

# Semantic Web Modelling: Challenges and Opportunities in Small and Large Museum Collections

**Dr Andreas Vlachidis – UCL Department of Information Studies**

[a.vlachidis@ucl.ac.uk](mailto:a.vlachidis@ucl.ac.uk)

# CrossCult – Identity Card

*Empowering reuse of digital cultural heritage in context-aware crosscuts of European history*

**Call:** EU H2020

**Topic:** REFLECTIVE-6-2015

**Type of Action:** Innovation Action

**Timeline:** 36 months, March 2016 – February 2019

**Budget (EU Contribution):** €3,503,358.25

11 core, 14 associate partners



# Archaeological Museum of Tripolis (Greece)

Personalising Associations

Pilot 3: Non-typical itineraries through museum items



# Pilot Case



Facilitate interconnections  
(Objects, Viewpoints,  
Venues)

Interactivity  
Recollection  
Reflection



Non-typical viewings

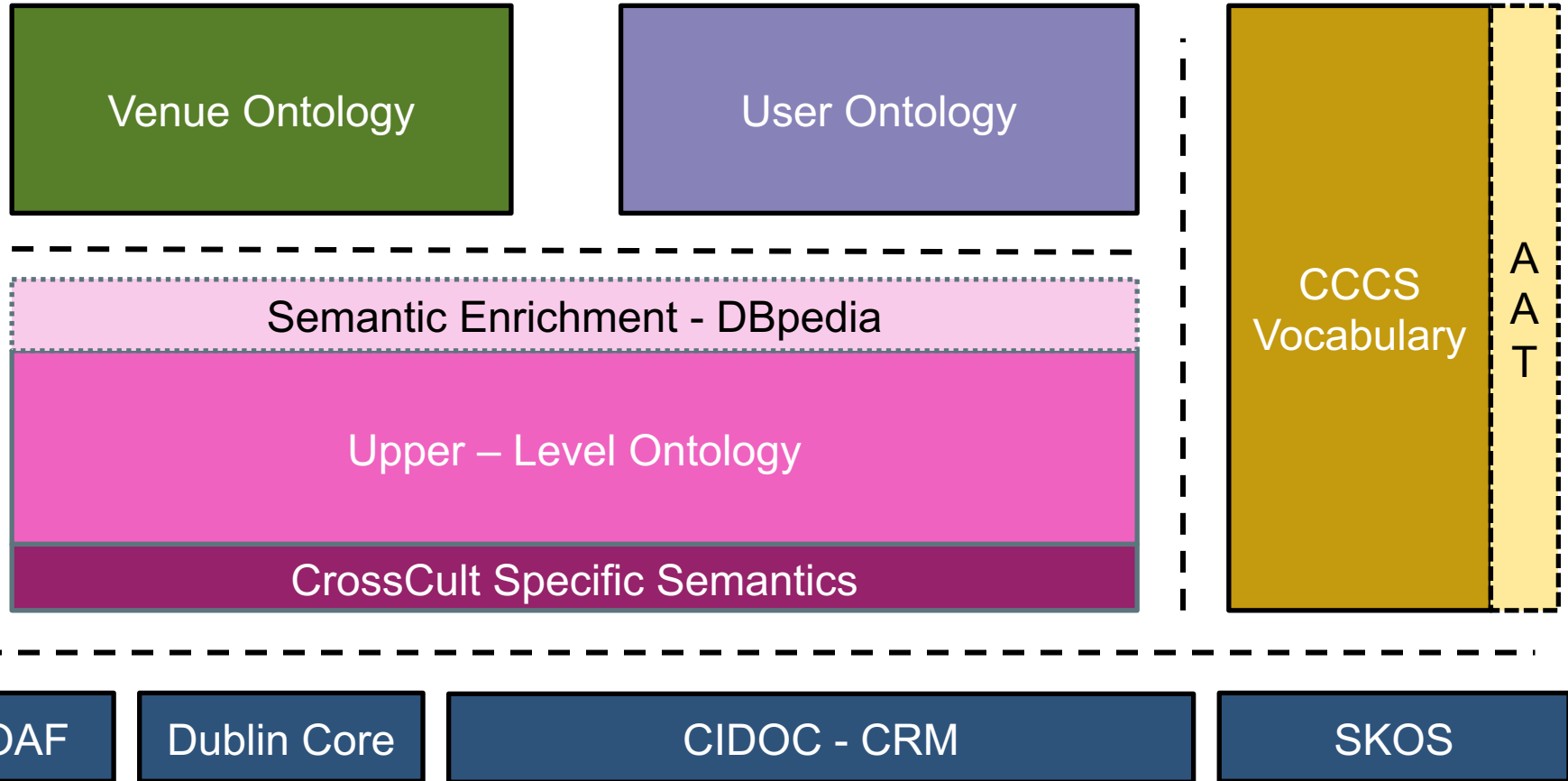
Crosscutting and transversal  
Beyond the typical  
presentation  
Deeper levels of reflection,  
• social aspects / power  
structures.



