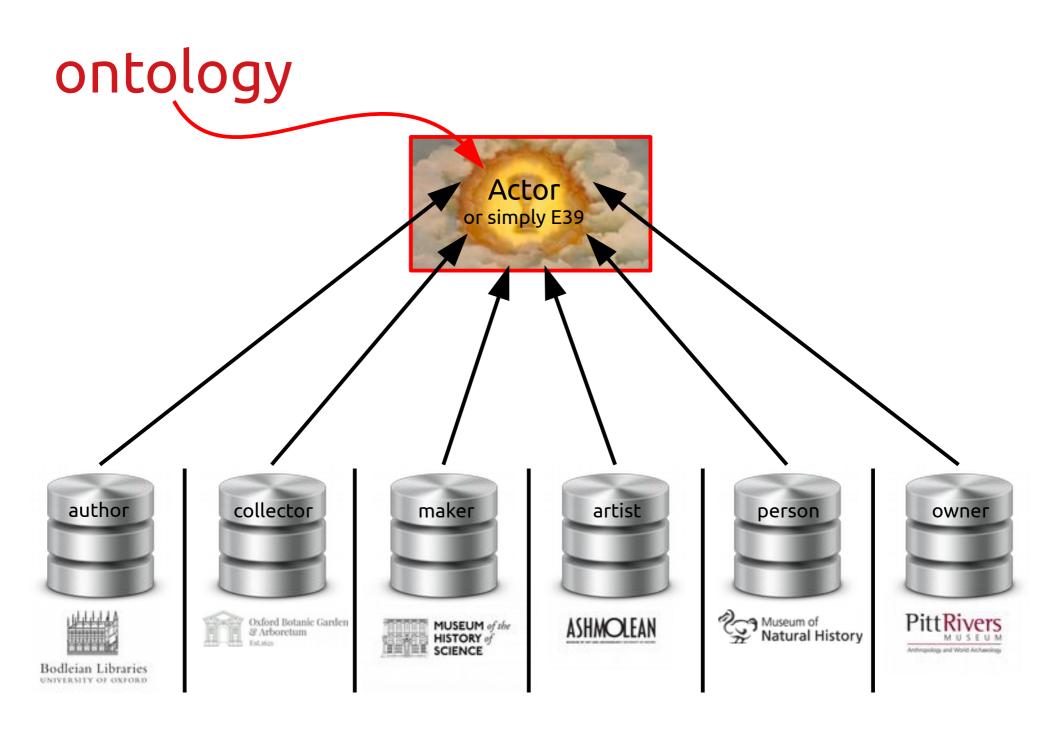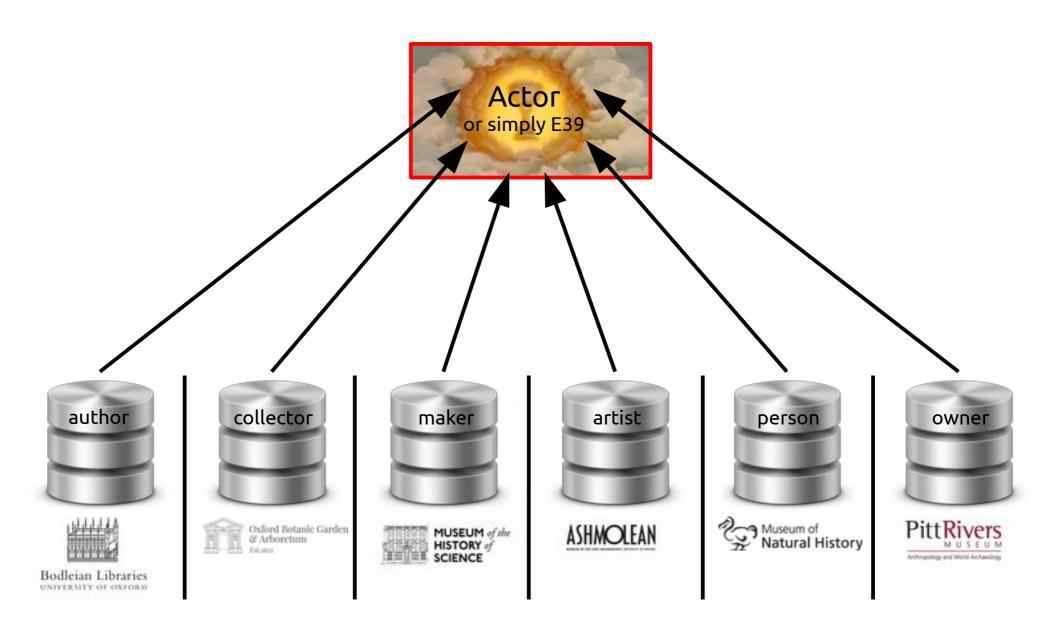*ACM/IEEE Joint Conference on Digital Libraries (2015)*

Integration
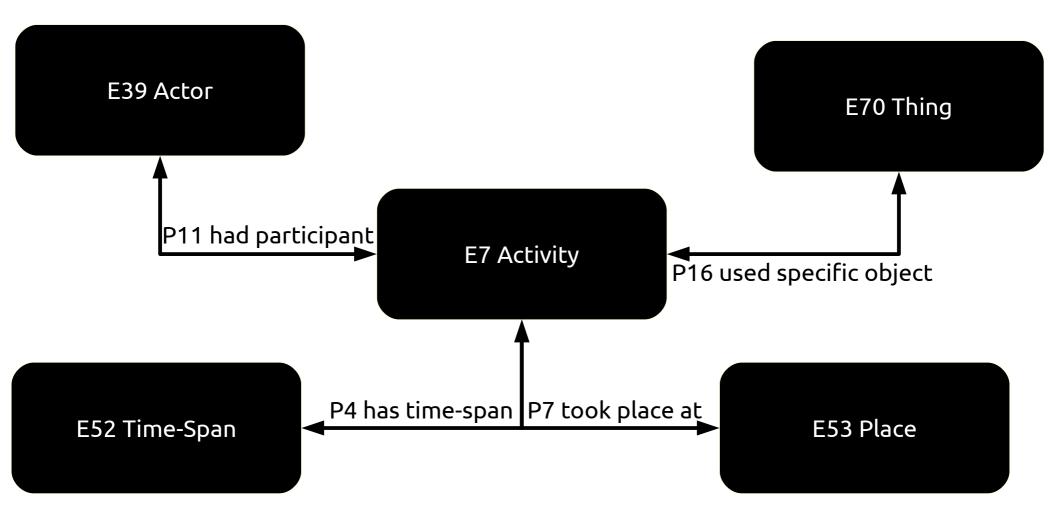
Larger sample →
Better conclusions

Data integration

Actor
or simply E39

author — Bodleian Libraries UNIVERSITY OF OXFORD

collector — Oxford Botanic Garden & Arboretum Est.1621

maker — MUSEUM of the HISTORY of SCIENCE

artist — ASHMOLEAN

person — Museum of Natural History

owner — PittRivers MUSEUM Anthropology and World Archaeology
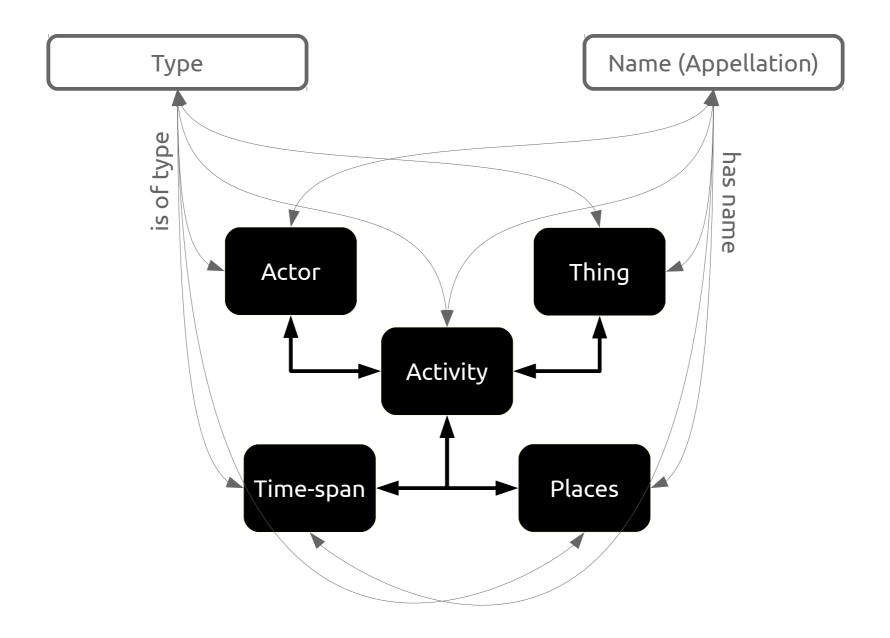
Ontology

- Provides properties to link entities
  - CIDOC-CRM
  - heritage ontology

OXLOD
cross-searching demo

Gardens, Libraries & Museums
University of Oxford

Search

Opening Times
What's On
Family Friendly Events
Shout Out For Women
Out in Oxford
Students & Researchers
Schools & Groups
Volunteering
Art Education
Venue Hire

What's On

Jade in Oxford
An exhibition about the history and properties of Jade with exhibits from the Oxford University Museums

Pre OXLOD system

Pre-OXLOD system

Specialist librarian

Collection data manager

Researcher / curator

GLAM contributors

Pre-OXLOD system

Specialist librarian

Collection data manager

Researcher / curator
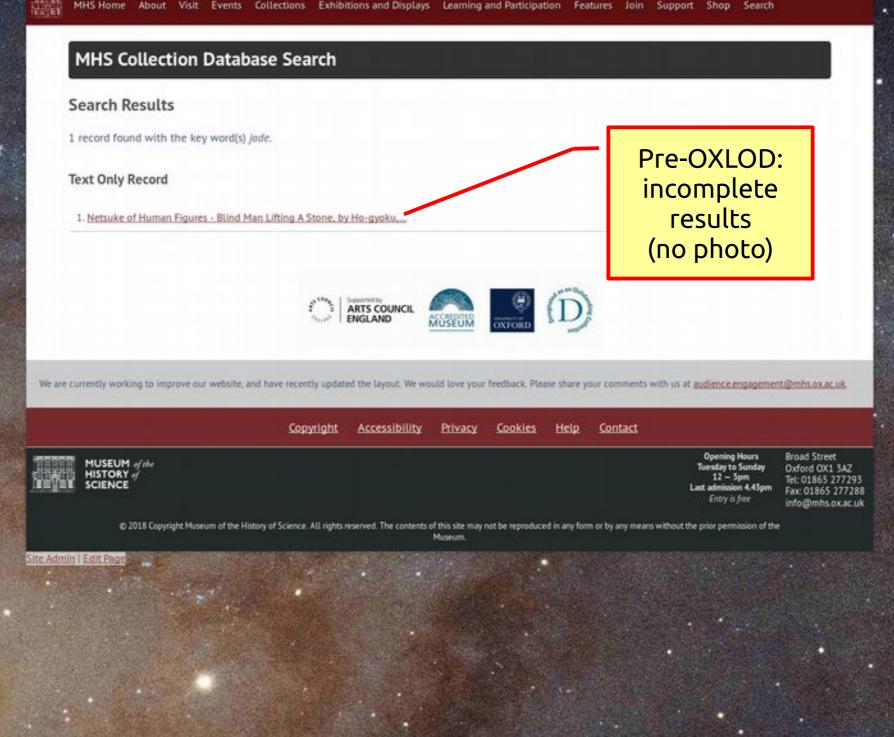
GLAM contributors

MHS Collection Database Search

Basic Search of over 20,000 records

jade

[Search]

☑ Show thumbnails

☐ Contains multimedia (including images, audio files, PDFs etc.)

Or simply click on an image below.