# The Crosscult Knowledge Base

- A semantic knowledge base capable of interrelating a wide set of (existing and future) disparate digital cultural heritage resources
  - Repository for storage
  - Management and retrieval of semantic information.
  - Common Layer for Cultural Heritage Concepts and their relationships,
  - Facilitate interoperability and linking with Linked Data resources
  - Building on standard Semantic Web technologies

# The CrossCult Knowledge Base



- **Ontology size:** 11,484 axioms, 2186 individuals
- **CCKB size:** 1,201,533 axioms, 282,010 individuals

# From Unstructured Data to Structured Data

- Example Item Description

2279: Marble pediment tombstone with a representation of a family (enface). The female figure bears a chiton and a cloak. The male figure and the boy bear a short chiton. On the architrave there is the inscription ANTIOXIC ΦΟΡΤΟΥΝΑΤΟΥ ΘΥΓΑΤΗΡ ΚΑΛΛΙΣΤΗ.

Found in Herod Atticus villa in Loukou, Kynouria. Roman era work (middle Antonine era, 161 A.D - 180 A.D.).

Dimensions: Height 1.60m, Width 0.82m. Location: Room 15, 1st floor.



# Mapping to CIDOC-CRM

- Item “2279” → E22.Human-made Object
  - P1\_is\_identified\_by → E42.Identifier “2979”
  - P2\_has\_type → E55.Type “Tombstone”
  - P3\_has\_note “the original textual description”
  - P43\_has\_dimension → E54.Dimension “height 1.60m” and “width 0.82m”
  - P45\_consists\_of → E55.Type “Marble”
  - P50\_has\_current\_keeper → E40.Legal Body “The Archaeological Museum of Tripoli”
  - P55.has\_current\_location → E53 Place “Room 15”
  - P102.has\_title → “Antiohis pediment tombstone”
  - P138i.has\_represenation → E38.Image “antiohis.jpg”.
  - *reflects* → *CCRT* “*Worship*”

# The Semantics of Reflective Topic

## Reflective Topic

- Create a network of points of view,
- Aid reflection and prospective interpretation over a topic,
- Enable interconnection between physical or conceptual things

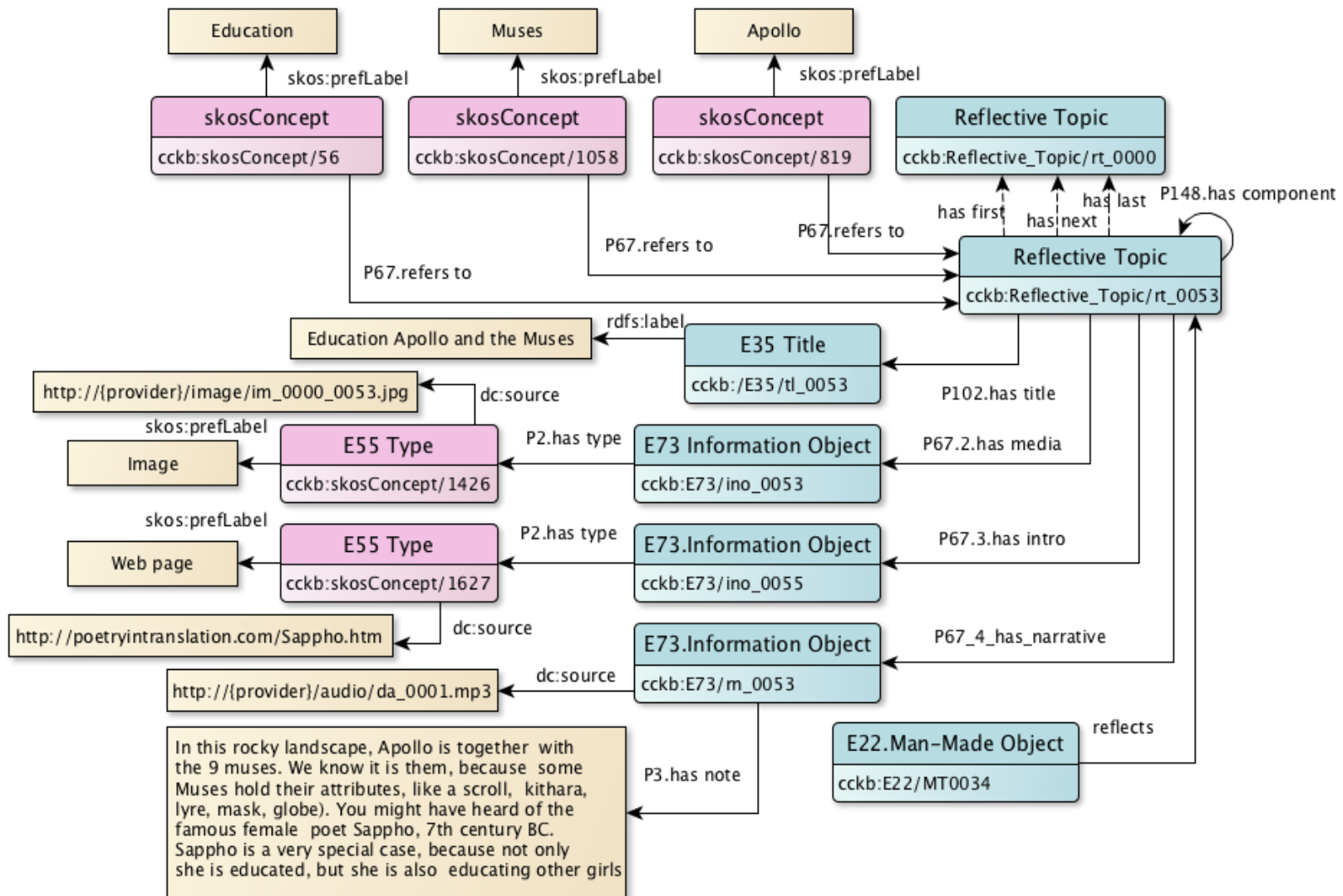
## is reflected by

- any **physical object** or immaterial thing,
- accommodate a true paradigm of reflection a **proposition**