Field Search

☐ "Or" Search

Inventory No
e.g. 49726

Title

Creation Date

Primary Inscription

Provenance

Subject Classification

Collection Group

Brief Description

Material

Accession Number

Object Type

[Search]   [Clear]

Pre-OXLOD:
free-text search
interface

MHS Home   About   Visit   Events   Collections   Exhibitions and Displays   Learning and Participation   Features   Join   Support   Shop   Search

MHS Collection Database Search

Basic Search

jade

Search

☑ Show thu

☐ Contains

Or simply cli

Basic Search of over 20,000 records

jade

Search

☑ Show thumbnails

☐ Contains multimedia (including im

Field Search

☐ "Or" Search

Inventory No
e.g. 49726

Collection Group

Title

Brief Description

Creation Date

Material

Primary Inscription

Accession Number

Provenance

Object Type

Subject Classification

Search   Clear

Pre-OXLOD:
free-text search
interface

MHS Home   About   Visit   Events   Collections   Exhibitions and Displays   Learning and Participation   Features   Join   Support   Shop   Search

MHS Collection Database Search

Basic Search of over 20,000 records

enter search string

Search

☑ Show thumbnails

☐ Contains multimedia (including images, audio files, PDFs etc.)

Or simply click on an image below.

Field Search

☐ "Or" Search

Inventory No
e.g. 49726

Title

Creation Date

Primary Inscription

Provenance

Subject Classification

Collection Group

Brief Description

Material
jade

Accession Number

Search   Clear

Pre-OXLOD: keyword search

MHS Home   About   Visit   Events   Collections   Exhibitions and Displays   Learning and Participation   Features   Join   Support   Shop   Search

MHS Collection Database Search

Search Results

Search for **all** terms: Material: *jade*

2 records found.

Records with Images

(Followed by: text only record)

1. Hsüan-chi, Formerly Identified as Early Chinese Astronomical...

Text Only Record

(Preceded by: records with image previews)

2. Netsuke of Human Figures - Blind Man Lifting A Stone, by Ho-gyoku....

ARTS COUNCIL ENGLAND   ACCREDITED MUSEUM   UNIVERSITY OF OXFORD

Pre-OXLOD: more results

MHS Collection Database Search

Record Details

Hsüan-chi, Formerly Identified as Early Chinese Astronomical Instrument

Inventory Number:	77193
Object Type:	
Place Created:	China Asia
Accession Number:	1959-33
Brief Description:	
Primary Inscriptions:	
Provenance:	Purchased from the Eastgate Galleries, Oxford, in 1959.
Collection Group:	
Material(s):	Jade

Narratives

- Inv. 77193 - rejection of astronomical significance
- Inventory no. 77193 - Former Display Label

MHS Collection Data

MHS Collection Database Search

Record Details

Record Details

Hsüan-chi, Formerly I

Hsüan-chi, Formerly Identified as Early Chinese Astronomical Instrument

Inventory Number:	77193
Object Type:	
Place Created:	China Asia
Accession Number:	1959-33
Brief Description:	
Primary Inscriptions:	
Provenance:	Purchased f
Collection Group:	
Material(s):	Jade

Inventory Number:	77193
Object Type:	
Place Created:	China Asia
Accession Number:	1959-33
Brief Description:	
Primary Inscriptions:	
Provenance:	Purchased from the Eastgate Galleries, Oxford, in 1959
Collection Group:	
Material(s):	Jade

Narratives

- Inv. 77193 - rejection of astr
- Inventory no. 77193 - Forme

Narratives

- Inv. 77193 - rejection of astronomical significance
- Inventory no. 77193 - Former Display Label

Pre-OXLOD system: step 1

Pre-OXLOD system: step 2

Pre-OXLOD system: step 3

Untitled 1 - LibreOffice Calc

File Edit View Insert Format Sheet Data Tools Window Help

D3

	A	B	C	D	E	F	G	H	I	J	K
1	Inventory Nu	77193									
2	Object Type:										
3	Place Create	China Asia									
4	Accession N	1959-33									
5	Brief Description:										
6	Primary Inscriptions:										
7	Provenance:	Purchased from the Eastgate Galleries, Oxford, in 1959.									
8	Collection Gr										
9	Material(s):	Jade									

Sheet1

Sheet 1 of 1   Default   Average: ; Sum: 0   100%

Pre-OXLOD system: step 5

Pre-OXLOD system: step 6

Pre-OXLOD system: step 7

Pre-OXLOD system: step 8

Pre-OXLOD: repeat all steps

Pre-OXLOD: repeat all steps

Current system: repeat all steps

Untitled 1 - LibreOffice Calc

Untitled 1 - LibreOffice Calc

Untitled 1 - LibreOffice Calc

File Edit View Insert Format Sheet Data Tools Window Help

B3:B11

	A	B	C	D	E	F	G	H	I	J	K
1	Inventory Nu	Object Type:	Place Create	Accession N	Brief Descrip	Primary Insc	Provenance:	Collection Gr	Material(s):		
2	77193		China As	1959-33			Purchased f		Jade		

Sheet1

Sheet 1 of 1   9 rows, 1 columns selected   Default   Average: ; Sum: 0   100%

Pre-OXLOD: repeat all steps

Pre-OXLOD: repeat all steps

Pre-OXLOD: repeat all steps

MUSEUM *of the* HISTORY *of* SCIENCE

+

Pre-OXLOD: added two records from MHS

Pre-OXLOD:
to search across collections each step needs to be repeated

ASHMOLEAN
MUSEUM OF ART AND ARCHAEOLOGY UNIVERSITY OF OXFORD
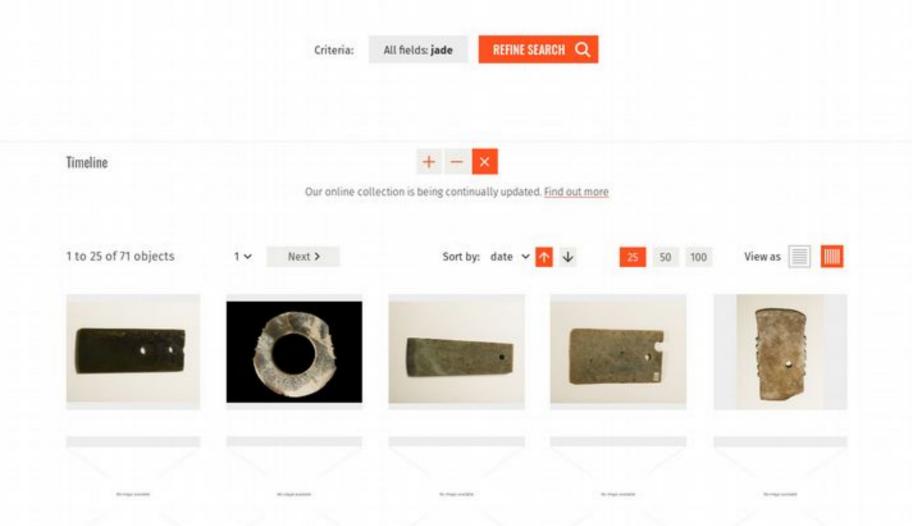
COLLECTION OVERVIEW   BROWSE   SEARCH   ABOUT

SEARCH RESULTS

Criteria: All fields: jade   REFINE SEARCH 🔍

Timeline   ➕ ➖ ✖️   ?

Our online collection is being continually updated. Find out more

1 to 25 of 71 objects   1 ⌄   Next ›   Sort by: date ⌄ ↑ ↓   25 50 100   View as ▤ ▥

ASHMOLEAN
MUSEUM OF ART AND ARCHAEOLOGY UNIVERSITY OF OXFORD

← ⠿ →

Notched disc axe, or xuanji

Criteria: All fields: **jade**   REFINE SEARCH 🔍

© Ashmolean Museum, University of Oxford (image)

Details +

Title	Notched disc axe, or xuanji
Associated place	China (place of creation)
Date	2000 - 1500 BC Neolithic Period (China) (c. 10,000 - c. 2000 BC) Erlitou Period (c. 1900 - c. 1500 BC)
Material and technique	jade
Material index	jade
Technique index	carved
Object type	funerary equipment; xuanji; ritual jade
Dimensions	9.2 x 9 x 0.5 cm max. (height x width x depth)
No. of items	1
Credit line	Presented by Sir Herbert Ingram, 1956.
Museum location	not on display
Museum department	Eastern Art
Accession no.	EA1956.1578

Details +

Title	Notched disc axe, or xuanji
Associated place	China (place of creation)
Date	2000 – 1500 BC Neolithic Period (China) (c. 10,000 – c. 2000 BC) Erlitou Period (c. 1900 – c. 1500 BC)
Material and technique	jade
Material index	jade
Technique index	carved
Object type	funerary equipment; xuanji; ritual jade
Dimensions	9.2 x 9 x 0.5 cm max. (height x width x depth)
No. of items	1
Credit line	Presented by Sir Herbert Ingram, 1956
Museum location	not on display
Museum department	Eastern Art
Accession no.	EA1956.1578

Our online collection is being continually updated. Find out more

MUSEUM *of the* HISTORY *of* SCIENCE

\+

ASHMOLEAN
MUSEUM OF ART AND ARCHAEOLOGY UNIVERSITY OF OXFORD

\+

Pre-OXLOD: added one record from the Ashmolean

Oxford University Museum of Natural History

Collections Online *BETA*

Scope: [all ▾] [jade] [🔍 Search]   Advanced Search | Help

Search Results

You searched for *jade*.
There were 18 results.

Mineralogy
MIN.9072, 'Jade'

Mineralogy
MIN.12281, 'Jade', New Zealand

Mineralogy
MIN.23728, 'Jade' (pale green, part of a container handle?)

Mineralogy
MIN.12283, Tremolite with 'jade', Easton, Northampton Co., Pennsylvania, USA

Mineralogy
MIN.23733, 'Jade' (ornamented blade holder)

Mineralogy
MIN.24672, 'Jade' (cabochon)

Mineralogy
MIN.23636, Nephrite 'jade' (wedge-shaped), New Zealand

Mineralogy
MIN.18592, Nephrite 'jade', New Zealand

Mineralogy
MIN.23635, Nephrite 'jade' (thin slice)

Mineralogy
MIN.19502, Actinolite var. nephrite 'New Zealand jade', New Zealand

MUSEUM *of the* HISTORY *of* SCIENCE + ASHMOLEAN + Museum of Natural History +

Pre-OXLOD: added records from MNH

PittRivers MUSEUM

1948.3.30 B   On Display?   Yes   **1**

Carved stone figure [SM (Verve) 22/6/2016]

M America   Mexico   Oaxaca

1943.10.66 B   On Display?   Yes   **2**

Unquantified assemblage of ?50 stone beads, strung together on a necklace

M America   Mexico   Northern Oaxaca

1897.38.1   On Display?   Yes   **3**

Polished jade axe, ground curved blade. [MJD (Verve) 30/3/2016]

M America   Guatemala

1933.85.104   On Display?   Yes   **4**

Stone bead or pendant

M America   Costa Rica

1936.3.1 .1 1936.3.1 .2   On Display?   Yes   **5**

Curved bone handle with bone socket [.1], with ground stone blade [.2].

N America   USA   Alaska   Point Hope

1991.17.64   On Display?   Yes   **6**

MUSEUM *of the* HISTORY *of* SCIENCE + ASHMOLEAN MUSEUM OF ART AND ARCHAEOLOGY UNIVERSITY OF OXFORD + Museum of Natural History +

+ PittRivers MUSEUM Anthropology and World Archaeology +

Pre-OXLOD: added records from the Pitt Rivers

Pre-OXLOD system

Specialist librarian

Collection data manager

Researcher / curator

GLAM contributors

SOLO

SEARCH OXFORD
LIBRARIES ONLINE

UNIVERSITY OF
OXFORD

NEW SEARCH   BROWSE   MAP SEARCH   FEEDBACK   HELP   TAGS

jade   ✕ | All libraries/collections ▾ 🔍   ADVANCED SEARCH

Sort & Filter Results

PAGE 1 *955 Results*   📌 **Save query**

Sort by Relevance ▾

Show only: ⌃

    Online Resources   (132)

    Physical Items   (861)

Library ⌄

Resource Type ⌄

Creator ⌄

Publication Date ⌃

From   To

-500   ⇅ 2018 ⇅ Refine

Topic ⌄

Language ⌄

Collection ⌄

Map Scale ⌄

---

1   BOOK

Jade

Zara, Louis, 1910-2001

c1969 | London : Studio Vista | 82, [2] p. : ill. ; 22 cm

Find & Request ›

    🔗 ✉ 📌 ⋯

---

2   BOOK

Jade .