## experienced as

- a composition or sequence of narrower Reflective Topics

# The Semantics of Reflective Topic

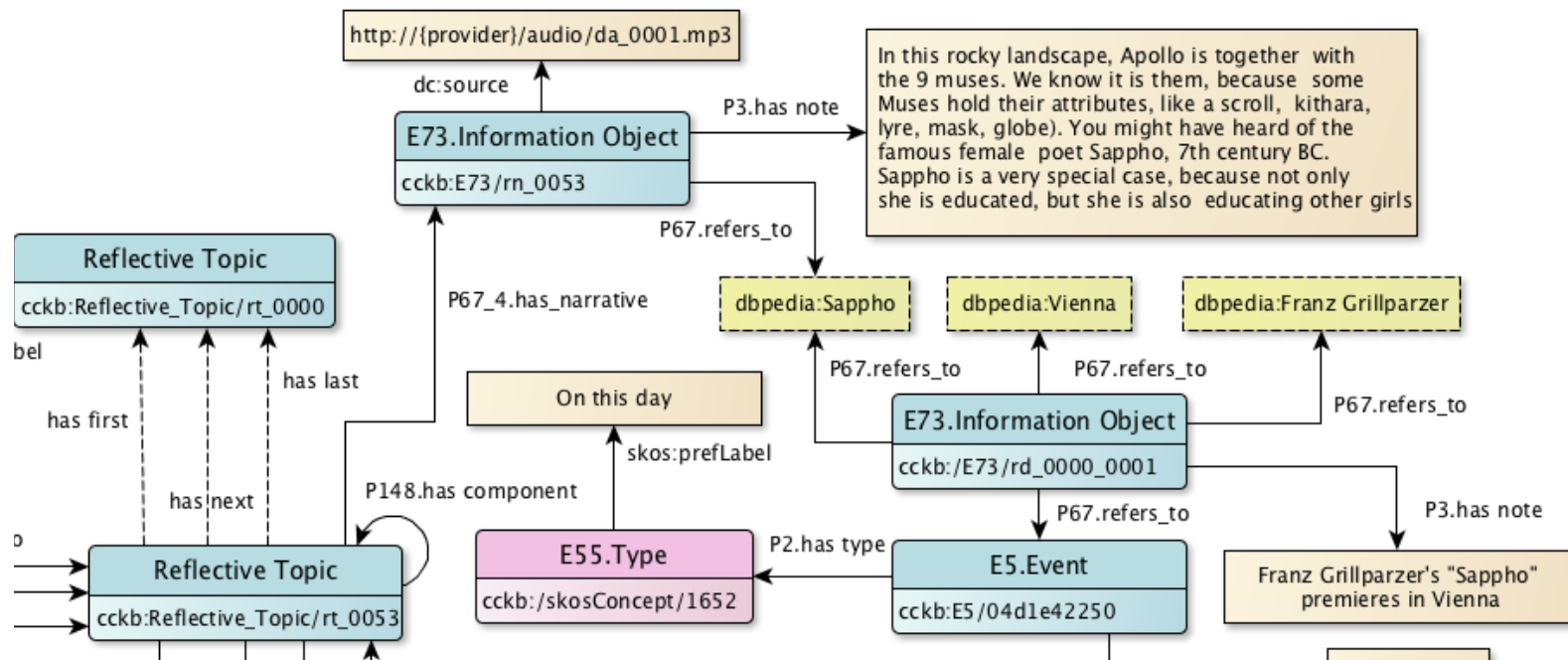




# Semantic Enrichment

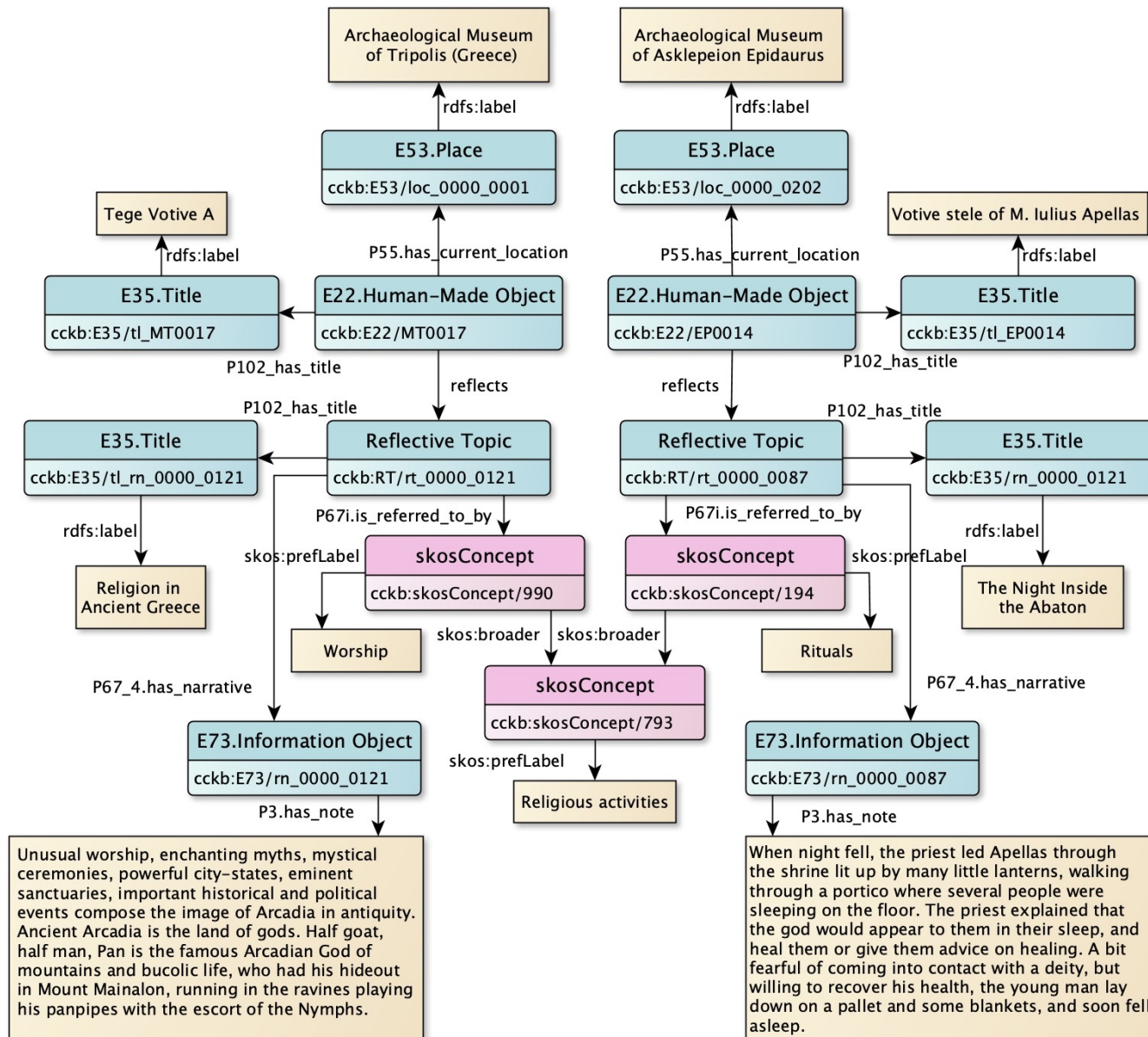
- **Trigger reflection through enrichment**
- Enriched a selected set of ontology individuals with links to standard and well-known Semantic Web resources such as
  - DBpedia
  - Getty AAT
- using appropriate properties from
  - OWL ([owl:sameAs](#))
  - SKOS ([skos:closeMatch](#))
- using Named Entity Linking
  - manually (by experts)
  - automatically (with DBpedia Spotlight)

# Example



## Search Scenario

- Related narratives of artefacts across museums
  - find more items that reflect topics relating to worship, which might be located else where
- Results
  - semantically expanding on the topic through its broader concept “Religious Activities”, the query retrieves the item EP0014, which is located in the Archaeological Museum of Asklepeion in Epidaurus



# Small Museums Challenges and Hindrances

---

Non-existent digital presence

---

Slow technology acceptance

---

Bureaucracy associated with decision-making

---

Geographical location

---

Low numbers of physical visitors

---

Lack of budget

---

Limited human resources

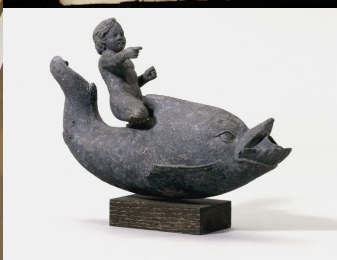


# The Sloane LAB

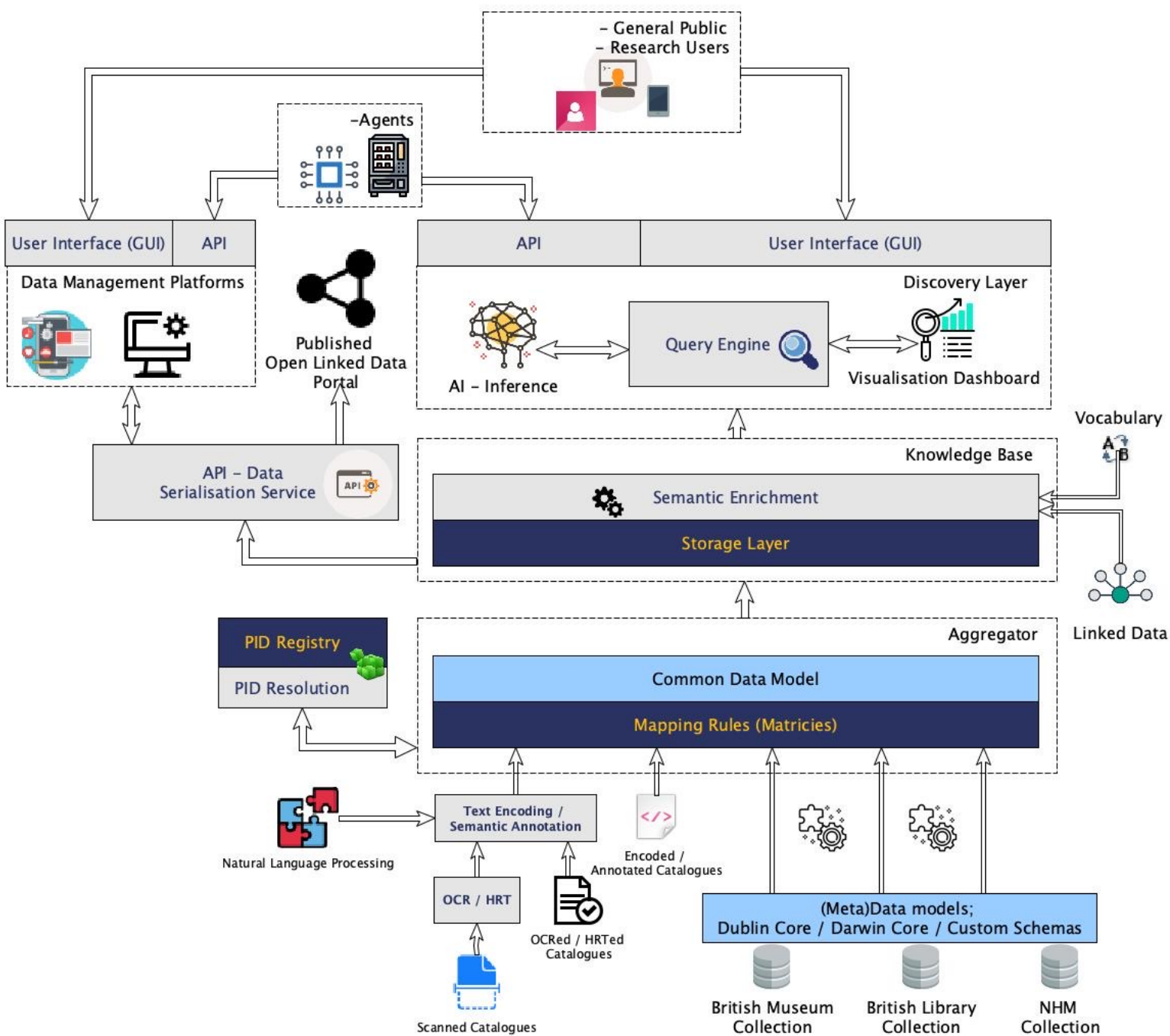


Sloane an **exemplary case study** to unite surviving objects (current digital collections) and records across the NHM, BM and BL from the original catalogues.