Wills, Geoffrey.

1964 | London : Arco Publications | 121 p. illus. 23 cm

Find & Request ›

    🔗 ✉ 📌 ⋯

---

3   BOOK

Jade

Palmer, Jock Pegler.

[1967] | London : Spring Books | 44 p. front., illus., 54 col. plates, map, table. 28 cm

Find & Request ›

    🔗 ✉ 📌 ⋯

---

4   MULTIPLE VERSIONS

Jade

Keverne, Roger

📖 2 versions of this record exist. See all versions ›

---

5   BOOK

Jade

Ward, Fred, 1935-2016. | Ward, Charlotte

1996 | Bethesda, MD : Gem Book Publishers | 64 p. : col. ill. ; 23 cm

Find & Request ›

    🔗 ✉ 📌 ⋯

---

6   BOOK

jade

BOOK
Jade
Ward, Fred, 1935-2016. | Ward, Charlotte
1996 | Bethesda, MD : Gem Book Publishers | 64 p. : col. ill. ; 23 cm

Find & Request >

TOP

EXPORT OPTIONS

FIND & REQUEST

DETAILS & LINKS

LINKS

MORE

TAGS

Export Options

REFWORKS	ENDNOTE	RIS	CITATION	PERMALINK	PRINT	E-MAIL

Find & Request

LOCATION ITEMS

Museum of Natural History Library (i)

549.642.26 / 1	Reference N11632020
Mineralogical Collections	Available

Details & Links

Title	Jade
Further information	text and photographs, Fred Ward ; editing, Charlotte Ward.
Author	Ward, Fred, 1935-2016. >
	Ward, Charlotte >
Publisher Details	Bethesda, MD : Gem Book Publishers
Publication Date	1996
Format	64 p. : col. ill. ; 23 cm
Language	English
Identifier	ISBN: 0963372343
	ISBN: 9780963372345
Subjects	Jade >
	Precious stone industry >
Related Titles	Series: Ward, Fred, 1935-2016. Gem Book series.
Aleph System Number	013470031
Type	book

Links

Sort & F

Sort by   R

Show only

Online R
Physical

Library ˅

Resource

Creator ˅

Publicatio

From
-500

Topic ˅

Language

Collection

Map Scale

TOP

EXPORT OPTIONS

FIND & REQUEST

DETAILS & LINKS

LINKS

MORE

TAGS

jade

Sort & F

Sort by   R

Show only

Online Re

Physical

Library ⌄

Resource

Creator ⌄

Publicatio

From
-500

Topic ⌄

Language

Collection

Map Scale

Export Options

| REFWORKS | ENDNOTE | RIS | CITATION | PERMALINK | PRINT | E-MAIL |

Encoding

UTF-8 ▾

DOWNLOAD

Find & Request

LOCATION ITEMS

Museum of Natural History Library ⓘ

549.642.26 / 1	Reference N11632020
Mineralogical Collections	Available

Details & Links

Title	Jade
Further information	text and photographs, Fred Ward ; editing, Charlotte Ward.
Author	Ward, Fred, 1935-2016. ›
	Ward, Charlotte ›
Publisher Details	Bethesda, MD : Gem Book Publishers
Publication Date	1996
Format	64 p. : col. ill. ; 23 cm
Language	English
Identifier	ISBN: 0963372343
	ISBN: 9780963372345
Subjects	Jade ›
	Precious stone industry ›
Related Titles	Series: Ward, Fred, 1935-2016. Gem Book series.
Aleph System Number	013470031
Type	book

SOLO

SEARCH OXFORD
LIBRARIES ONLINE

jade

TOP

EXPORT OPTIONS

FIND & REQUEST

DETAILS & LINKS

LINKS

MORE

TAGS

Export Options

REFWORKS　　ENDNOTE　　RIS　　CITATION　　PERMALINK　　PRINT　　E-MAIL

Encoding

UTF-8 ▾

DOWNLOAD

Sort & F

Sort by

Show only

Online R
Physical

Library ⌄

Resource

Creator ⌄

Publicatio

From
-500

Topic ⌄

Language

Collection

Map Scale

Find & Request

LOCATION ITEMS

Museum of Natural History Library ⓘ

549.642.26 / 1
Mineralogical Collections

Refere
Availab

Details & Links

Title	Jade
Further information	text and photographs, Fred Ward ; editing, Charlotte Ward.
Author	Ward, Fred, 1935-2016. ›
	Ward, Charlotte ›
Publisher Details	Bethesda, MD : Gem Book Publishers
Publication Date	1996
Format	64 p. : col. ill. ; 23 cm
Language	English
Identifier	ISBN: 0963372343
	ISBN: 9780963372345
Subjects	Jade ›
	Precious stone industry ›
Related Titles	Series: Ward, Fred, 1935-2016. Gem Book series.
Aleph System Number	013470031
Type	book

Opening Primo_RIS_Export.ris ✕

You have chosen to open:

☐ **Primo_RIS_Export.ris**
which is: plain text document (267 bytes)
from: blob:

What should Firefox do with this file?

◉ Open with [Text Editor (default) ▾]

◯ Save File

☐ Do this automatically for files like this from now on.

[Cancel]　[OK]

MUSEUM *of the* HISTORY *of* SCIENCE

+

ASHMOLEAN
MUSEUM OF ART AND ARCHAEOLOGY UNIVERSITY OF OXFORD

+

Museum of Natural History

+

+

PittRivers MUSEUM
Anthropology and World Archaeology

+

Bodleian Libraries
UNIVERSITY OF OXFORD

+

Pre-OXLOD: added references from the Bodleian

MUSEUM *of the* HISTORY *of* SCIENCE + ASHMOLEAN + Museum of Natural History +

+ PittRivers MUSEUM + Bodleian Libraries UNIVERSITY OF OXFORD + Oxford Botanic Garden & Arboretum Est.1621

Anthropology and World Archaeology

• • •

Pre-OXLOD: jade (mineral) not featured here

Pre-OXLOD system

Specialist librarian

Collection data manager

Researcher / curator

GLAM contributors
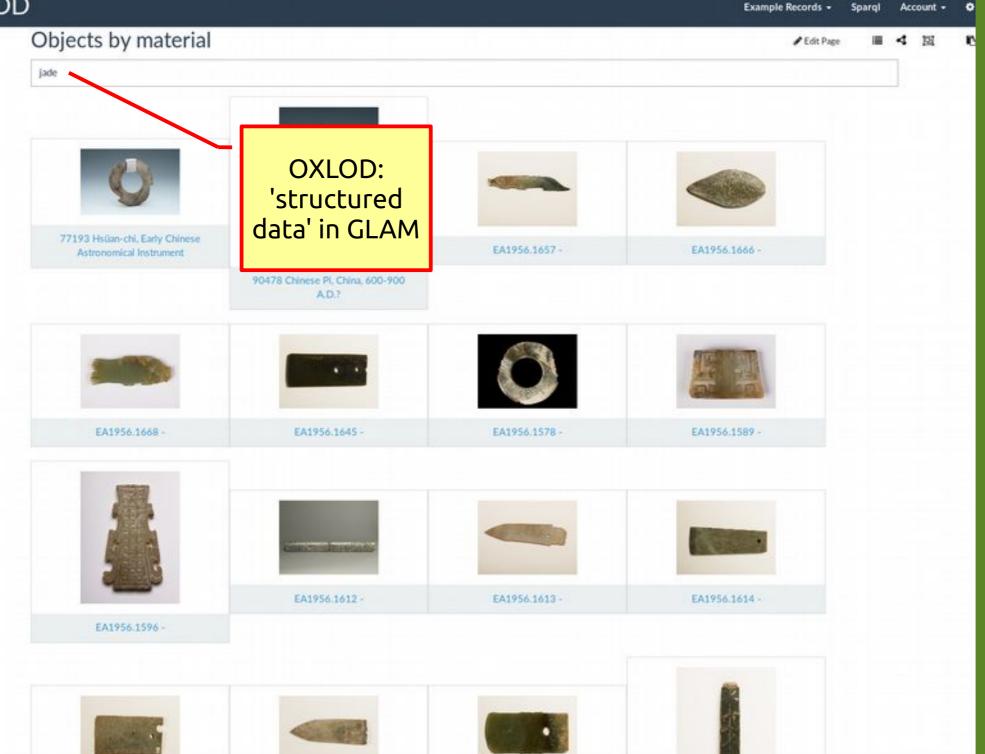
Pre-OXLOD summary

- No means of searching across GLAM datasets

- Results from different collections are difficult to compare

- Data manager's effort will need to be repeated for future exhibitions/projects

OXLOD taster
replay of
GLAM jade exhibition

OXLOD Search

✎ Edit Page

Simple search | Search based on authority | Place search

Objects by title

Search for objects e.g. type "Hsüan-chi" or "EA2012.185"

Objects by type

Search for types of objects e.g. type "xuanji"

Objects by material

Search for materials e.g. type "jade"

OXLOD:
'structured
data' enables
accurate/deep search
across collections

OXLOD: 'structured data' in GLAM

Current system

Specialist librarian

Collection data manager

Researcher / curator

GLAM contributors

OXLOD

Example Records ▾   Sparql   Account ▾   ⚙

Summary   Annotations ⓪   Explore Related

✏ Edit Page   ☰   ⬠   ▣   📑

Reconciled entities   Links to here   **Download relevant objects**   References

Download objects of this type

Download CSV
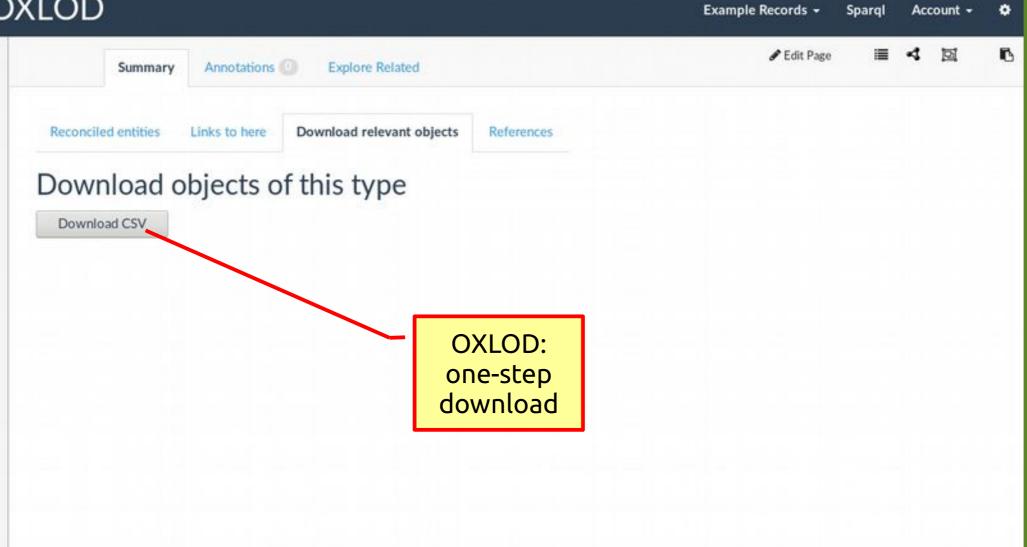
OXLOD:
one-step
download

Terms of Use

CIDOC CRM

Acknowledgements

Pre-OXLOD: 8 steps per result. OXLOD: 1 step for all.

Pre-OXLOD: 8 steps per result. OXLOD: 1 step for all.

OXLOD architecture

Base system

Base sys

Triple store

Web server

Operating system

PowerEdge R640

The fundamental building block of Linked Data is RDF

Subject — Predicate → Object

That's it.

An RDF "triple"

(RDF is the Resource Description Framework)

Base system

Workflow

Bodleian Libraries
UNIVERSITY OF OXFORD

Oxford Botanic Garden
& Arboretum
Est.1621

MUSEUM *of the*
HISTORY *of*
SCIENCE

ASHMOLEAN

Museum of
Natural History

PittRivers
MUSEUM
Anthropology and World Archaeology

export

convert
lookup
fields

Dataset
authority
lists

Automatic
(fast, low quality)

OpenRefine

convert to
XML

convert to
XML

convert to
XML

Manual
(slow, high quality)

WIKIDATA

External reconciliation

Mapping - 3M

export

RDF triples

upload

blazegraph
by SYSTAP

metaphacts

RDF Visualizer

search interface

data
export

HTML

OXLOD

Bodleian Libraries UNIVERSITY OF OXFORD

Oxford Botanic Garden & Arboretum Est.1621

MUSEUM of the HISTORY of SCIENCE

ASHMOLEAN

Museum of Natural History

PittRivers MUSEUM Anthropology and World Archaeology

export

convert lookup fields → Dataset authority lists

convert to XML

convert to XML

convert to XML

Automatic (fast, low quality)

OpenRefine

Manual (slow, high quality)

WIKIDATA

External reconciliation

Mapping - 3M

export

RDF triples

upload

blazegraph by SYSTAP

metaphacts

RDF Visualizer

search interface

data export

HTML

OXLOD

Export formats

- Allegro – key, value pairs in blocks

- BRAHMS – CSV

- EMu – XML

- MuseumPlus – SQL Server

- Bodleian MSs – Text Encoding Initiative XML

- FileMaker – CSV, XML

- MS Excel – CSV

- …

Pre-processing

- Separate lookup fields from other data

Pre-processing

- Separate lookup fields from other data

- Specific to the dataset

- Often requires programming
  - parsing records

Inconsistent input

- 01/1976

- 01/09/1993

- 02 October 2013

- 02/1983

- 1974

- 1998?

- 1999 - 2000

- 1999-2000

- 28.01.15

- 30/11 - 01/12/1991

- February 1983

- mid 20th century

- c. 1950

- 2nd half of 20th century

- early 19th century

- before 1880

- post 1910

- pre 1790
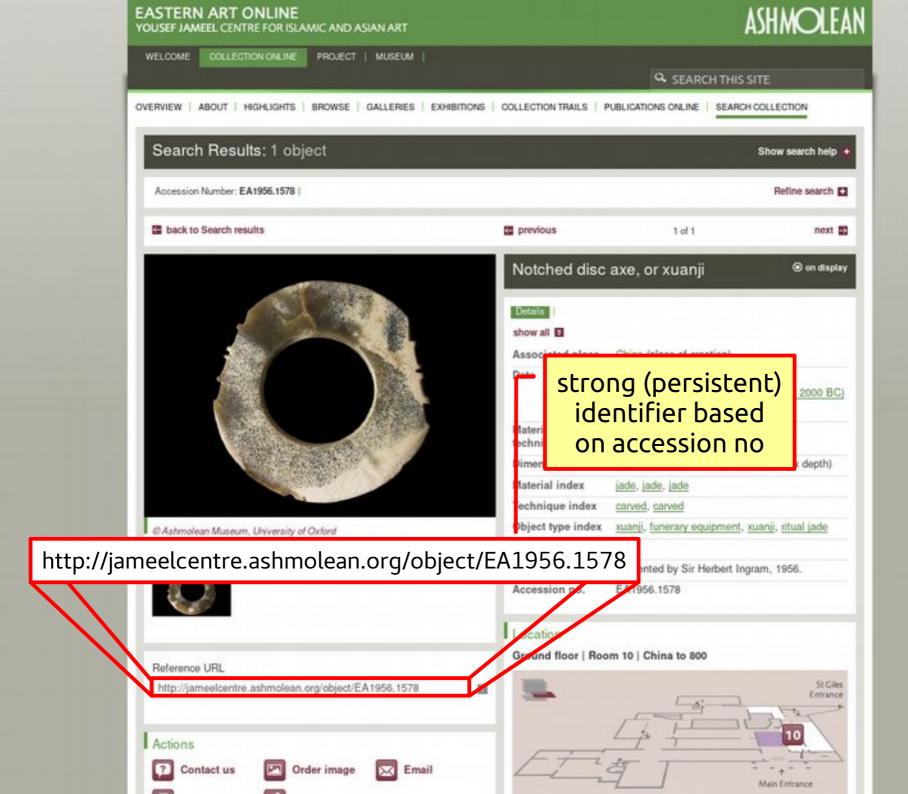
```
<date>
  <from>19510101T00:00:00</from>
  <to>20001231T23:59:59</to>
</date>
```

Pre-processing
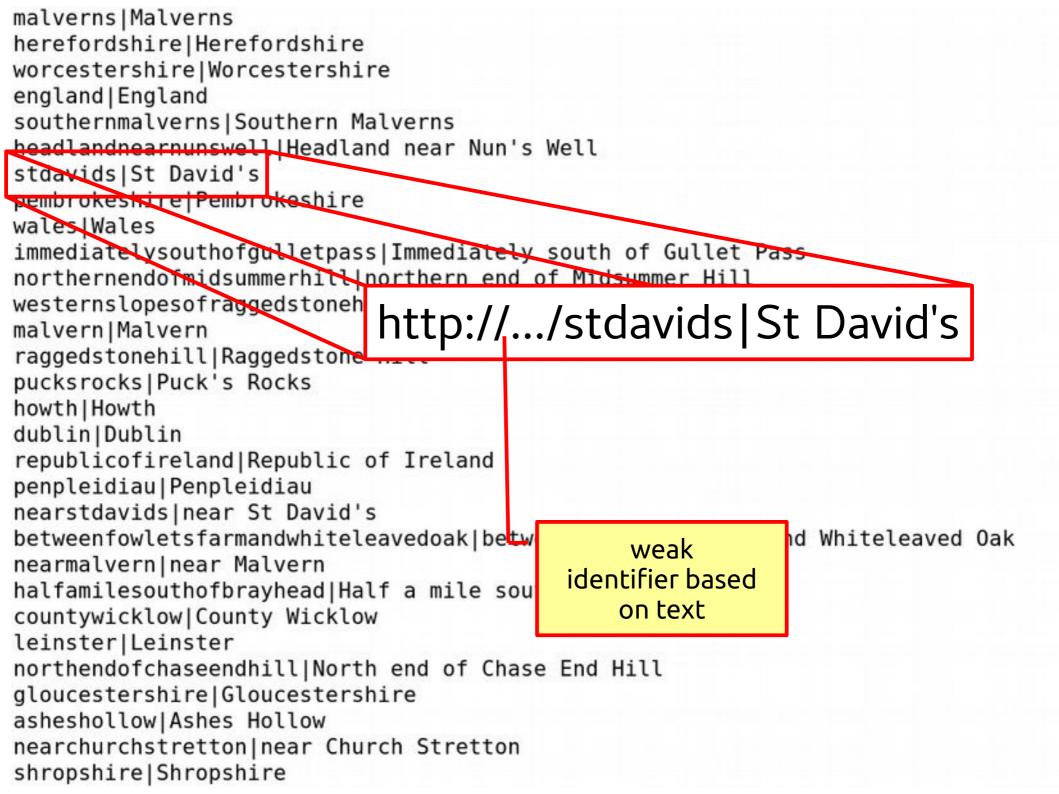
- Separate lookup fields from other data
- Specific to the dataset
- Often requires programming
  - parsing records
  - locating/creating identifers

Pre-processing
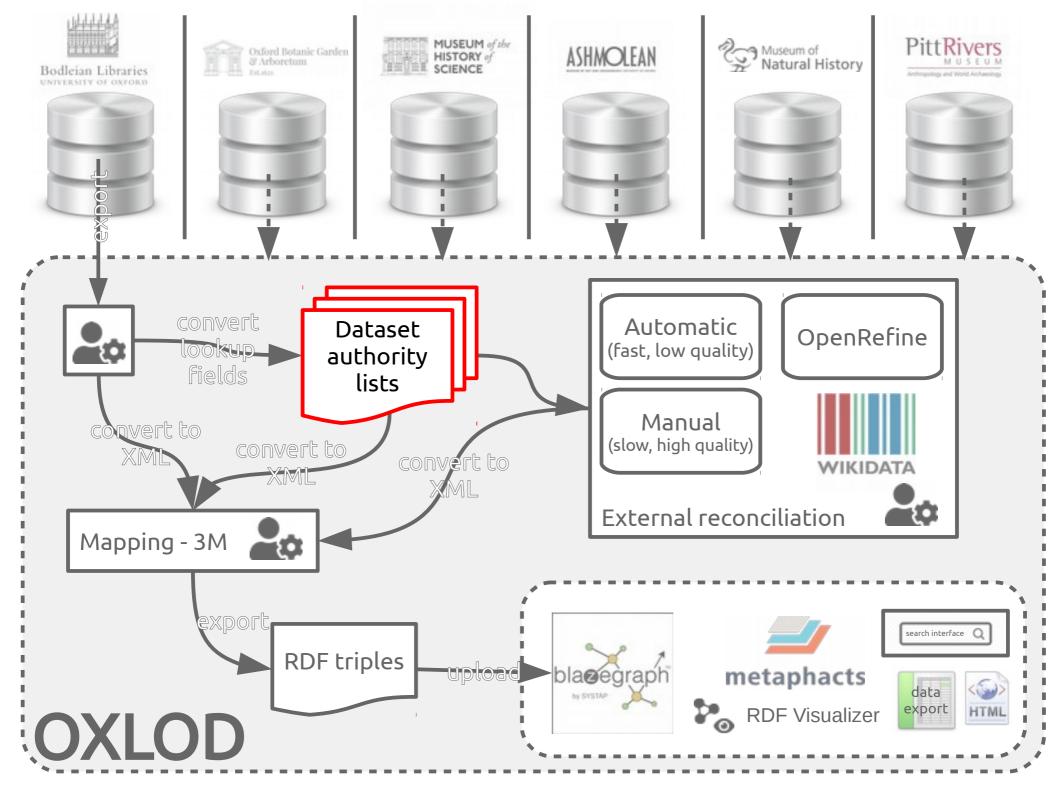
- Separate lookup fields fr

- Specific to the dataset

- Often requires programm
  - parsing records
  - locating/creating identifers

EASTERN ART ONLINE
YOUSEF JAMEEL CENTRE FOR ISLAMIC AND ASIAN ART

ASHMOLEAN

WELCOME | COLLECTION ONLINE | PROJECT | MUSEUM |

Q SEARCH THIS SITE

OVERVIEW | ABOUT | HIGHLIGHTS | BROWSE | GALLERIES | EXHIBITIONS | COLLECTION TRAILS | PUBLICATIONS ONLINE | SEARCH COLLECTION

Search Results: 1 object

Show search help +

Accession Number: EA1956.1578 |

Refine search ▣

back to Search results   previous   1 of 1   next ▣

Notched disc axe, or xuanji   ◉ on display

Details |

show all ▣

Associated place   China (place of creation)

Date   2000 BC)

strong (persistent) identifier based on accession no

Materi...
echni...

...depth)

Material index   jade, jade, jade

Technique index   carved, carved

Object type index   xuanji, funerary equipment, xuanji, ritual jade

http://jameelcentre.ashmolean.org/object/EA1956.1578

...nted by Sir Herbert Ingram, 1956.

Accession no.   EA1956.1578

Location...