Reunify and Extend digital collection  
Searchable and amenable to further analysis and interpretation







# Sloane Lab Data Atlas

## Charting all data available for Sir Hans Sloane and his collections

### Colour index

Colour-coding is only applicable to cells relating to data sources

- Digital - Available digitally and available publicly
- Digital - Available digitally but within the institution
- Digital - In the process of becoming available
- Tangible artefact
- Permanently destroyed

### Enlightenment Architectures

- Manuscript Catalogues - Fully Digitised (TEI)
- <sup>1</sup> Digital edition of MS 3972C Vol VI, Sloane's catalogue of books and printed ephemera
- <sup>2</sup> Digital edition of Sloane MS 3972B of folios listing manuscript material removed from 3972C
- <sup>3</sup> Digital edition of Sir Hans Sloane's Catalogues of Miscellanies, Antiquities, Seals, Pictures
- <sup>4</sup> Digital edition of Sir Hans Sloane's Catalogue of Fossils including Corals, Serpents, Eggs, etc.
- <sup>5</sup> Digital edition of Sir Hans Sloane's Catalogue of Fossils including Fishes, Birds, Eggs, etc.

### Sloane Letters

- Digitised Letters written to and from Sloane
- <sup>1</sup> 5,320 Searchable Letters

### Footnotes for multiple instances

- 1 Also available in Enlightenment Architectures as "Digital edition of Sir Hans Sloane's Catalogue of Fossils including Fishes, Birds, Eggs, etc."
- 2 Also available in Enlightenment Architectures as "Digital edition of Sir Hans Sloane's Catalogue of Fossils including Fishes, Birds, Eggs, etc."
- 3 Available in Sloane Letters as "5,320 Searchable Letters"
- 4 Also available in Enlightenment Architectures as "Digital edition of Sloane MS 3972B of folios listing manuscript material removed from 3972C"
- 5 Also available in Enlightenment Architectures as "Digital edition of Sir Hans Sloane's Catalogue of Fossils including Corals, Serpents, Eggs, etc."
- 6 Also available in Enlightenment Architectures as "Digital edition of Sir Hans Sloane's Catalogue of Fossils including Fishes, Birds, Eggs, etc."
- 7 Also available in British Library as "Sloane MS 3972 C VI"
- 8 Also available in British Library as "Sloane MS 3972 B"
- 9 Also available in British Museum as "Miscellanea vol"
- 10 Also available in Natural History Museum as Fossils vol 1
- 11 Also available in Natural History Museum as Fossils vol 5
- 12 Available in British Library as "Sloane Correspondence"

Insects vol 1	→	Insects vol 1	→	Not transcribed
Insects vol 2	→	Insects vol 2	→	Not transcribed
Minerals vol 1	→	Minerals vol 1	→	Not transcribed
Minerals vol 2	→	Minerals vol 2	→	Not transcribed
Minerals vol 3a	→	Minerals vol 3a	→	Not transcribed
Minerals vol 3b	→	Minerals vol 3b	→	Not transcribed
Minerals vol 4	→	Minerals vol 4	→	Not transcribed
Echinoderms	→	Echinoderms	→	Not transcribed
Supplement	→	Supplement	→	Not transcribed
Catalogus marmorum	→	Catalogus marmorum	→	Not transcribed
Collection of Vegetables 1-3000	→	Collection of Vegetables 1-3000	→	Collection of Vegetables 1-3000
Collection of Vegetables 3001-6000	→	Collection of Vegetables 3001-6000	→	Collection of Vegetables 3001-6000
Collection of Vegetables 6001-12523	→	Collection of Vegetables 6001-12523	→	Collection of Vegetables 6001-12523
Collection of Vegetables Index [part 1]	→	Collection of Vegetables Index [part 1]	→	Not transcribed
John Ray, Historia Plantarum Vol I (HS copy)	→	John Ray, Historia Plantarum Vol I	→	John Ray, Historia Plantarum Vol I
John Ray, Historia Plantarum Vol II (HS copy)	→	John Ray, Historia Plantarum Vol II	→	John Ray, Historia Plantarum Vol II
John Ray, Historia Plantarum Vol III (HS copy)	→	John Ray, Historia Plantarum Vol III	→	John Ray, Historia Plantarum Vol III

### Collection Guides

<b>Collection Guides - Print</b>	<b>Collection Guides - Scanned</b>
J.E. Dandy, The Sloane herbarium : an annotated list of the Horti sicci composing it	J.E. Dandy, The Sloane herbarium : an annotated list of the Horti sicci composing it
A catalogue and historical account of the Sloane Shell Collection by Guy L. Wilkins, Bulletin of the British Museum (Natural History) Historical Series Vol I, 1953-1959	A catalogue and historical account of the Sloane Shell Collection by Guy L. Wilkins, Bulletin of the British Museum (Natural History) Historical Series Vol I, 1953-1959
V. Pickering, Putting nature in a box: Hans Sloane's Vegetable Substances collection, unpublished thesis 2017	V. Pickering, Putting nature in a box: Hans Sloane's Vegetable Substances collection, unpublished thesis 2017