Ground floor | Room 10 | China to 800

© Ashmolean Museum, University of Oxford

Reference URL

http://jameelcentre.ashmolean.org/object/EA1956.1578

Actions

? Contact us   🖼 Order image   ✉ Email

St Giles Entrance

10

Main Entrance

```
malverns|Malverns
herefordshire|Herefordshire
worcestershire|Worcestershire
england|England
southernmalverns|Southern Malverns
headlandnearnunswell|Headland near Nun's Well
stdavids|St David's
pembrokeshire|Pembrokeshire
wales|Wales
immediatelysouthofgulletpass|Immediately south of Gullet Pass
northernendofmidsummerhill|northern end of Midsummer Hill
westernslopesofraggedstonehill|Western slopes of Raggedstone Hill
malvern|Malvern
raggedstonehill|Raggedstone Hill
pucksrocks|Puck's Rocks
howth|Howth
dublin|Dublin
republicofireland|Republic of Ireland
penpleidiau|Penpleidiau
nearstdavids|near St David's
betweenfowletsfarmandwhiteleavedoak|between Fowlet's Farm and Whiteleaved Oak
nearmalvern|near Malvern
halfamilesouthofbrayhead|Half a mile south of Bray Head
countywicklow|County Wicklow
leinster|Leinster
northendofchaseendhill|North end of Chase End Hill
gloucestershire|Gloucestershire
asheshollow|Ashes Hollow
nearchurchstretton|near Church Stretton
shropshire|Shropshire
```

```
malverns|Malverns
herefordshire|Herefordshire
worcestershire|Worcestershire
england|England
southernmalverns|Southern Malverns
headlandnearnunswell|Headland near Nun's Well
stdavids|St David's
pembrokeshire|Pembrokeshire
wales|Wales
immediatelysouthofgulletpass|Immediately south of Gullet Pass
northernendofmidsummerhill|northern end of Midsummer Hill
westernslopesofraggedstoneh
malvern|Malvern
raggedstonehill|Raggedstone
pucksrocks|Puck's Rocks
howth|Howth
dublin|Dublin
republicofireland|Republic of Ireland
penpleidiau|Penpleidiau
nearstdavids|near St David's
betweenfowletsfarmandwhiteleavedoak|betw                    nd Whiteleaved Oak
nearmalvern|near Malvern
halfamilesouthofbrayhead|Half a mile sou
countywicklow|County Wicklow
leinster|Leinster
northendofchaseendhill|North end of Chase End Hill
gloucestershire|Gloucestershire
asheshollow|Ashes Hollow
nearchurchstretton|near Church Stretton
shropshire|Shropshire
```

http://.../stdavids|St David's

weak
identifier based
on text

Bodleian Libraries
UNIVERSITY OF OXFORD

Oxford Botanic Garden
& Arboretum
Est.1621

MUSEUM *of the*
HISTORY *of*
SCIENCE

ASHMOLEAN

Museum of
Natural History

PittRivers
MUSEUM
Anthropology and World Archaeology

export

convert
lookup
fields

Dataset
authority
lists

Automatic
(fast, low quality)

OpenRefine

convert to
XML

convert to
XML

convert to
XML

Manual
(slow, high quality)

WIKIDATA

External reconciliation

Mapping - 3M

export

RDF triples

upload

blazegraph
by SYSTAP

metaphacts

search interface

data
export

HTML

RDF Visualizer

OXLOD

Pre-processing

- Separate lookup fields from other data
- Specific to the dataset
- Often requires programming
  - parsing records
  - locating/creating identifers
- Bash, AWK, Python, XQuery, XSLT

Preprocessing records

Andrew Morrison
Software Engineer
Bodleian Digital Library Systems and Services

Chinese catalogue and medieval MSs

- Chinese catalogue
  - Allegro

- Export in text format
  - Records separate by blank lines
  - Records mixed together, linked by ad hoc methods
  - Fields identified by first three characters
  - Subfields delimited by pipe symbols

- Medieval Manuscripts
  - Text Encoding Initiative

- In XML format
  - Emphasis on provenance
    - Mapping Manuscript Migrations
  - Local authority files
    - Work titles
    - Person names
    - Places

Need converting into a suitable structure for mapping

Solution: XQuery

- Filter out data
  - Administrative info, tracking barcodes, etc
- Keep data to share
  - Titles, authors, publication dates and places, provenance, languages, subject classifications, identifiers
- Build XML export
  - Separate records by type
  - Normalize dates
  - Manuscripts containing parts containing works
  - Former owners linked to periods of provenance (where known)
- Identifiers
  - Match countries and languages to Getty identifiers
  - Build local URIs

Example of Allegro data

```
#0rr 19930601/12:27:54|odh
#0ca d1961
#0cc us|teng
#200 Japanese botany during the period of wood-block printing
#290 by Harley Harris Bartlett & Hide Shohara
#300 Bartlett, Harley Harris
#300aShohara Hide
#410 Los Angeles
#420 Dawson's Book Shop
#430 1961
#470 271 p. : ill. ; 24 cm
#600 Repr. from Asa Gray bulletin NS3[1961], 289-561
#803 19165 d.28

#000 19921029/15:03:09|osn
#0rr 20130520/13:53:51|osn
#0ca d1983
#0cc cc|tchi
#130 11285.6
#200 自立會史料集|9zi li hui shi liao ji
#290 杜邁之，劉泱泱，李龍如輯
#300 杜邁之|9du mai zhi
#300a劉泱泱|9liu yang yang
#300b李龍如|9li long ru
#410 長沙|9chang sha
#420 岳麓書社|9yue lu shu she
#430 1983
#470 平裝1冊(371頁) : 圖 ; 20公分
#800 Chin.e.7731
#900 r
```

Example of resulting XML

```xml
<?xml version="1.0" encoding="UTF-8"?>
<books>
  <book>
    <uri>http://allegro.bodleian.ox.ac.uk/chinese/19920109163415odh</uri>
    <title>
      <main>Japanese botany during the period of wood-block printing</main>
    </title>
    <author>
      <uri>http://allegro.bodleian.ox.ac.uk/chinese/author/name/bartlett%2C%20harley%20harris</uri>
      <main>Bartlett, Harley Harris</main>
    </author>
    <author>
      <uri>http://allegro.bodleian.ox.ac.uk/chinese/author/name/shohara%20hide</uri>
      <main>Shohara Hide</main>
    </author>
    <shelfmark>RSL 19165 d.28</shelfmark>
    <date>
      <from>1961-01-01T00:00:00</from>
      <to>1961-12-31T23:59:59</to>
    </date>
    <pubplace>
      <main>Los Angeles</main>
    </pubplace>
    <pubcountry>
      <uri>http://vocab.getty.edu/tgn/7012149</uri>
      <label>United States</label>
      <code>us</code>
    </pubcountry>
    <language>
      <uri>http://vocab.getty.edu/aat/300388277</uri>
      <label>English</label>
      <code>eng</code>
    </language>
  </book>
  <book>
    <uri>http://allegro.bodleian.ox.ac.uk/chinese/19921029150309osn</uri>
    <title>
      <main>自立會史料集</main>
      <alt>zi li hui shi liao ji</alt>
    </title>
    <author>
      <uri>http://allegro.bodleian.ox.ac.uk/chinese/author/name/%E6%9D%9C%E9%82%81%E4%B9%8B</uri>
      <main>杜邁之</main>
```

Example of provenance in TEI
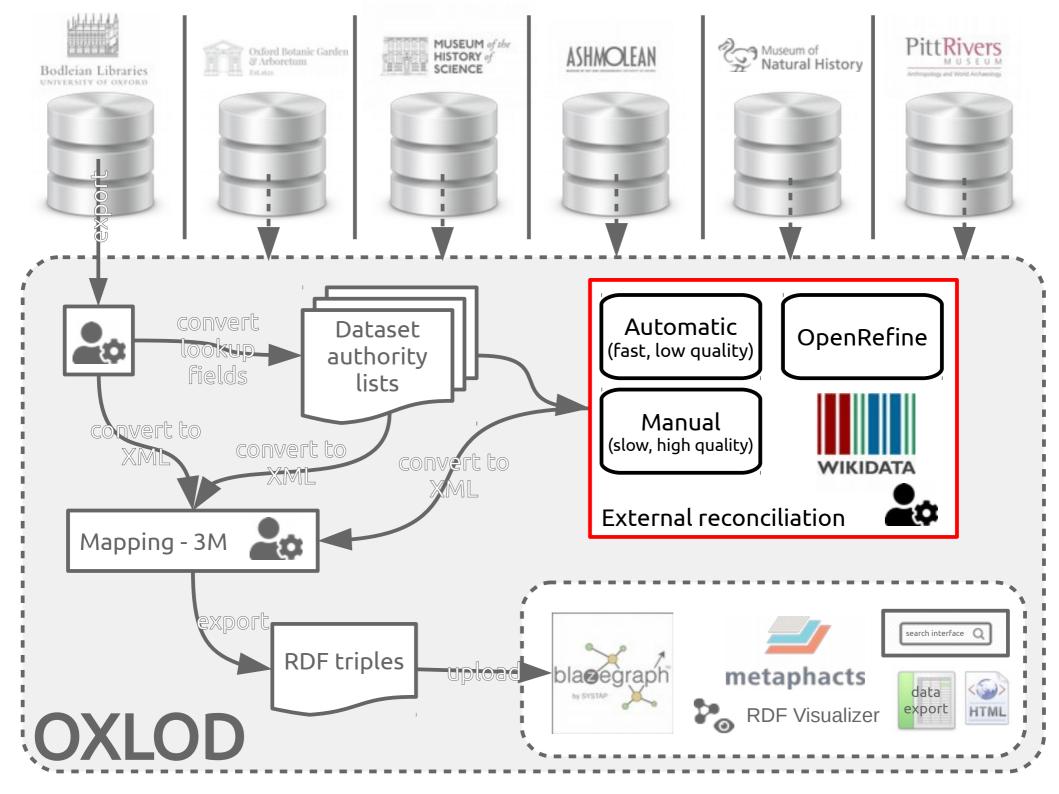
```xml
<history>
   <origin resp="#PA">
      <origDate calendar="Gregorian" notAfter="1450" notBefore="1400">15th century, first half</origDate>
      <origPlace>
         <country key="place_7002445">English</country>
      </origPlace>
   </origin>
   <provenance resp="#VD">Liturgical evidence for <placeName key="place_7012044">Canterbury</placeName> (?):
      tag for the mass of St. Augustine of Canterbury, and, in the calendar, the Translations of St. Thomas,
      St. Edmund and St. Richard, all in red.</provenance>
   <provenance resp="#SC">'Willelmus Brocas', p. 160, 16th-17th cent.</provenance>
   <provenance resp="#SC">'Willelmus Fitz Willelmus', p. 260 rev, 16th-17th cent.</provenance>
   <provenance notBefore="1633" notAfter="1660"><persName role="formerOwner" key="person_76667490">William
      Sprigg, c. 1633 - c. 1701</persName>; given by him to Anthony Wood.</provenance>
   <!-- Sprigg was ejected from Lincoln College in 1660: ODNB -->
   <provenance resp="#SC" notAfter="1690" notBefore="1660">'Liber <persName role="formerOwner"
      key="person_14803186">Antonii Wood</persName> [1632-1695] e coll' Merton Oxonie ex dono [Willelmi]
      Sprigges nuper socii coll' Lync' Oxon' Novembris' [after 1657], p. iii.</provenance>
   <acquisition when="1690">Bought by the library in 1690.</acquisition>
   <!-- check: index other names -->
</history>
```

Example of resulting XML

```xml
<date context="origin"><from>1400-01-01T00:00:00</from><to>1450-12-31T23:59:59</to></date>
<date context="acquisition"><from>1690-01-01T00:00:00</from><to>1690-12-31T23:59:59</to></date>
<place context="origin">
    <uri>https://medieval.bodleian.ox.ac.uk/catalog/place_7002445</uri>
    <label>English</label>
</place>
<provenance xml:id="manuscript_9055_prov1">
    <place>
        <uri>https://medieval.bodleian.ox.ac.uk/catalog/place_7012044</uri>
        <label>Canterbury</label>
    </place>
    <text>Liturgical evidence for Canterbury (?): tag for the mass of St. Augustine of Canterbury, and, in the calendar, the
        Translations of St. Thomas, St. Edmund and St. Richard, all in red.</text>
</provenance>
<inscription xml:id="manuscript_9055_prov2">
    <text>'Willelmus Brocas', p. 160, 16th-17th cent.</text>
</inscription>
<inscription xml:id="manuscript_9055_prov3">
    <text>'Willelmus Fitz Willelmus', p. 260 rev, 16th-17th cent.</text>
</inscription>
<provenance xml:id="manuscript_9055_prov4">
    <date><from>1633-01-01T00:00:00</from><to>1660-12-31T23:59:59</to></date>
    <person>
        <uri>https://medieval.bodleian.ox.ac.uk/catalog/person_76667490</uri>
        <label>William Sprigg, c. 1633 - c. 1701</label>
        <role>formerOwner</role>
    </person>
    <text>William Sprigg, c. 1633 - c. 1701; given by him to Anthony Wood.</text>
</provenance>
<inscription xml:id="manuscript_9055_prov5">
    <date><from>1660-01-01T00:00:00</from><to>1690-12-31T23:59:59</to></date>
    <person>
        <uri>https://medieval.bodleian.ox.ac.uk/catalog/person_14803186</uri>
        <label>Antonii Wood</label>
        <role>formerOwner</role>
    </person>
    <text>'Liber Antonii Wood [1632-1695] e coll' Merton Oxonie ex dono [Willelmi] Sprigges nuper socii coll' Lync' Oxon'
        Novembris' [after 1657], p. iii.</text>
</inscription>
```

```
malverns|Malverns
herefordshire|Herefordshire
worcestershire|Worcestershire
england|England
southernmalverns|Southern Malverns
headlandnearnunswell|Headland near Nun's Well
stdavids|St David's
pembrokeshire|Pembrokeshire
wales|Wales
immediatelysouthofgulletpass|Immediately south of Gullet Pass
northernendofmidsummerhill|northern end of Midsummer Hill
westernslopesofraggedstoneh
malvern|Malvern
raggedstonehill|Raggedstone
pucksrocks|Puck's Rocks
howth|Howth
dublin|Dublin
republicofireland|Republic of Ireland
penpleidiau|Penpleidiau
nearstdavids|near St David's
betweenfowletsfarmandwhiteleavedoak|betw                   d Whiteleaved Oak
nearmalvern|near Malvern
halfamilesouthofbrayhead|Half a mile sou
countywicklow|County Wicklow
leinster|Leinster
northendofchaseendhill|North end of Chase End Hill
gloucestershire|Gloucestershire
asheshollow|Ashes Hollow
nearchurchstretton|near Church Stretton
shropshire|Shropshire
```

http://.../stdavids|St David's

weak
identifier based
on text

Saint David's

Source: http://vocab.getty.edu/tgn/7008839

Subject (38)   Predicate   Object   All   Website | Hierarchy | Download in: JSON | JSONLD | RDF | N3/Turtle | N-Triples

http://vocab.getty.edu/tgn/7008839

Inference   [Explicit only ▾]

Statements in which the resource exists as a subject.

Predicate	Object
rdf:type	gvp:AdminPlaceConcept
rdfs:seeAlso	http://www.getty.edu/vow/TGNFullDisplay?find=&place=&nation=&subjectid=7008839
dcterms:created	1991-09-13T04:48:00
dcterms:modified	2011-09-16T12:55:43, 2011-09-16T12:55:44
skos:changeNote	tgn_rev:5000891297, tgn_rev:5013622493, tgn_rev:5013622500
gvp:parentString	Pembrokeshire, Wales, United Kingdom, Europe, World
gvp:parentStringAbbrev	Pembrokeshire, Wales, ... World
gvp:displayOrder	226
xl:prefLabel	tgn_term:65407-en
xl:altLabel	tgn_term:1001833420-cy, tgn_term:170429, tgn_term:173996
gvp:broaderPartitive	tgn:7018993
gvp:broaderPreferred	tgn:7018993
gvp:prefLabelGVP	tgn_term:65407-en
skos:inScheme	tgn
skos:scopeNote	tgn_scopeNote:48292
dcterms:contributor	tgn_contrib:10000000, tgn_contrib:10000001, tgn_contrib:10000003
dcterms:source	tgn_source:2009007144-subject-7008839, tgn_source:9005218-subject-7008839, tgn_source:9006249, tgn_source:9006449, tgn_source:9006548-subject-7008839, tgn_source:9006561
gvp:placeTypePreferred	aat:300008347

External authorities in OXLOD

- Virtual Authority Files (VIAF)
- International Plant Names Index (IPNI)
- Getty Thesauri
  - Arts and Architecture Thesaurus (AAT)
  - Thesaurus of Geographic Names (TGN)
- British Museum Thesaurus
- China Bibliographical Database (CBDB)
- China Historical Geographic Information System (CHGIS)
- Geonames
- Wikidata
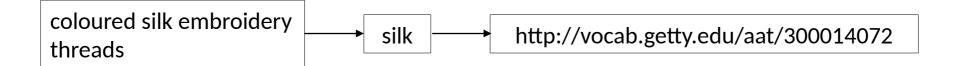
Reconciling records
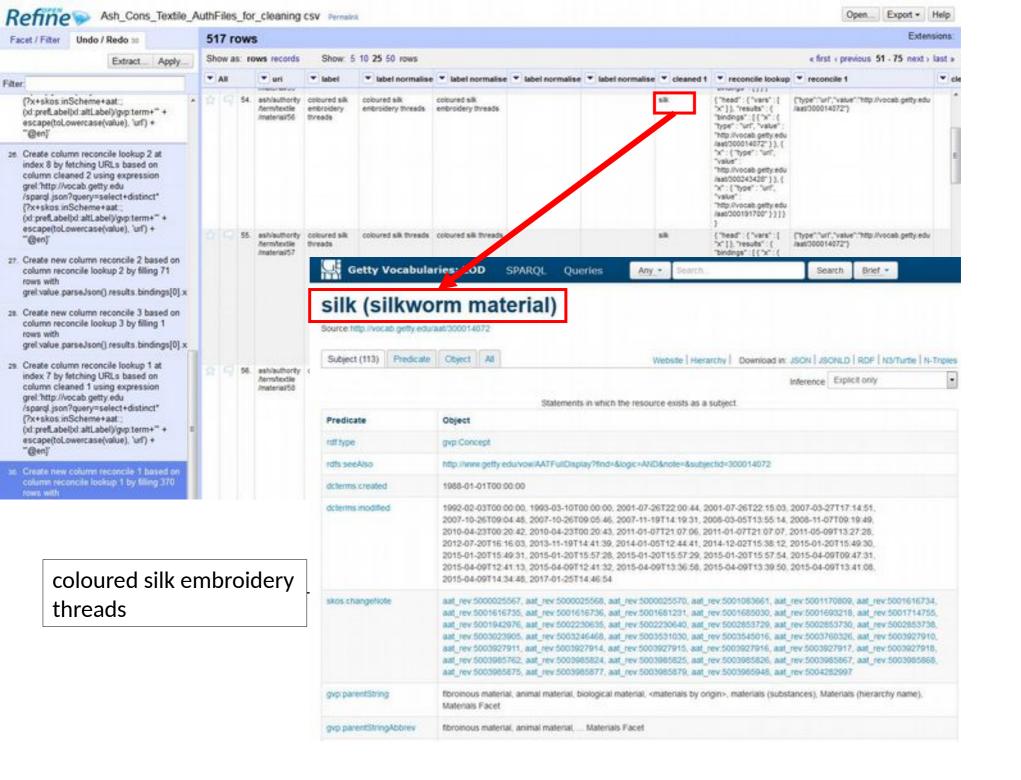
Sarah Joomun
Digital Collections Manager
Oxford University Museum of Natural History

OpenRefine interface screenshot:
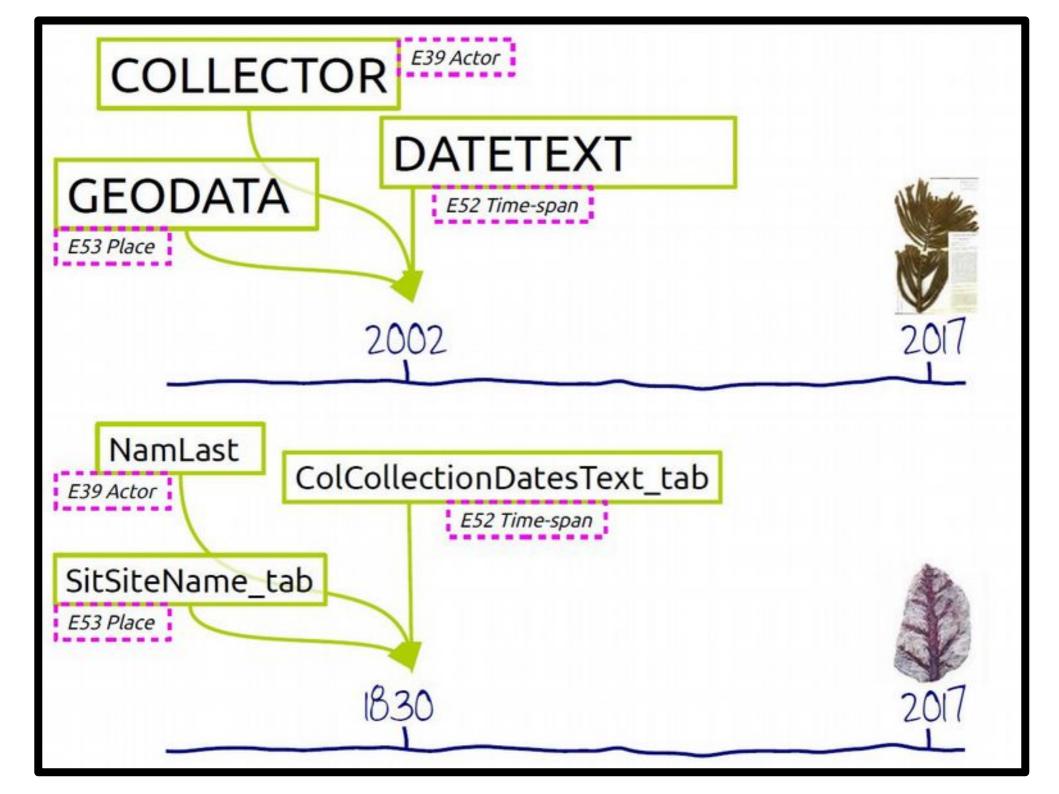
Refine — Ash_Cons_Textile_AuthFiles_for_cleaning csv Permalink

Open... Export ▾ Help

Facet / Filter   Undo / Redo 30   Extensions:

Extract... Apply...

Filter:

517 rows

Show as: rows records   Show: 5 10 25 50 rows   « first ‹ previous 51 - 75 next › last »

Undo/Redo history:

- [?x+skos:inScheme+aat:; (xd:prefLabel|xd:altLabel)/gvp:term+"' + escape(toLowercase(value), 'url') + "@en)'

26. Create column reconcile lookup 2 at index 8 by fetching URLs based on column cleaned 2 using expression grel:'http://vocab.getty.edu /sparql.json?query=select+distinct" {?x+skos:inScheme+aat:; (xd:prefLabel|xd:altLabel)/gvp:term+"' + escape(toLowercase(value), 'url') + "@en)'

27. Create new column reconcile 2 based on column reconcile lookup 2 by filling 71 rows with grel:value.parseJson().results.bindings[0].x

28. Create new column reconcile 3 based on column reconcile lookup 3 by filling 1 rows with grel:value.parseJson().results.bindings[0].x

29. Create column reconcile lookup 1 at index 7 by fetching URLs based on column cleaned 1 using expression grel:'http://vocab.getty.edu /sparql.json?query=select+distinct" {?x+skos:inScheme+aat:; (xd:prefLabel|xd:altLabel)/gvp:term+"' + escape(toLowercase(value), 'url') + "@en)'

30. Create new column reconcile 1 based on column reconcile lookup 1 by filling 370 rows with

Data table columns: All | uri | label | label normalise | label normalise | label normalise | label normalise | cleaned 1 | reconcile lookup | reconcile 1 | cl...