### Objects

<b>Bound Volumes</b>	<b>Bound Volumes - Catalogued</b>
Sloane Herbarium bound volumes	Sloane Herbarium bound volumes
<b>Folios</b>	<b>Folios - Scanned? and/or Catalogued?</b>
Sloane volume: folio and specimen level cat	Sloane volume: folio and specimen level catalog
<b>Algae, Fungi and Plants</b>	<b>Algae, Fungi and Plants - DOI-registered collect</b>
Sloane herbarium: Sloane's Jamaican material	Sloane herbarium: Sloane's Jamaican material
Sloane herbarium: Mark Catesby material	Sloane herbarium: Mark Catesby material
Sloane herbarium: Engelbert Kaempfer material	Sloane herbarium: Engelbert Kaempfer material
<b>Vegetable Substances</b>	<b>Vegetable Substances</b>
<b>Insects</b>	<b>Invertebrates - DOI-registered collections</b>
FlukeNet bound volume	Molluscs
Petiver insect collection	Other invertebrates
Sloane specimens (in boxes on shelves)	
<b>Invertebrates</b>	<b>Palaeontology - DOI-registered collections</b>
Molluscs	Fossils
Other invertebrates	
<b>Palaeontology</b>	<b>Mineralogy - DOI-registered collections</b>
Fossils	Mineral specimens
<b>Mineralogy</b>	<b>Vertebrates - DOI-registered collections</b>
Mineral specimens	Fish
<b>Vertebrates</b>	Herps
Fish	Mammals
Herps	Birds
Mammals	
Birds	

<b>Sloane MS 3972 C.VI</b>	<b>not digitised?</b>	<b>not digitised?</b>
Sloane MS 3972 C.VI	→	Sloane MS 3972 C.VI <sup>1</sup>
Sloane MS 3972 C.VII	→	not digitised?
Sloane MS 3972 C.VIII	→	not digitised?
Sloane MS 3972 D	→	not digitised?
Lindenius renovates	→	not digitised?
<b>Objects</b>		
<b>Collected Manuscripts - Print</b>	<b>Collected Manuscripts - Catalogued</b>	<b>Books and Manuscripts - Scanned</b>
SL MSS 1-4100 (Manuscripts collected by Sloane)	→	Not scanned
<b>Books and Manuscripts collected by Sloane</b>	<b>Books and Manuscripts - Catalogued</b>	<b>Sloane Renaissance Manuscripts</b>
Printed Books	→	→
Sloane Renaissance Manuscripts	→	→
Other Books & Manuscripts collected by HS	→	→
<b>Other Catalogues</b>	<b>Other Catalogues - Digitised</b>	<b>Some letters digitised<sup>2</sup></b>
Printed Books Catalogue	→	→
<b>Sloane Correspondence - Print</b>	<b>Printed Books Catalogue</b>	<b>Some letters digitised<sup>4</sup></b>
Sloane Letters MSS 4036-4069	→	→
Sloane Letters MSS 4075-4078	→	→

### Royal Society

<b>Draft Minutes - Manuscripts</b>	<b>Draft Minutes - Fully Digitised</b>	
Meetings of the Royal Society, 1686-1691	→	Meetings of the Royal Society, 1686-1691
Meetings of the Royal Society, 1693-1698	→	Meetings of the Royal Society, 1693-1698
Meetings of the Royal Society, 1699-1702	→	Meetings of the Royal Society, 1699-1702
Meetings of the Royal Society, 1702-1707	→	Meetings of the Royal Society, 1702-1707
Meetings of the Royal Society, 1707-1711	→	Meetings of the Royal Society, 1707-1711

Credit: Dr Foteini Valeonti

### Sloane's Manuscript Catalogues Today

<b>Natural History Museum</b>		
<b>Sloane Manuscript Catalogues</b>	<b>Manuscript Catalogues - Digitised</b>	<b>Manuscript Catalogues - Transcriptions</b>
Fossils vol 1	→	Fossils vol 1 <sup>1</sup>
Fossils vol 2	→	Fossils vol 2
Fossils vol 3	→	Fossils vol 3
Fossils vol 4	→	Fossils vol 4
Fossils vol 5	→	Fossils vol 5 <sup>2</sup>
Fossils vol 6	→	Fossils vol 6