54.	ash/authority /term/textile /material/56	coloured silk embroidery threads	coloured silk embroidery threads	coloured silk embroidery threads			silk	{ "head" : { "vars" : ["x"]}, "results" : { "bindings" : [{ "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300014072" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300243428" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300191700" }}]}}	{"type":"uri","value":"http://vocab.getty.edu /aat/300014072"}
55.	ash/authority /term/textile /material/57	coloured silk threads	coloured silk threads	coloured silk threads			silk	{ "head" : { "vars" : ["x"]}, "results" : { "bindings" : [{ "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300014072" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300243428" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300191700" }}]}}	{"type":"uri","value":"http://vocab.getty.edu /aat/300014072"}
56.	ash/authority /term/textile /material/58	coloured silks	coloured silks	coloured silks			silk	{ "head" : { "vars" : ["x"]}, "results" : { "bindings" : [{ "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300014072" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300243428" }}, { "x" : { "type" : "uri", "value" : "http://vocab.getty.edu /aat/300191700" }}]}}	{"type":"uri","value":"http://vocab.getty.edu /aat/300014072"}

Diagram:

coloured silk embroidery threads → silk → http://vocab.getty.edu/aat/300014072

coloured silk embroidery threads

COLLECTOR — *E39 Actor*

DATETEXT — *E52 Time-span*

GEODATA — *E53 Place*

2002

2017

NamLast — *E39 Actor*

ColCollectionDatesText_tab — *E52 Time-span*

SitSiteName_tab — *E53 Place*

1830

2017

MHS EMu catalogue module, Summary view, 2007-2018

the DoDo & Civet by G. Edwards F.R.S AD 1759.

Info   Matching Table   Generators   Analysis   Transformation   Configuration   About

▲ TOP   ▼ BOTTOM   ◉ VIEW MODE   </> XML

(ALL) SOURCES ↔   (ALL) TARGETS ↔   (ALL) IF RULES ↔   (ALL) COMMENTS ↔   (ALL) MAPS ↕

Click on a row to edit the matching table

#		SOURCE		TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
1	D	/dataroot/dbo_tObject		crm:E22_Man-Made_Object			
1.1	P	↓ Accno	↓	crm:P1_is_identified_by			
	R	☐ Accno	☐	crm:E42_Identifier			
1.2	P	↓ Collection	↓	crm:P46i_forms_part_of			
	R	☐ Collection	☐	crm:E78_Collection			
1.3	P	↓ CreditLine	↓	crm:P3_has_note			
	R	☐ CreditLine	☐	http://www.w3.org/2000/01/rdf-schema#Literal			
1.4	P	↓ dbo_tObjectDates/DateText	↓	crm:P108i_was_produced_by			
			▩	crm:E12_Production ⊛ [p]			
			↓	crm:P4_has_time-span			
			▩	crm:E52_Time-Span ⊛ [ts]			
			↓	crm:P3_has_note			
	R	☐ dbo_tObjectDates/DateText	☐	http://www.w3.org/2000/01/rdf-schema#Literal	../DateType/text() = display	(We do not want to include peri..)
1.5	P	↓ dbo_tObjectDates/DateFrom	↓	crm:P108i_was_produced_by			
			▩	crm:E12_Production ⊛ [p]			
			↓	crm:P4_has_time-span			
			▩	crm:E52_Time-Span ⊛ [ts]			
			↓	crm:P82a_begin_of_the_begin			
	R	☐ dbo_tObjectDates/DateFrom	☐	http://www.w3.org/2000/01/rdf-schema#Literal	../DateType/text() = display	
1.6	P	↓ dbo_tObjectDates/DateTo	↓	crm:P108i_was_produced_by			
			▩	crm:E12_Production ⊛ [p]			
			↓	crm:P4_has_time-span			
			▩	crm:E52_Time-Span ⊛ [ts]			
			↓	crm:P82b_end_of_the_end			
	R	☐ dbo_tObjectDates/DateTo	☐	http://www.w3.org/2000/01/rdf-schema#Literal	../DateType/text() = display	
1.7	P	↓ dbo_uri_Technique_REL	↓	crm:P108i_was_produced_by			
			▩	crm:E12_Production ⊛ [p]			
			↓	crm:P32_used_general_technique			
	R	☐ dbo_uri_Technique_REL	☐	crm:E55_Type			
1.8	P	↓ dbo_uri_Material_REL	↓	crm:P45_consists_of			
	R	☐ dbo_uri_Material_REL	☐	crm:E57_Material			

Info   **Matching Table**   Generators   Analysis   Transformation   Configuration   About

▲ TOP   ▼ BOTTOM   👁 VIEW MODE   </> XML

(ALL) SOURCES ▸▸   (ALL) TARGETS ▸▸   (ALL) IF RULES ▸▸   (ALL) COMMENTS ▸▸   (ALL) MAPS ⇕

Click on a row to edit the matching table

#		SOURCE		TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
1	D	/dataroot/dbo_tObject		crm:E22_Man-Made_Object			
1.1	P	Accno	↓	crm:P1_is_identified_by			
	R	Accno		crm:E42_Identifier			
1.2	P	Collection	↓	crm:P46i_forms_part_of			
	R	Collection		crm:E78_Collection			
1.3	P	CreditLine	↓	crm:P3_has_note			
	R	CreditLine		http://www.w3.org/2000/01/rdf-schema#Literal			
1.4	P	dbo_tObjectDates/DateText	↓	crm:P108i_was_produced_by			
				crm:E12_Production ☀ [p]			
			↓	crm:P4_has_time-span			
				crm:E52_Time-Span ☀ [ts]			
			↓	crm:P3_has_note			
	R	dbo_t...		http://www...	../DateType/text() = display	(We do not want to include peri...)	
1.5	P	dbo_t...	↓	crm:P108i_...			
				crm:... Pr...			
			↓	crm:P4_has...			
				crm:E52_Ti...			
			↓	crm:P82a_b...			
	R	dbo_tObjectDates/DateFrom		http://www.w3.org/2000/01/rdf-schema#Literal	../DateType/text() = display		
1.6	P	dbo_tObjectDates/DateTo	↓	crm:P108i_was_produced_by			
				crm:E12_Production ☀ [p]			
			↓	crm:P4_has_time-span			
				crm:E52_Time-Span ☀ [ts]			
			↓	crm:P82b_end_of_the_end			
	R	dbo_tObjectDates/DateTo		http://www.w3.org/2000/01/rdf-schema#Literal	../DateType/text() = display		
1.7	P	dbo_uri_Technique_REL	↓	crm:P108i_was_produced_by			
				crm:E12_Production ☀ [p]			
			↓	crm:P32_used_general_technique			
	R	dbo_uri_Technique_REL		crm:E55_Type			
1.8	P	dbo_uri_Material_REL	↓	crm:P45_consists_of			
	R	dbo_uri_Material_REL		crm:E57_Material			

Mapping records

Chris Powell
Digital Collections Team
Ashmolean Museum

Regular tasks for the Digital Collections Team

- Improving the quantity and quality of data stored in the Collections Management System (CMS).

- Mapping legacy databases to the CMS.

- Devising appropriate schemes for transcribing and importing paper records.

- Measurable validation of data in the CMS.

- Turning down requests by other institutions for our data.

Relational Table data

dbo_tObjectDates

OdaId	OdaObjId	DateType	DateText	DateFrom	DateTo
93645	389571	display	1700 - 1780	1700	1780
93646	389572	display	1700 - 1740	1700	1740
93647	389573	display	1680 - 1720	1680	1720
93670	389595	display (peri	Ming Dynasty,	1522	1567
93671	389597	display	18th century	1701	1800
93673	389599	display	16th century	1501	1600
93676	389602	display	16th century	1501	1600
93679	389605	display	1755	1755	1755
93680	389606	display (peri	Qing Dynasty,	1736	1795
93681	389607	display	1724	1724	1724
175768	361806	display	1961	1961	1961
168541	696655	display	1740 - 1780	1740	1780
139823	738084	display (peri	Yuan Dynasty	1279	1368
202153	385346	display	1946 - 1948	1946	1948
304983	789394	display	20th century	1901	2000

Relational Table data

dbo_tObjectDates

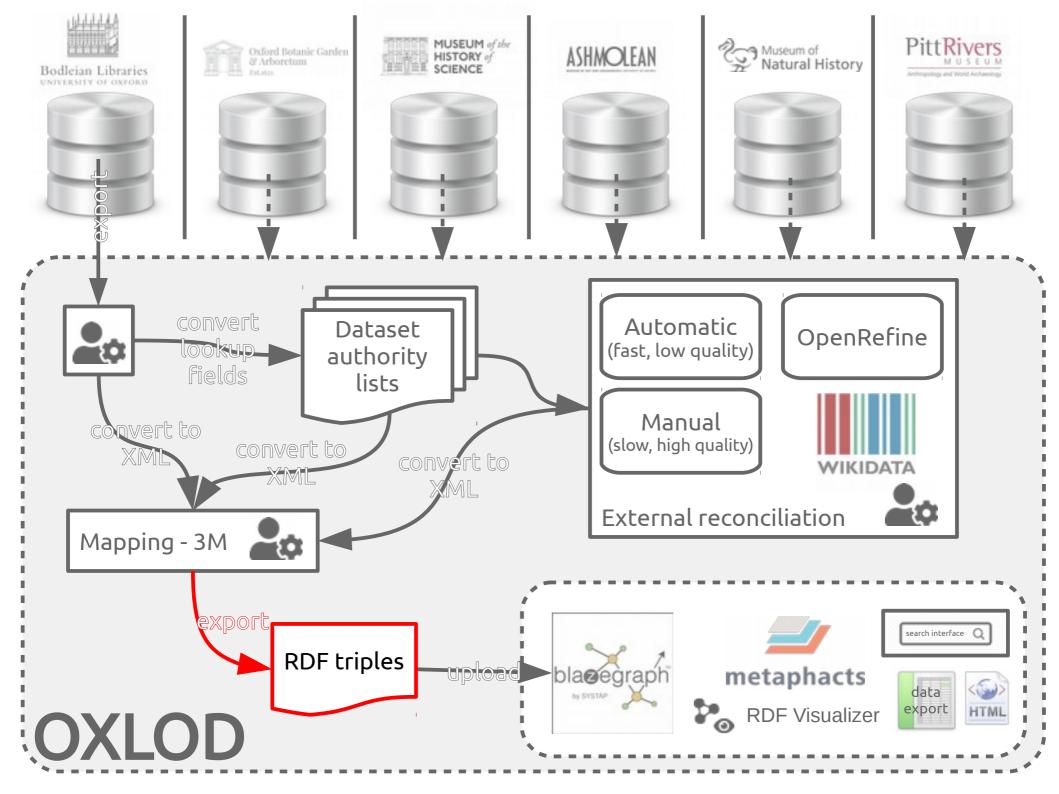| OdaId	OdaObjId	DateType	DateText	DateFrom	DateTo
93645	389571	display	1700 - 1780	1700	1780
93646	389572	display	1700 - 1740	1700	1740
93647	389573	display	1680 - 1720	1680	1720
93670	389595	display (peri	Ming Dynasty,	1522	1567
93671	389597	display	18th century	1701	1800
93673	389599	display	16th century	1501	1600
93676	389602	display	16th century	1501	1600
93679	389605	display	1755	1755	1755
93680	389606	display (peri	Qing Dynasty,	1736	1795
93681	389607	display	1724	1724	1724
175768	361806	display	1961	1961	1961
168541	696655	display	1740 - 1780	1740	1780
139823	738084	display (peri	Yuan Dynasty	1279	1368
202153	385346	display	1946 - 1948	1946	1948
304983	789394	display	20th century	1901	2000

Export to XML

Describing column semantics: Mapping

Impact on tasks of the Digital Collections Team

In addition to gaining hands-on experience of the tools and techniques needed to publish LOD I have learned that:

- Mapping to CIDOC:CRM – retain source database structure (if normalised!)
- We should record the uri of reconciled data in our CMS (Getty).
- Combining our data with that of other institutions in new ways becomes possible (BM/Egypt)
- LOD expands possibilities for collaborative research (BM).
- Meet external demands for transportable exports of our data.

Bodleian Libraries
UNIVERSITY OF OXFORD

Oxford Botanic Garden
& Arboretum
Est.1621

MUSEUM *of the*
HISTORY *of*
SCIENCE

ASHMOLEAN

Museum of
Natural History

PittRivers
MUSEUM
Anthropology and World Archaeology

export

convert
lookup
fields

Dataset
authority
lists

convert to
XML

convert to
XML

convert to
XML

Automatic
(fast, low quality)

OpenRefine

Manual
(slow, high quality)

WIKIDATA

External reconciliation

Mapping - 3M

export

RDF triples

upload

blazegraph
by SYSTAP

metaphacts

search interface

RDF Visualizer

data
export

HTML

OXLOD

```xml
<mapping>
    <domain>
        <source_node>/dataroot/dbo_tObject</source_node>
        <target_node>
            <entity>
                <type>crm:E22_Man-Made_Object</type>
                <instance_generator name="Entity">
                    <arg name="identifier" type="xpath">ObjId/text()</arg>
                </instance_generator>
                <label_generator name="Literal">
                    <arg name="text" type="xpath">concat(Accno/text(),' - ',Title/text())</arg>
                    <arg name="language" type="constant">en</arg>
                </label_generator>
            </entity>
        </target_node>
    </domain>
</mapping>
```
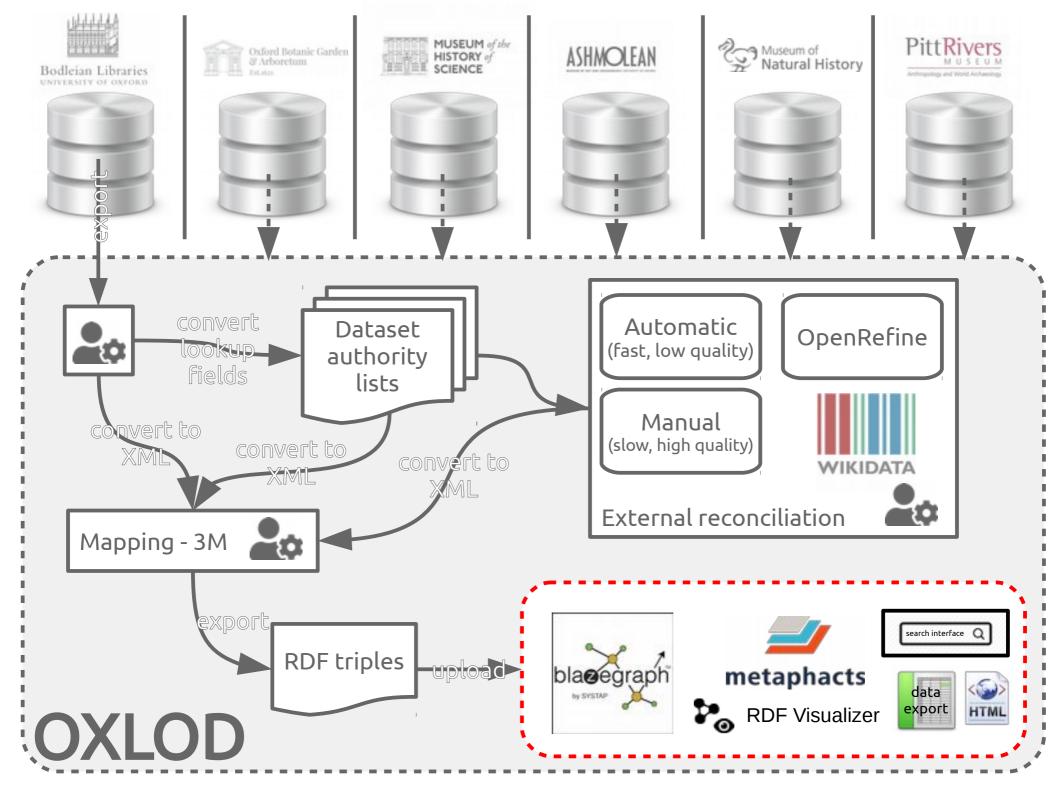
X3ML Engine supports the data transformation which is part of the data provisioning and aggregation process.

⟳ **943** commits   ⎇ **6** branches   ◌ **20** releases   👥 **4** contributors   ⚖ Apache-2.0

Branch: **master ▾**   New pull request   Create new file   Upload files   Find file   **Clone or download ▾**

This branch is 759 commits ahead of delving:master.   ⓘ Pull request ⑇ Compare

🖻 **ymark** Merge branch 'new_generator-issue_129'   Latest commit 3f99022 20 days ago

📁 docs	new issue	20 days ago
📁 example	fixed examples	3 months ago
📁 resources/X3ML Schema	Merge branch 'master' into namedgraphs	3 months ago
📁 src	added junit tests for testing UriExistingOrNew generator	20 days ago
📄 .gitignore	updated gitignore	5 years ago
📄 LICENSE	Initial commit	5 years ago
📄 README.md	Update README.md	4 months ago
📄 pom.xml	removed duplicate dependency	3 months ago

▤▤ **README.md**

X3ML Engine

Introduction

The X3ML engine handles the URI generation and the data transformation steps of the data provision and aggregation process. It realizes the transformation of the source records to the target's format. The engine takes as input the source data (currently in the form of an XML document), the description of the mappings in the X3ML mapping definition file and the URI generation policy file and is responsible for transforming the XML document into a valid RDF document which is equivalent with the XML

5min break

OXLOD live demo

OXLOD:
enriched data
e.g. links from
Ashmolean...

OXLOD

Example Records ▾   Sparql   Account ▾   ⚙

✎ Edit Page   ☰   ⬍   ⬚   ⎘

Suzhou Fu   ◷ Place record, (http://maps.cga.harvard.edu/tgaz/)

Summary   Annotations ⓪   Explore Related

Map   People   Objects   References

Map

OXLOD:
enriched data
e.g. links to China
Historical GIS
(Harvard)

OXLOD

Example Records ▾   Sparql   Account ▾   ⚙

✏ Edit Page

Suzhou Fu ⦿ Place record, (http://maps.cga.harvard.edu/tgaz/)

Summary   Annotations ⓪   Explore Related

Map   People   Objects   References

OXLOD: cross-collection linking to Bodleian

Map

OXLOD

Example Records ▾   Sparql   Account ▾   ⚙

✎ Edit Page   ☰   ⋖   ⌘   ▣

Suzhou Fu   ⊙ Place record, (http://maps.cga.harvard.edu/tgaz/)

Summary   Annotations ⓪   Explore Related

Map   People   Objects   **References**

Shown on

No recorded depictions.

Reference in

Filter Results

Object	Dataset
真理問答	http://allegro.bodleian.ox.ac.uk/chinese/
聖書典論	http://allegro.bodleian.ox.ac.uk/chinese/
真道略論. 上卷	http://allegro.bodleian.ox.ac.uk/chinese/
蒙童訓 : 蘇州土白	http://allegro.bodleian.ox.ac.uk/chinese/
蘇州舊街巷圖錄 = Traditional Suzhou streets and lanes	http://allegro.bodleian.ox.ac.uk/chinese/
蘇州博物館藏明清書畫 = ^The^ paintings and calligraphy of Ming and Qing dynasties collected by Suzhou Museum	http://allegro.bodleian.ox.ac.uk/chinese/
宿州文物 = Cultural relics in Suzhou	http://allegro.bodleian.ox.ac.uk/chinese/
東吳大學中國藝術史集刊 = Soochow university journal of Chinese art history	http://allegro.bodleian.ox.ac.uk/chinese/
東吳文史學報 = Soochow journal of humanities	http://allegro.bodleian.ox.ac.uk/chinese/
東吳學報 = Soochow journal of literature & social studies	http://allegro.bodleian.ox.ac.uk/chinese/

OXLOD: **cross-collection linking** to Bodleian

OXLOD

✏ Edit Page   ☰   ◁   ▣   ▣

OXF - (Wall.) Makino - A - s.n. - A - 2008 - India

Object (Man-Made Object) record,
(http://herbaria.plants.ox.ac.uk/)

Summary   Annotations ⓪   Explore Related
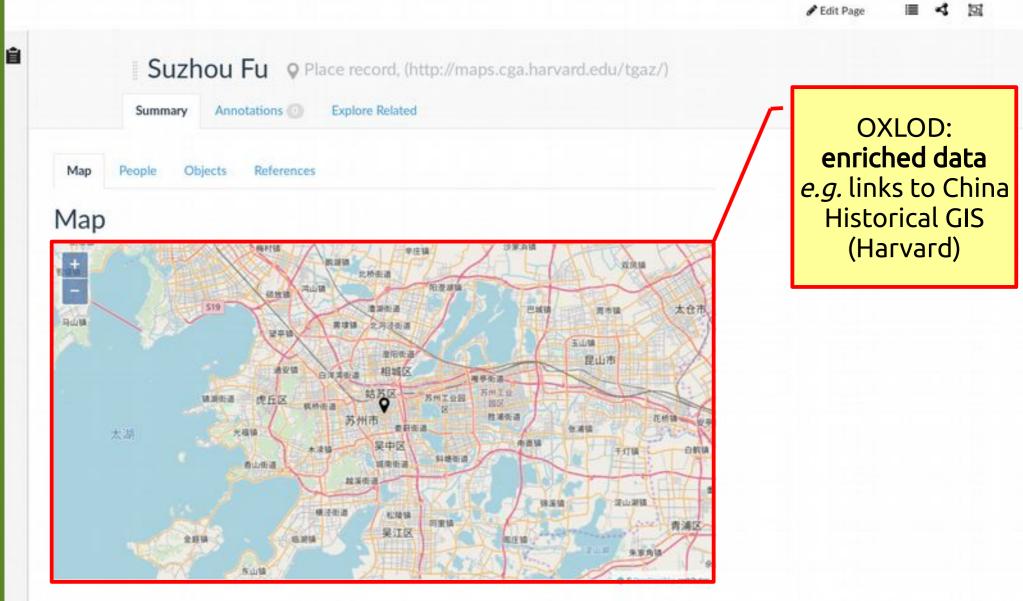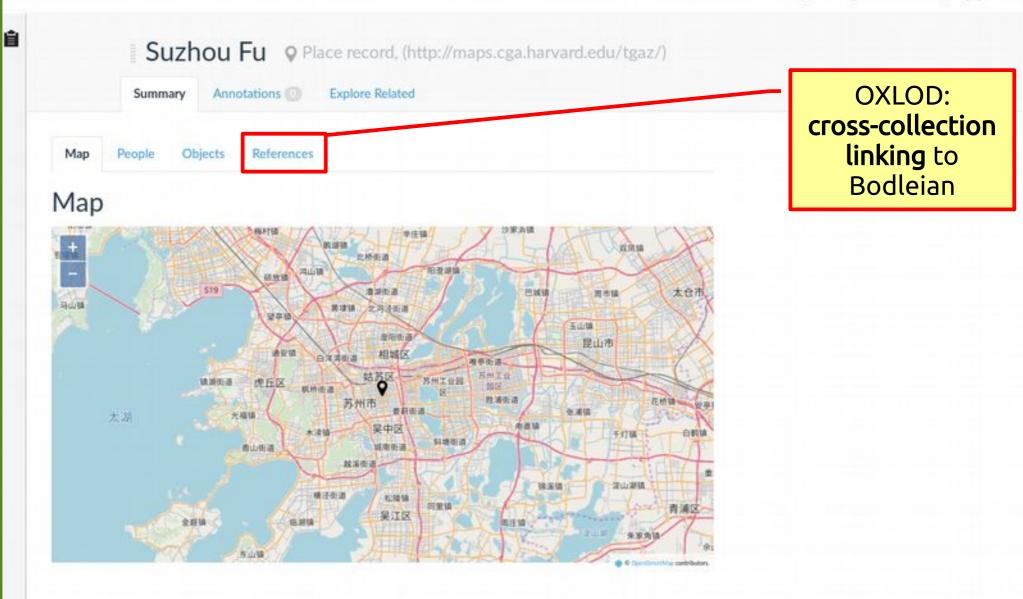
Links from here   Links to here   Production

Filter Results

Property	Object
has type	Cardiocrinum giganteum
has type	Liliaceae
has type	Cardiocrinum
has type	NOT CULTIVATED
has type	NOT PHENOTYPE IN CULTIVATION
has type	V
has preferred identifier	OXF - (Wall.) Makino - A - s.n. - A - 2008 - India
forms part of	Container for OXF - (Wall.) Makino - A - s.n. - A - 2008 - India
O5i_was_removed_by	Collection of specimen OXF - (Wall.) Makino - A - s.n. - A - 2008 - India
type	Man-Made Object

« **1** 2 »

OXLOD: record of specimen