## Sloane Manuscript Catalogues

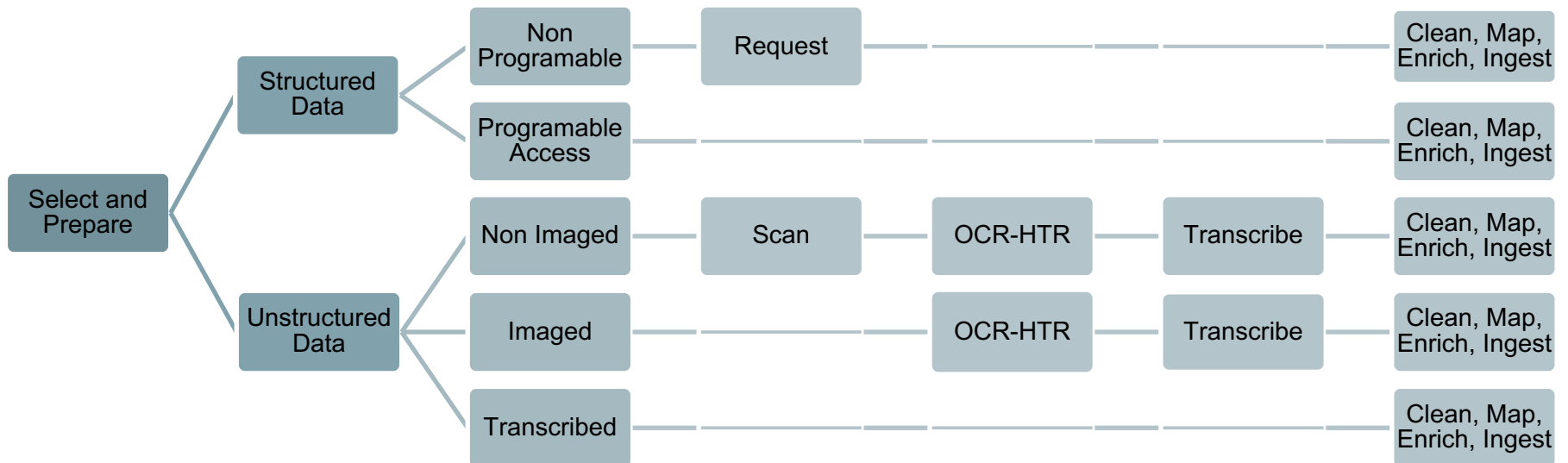
Sloane Manuscript Catalogues		Manuscript Catalogues - Digitised		Manuscript Catalogues - Transcriptions
Fossils vol 1	→	Fossils vol 1 <sup>1</sup>	→	Fossils vol 1 <sup>1</sup>
Fossils vol 2	→	Fossils vol 2		<i>Not transcribed</i>
Fossils vol 3	→	Fossils vol 3		<i>Not transcribed</i>
Fossils vol 4	→	Fossils vol 4		<i>Not transcribed</i>
Fossils vol 5	→	Fossils vol 5 <sup>2</sup>	→	Fossils vol 5 <sup>2</sup>
Fossils vol 6	→	Fossils vol 6		<i>Not transcribed</i>
Insects vol 1	→	Insects vol 1		<i>Not transcribed</i>
Insects vol 2	→	Insects vol 2		<i>Not transcribed</i>
Minerals vol 1	→	Minerals vol 1		<i>Not transcribed</i>
Minerals vol 2	→	Minerals vol 2		<i>Not transcribed</i>
Minerals vol 3a	→	Minerals vol 3a		<i>Not transcribed</i>
Minerals vol 3b	→	Minerals vol 3b		<i>Not transcribed</i>
Minerals vol 4	→	Minerals vol 4		<i>Not transcribed</i>
Echinoderms	→	Echinoderms		<i>Not transcribed</i>
Supplement	→	Supplement		<i>Not transcribed</i>
Catalogus marmorum	→	Catalogus marmorum		<i>Not transcribed</i>
Collection of Vegetables 1-3000	→	Collection of Vegetables 1-3000	→	Collection of Vegetables 1-3000
Collection of Vegetables 3001-6000	→	Collection of Vegetables 3001-6000	→	Collection of Vegetables 3001-6000
Collection of Vegetables 6001-12523	→	Collection of Vegetables 6001-12523	→	Collection of Vegetables 6001-12523
Collection of Vegetables Index [part 1]	→	Collection of Vegetables Index [part 1]		<i>Not transcribed</i>
John Ray, Historia Plantarum Vol I (HS copy)	→	John Ray, Historia Plantarum Vol I	→	John Ray, Historia Plantarum Vol I
John Ray, Historia Plantarum Vol II (HS copy)	→	John Ray, Histori		ol II
John Ray, Historia Plantarum Vol III (HS copy)	→	John Ray, Histori		ol III

### Colour index

*Colour-coding is only applicable to cells relating to data sources*

<span style="background-color: #ADD8E6; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>	Digital - Available digitally and available publicly
<span style="background-color: #FF00FF; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>	Digital - Available digitally but within the institution
<span style="background-color: #D3D3D3; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>	Digital - In the process of becoming available
<span style="background-color: #FFD700; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>	Tangible artefact
<span style="background-color: #FF0000; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>	Permanently destroyed

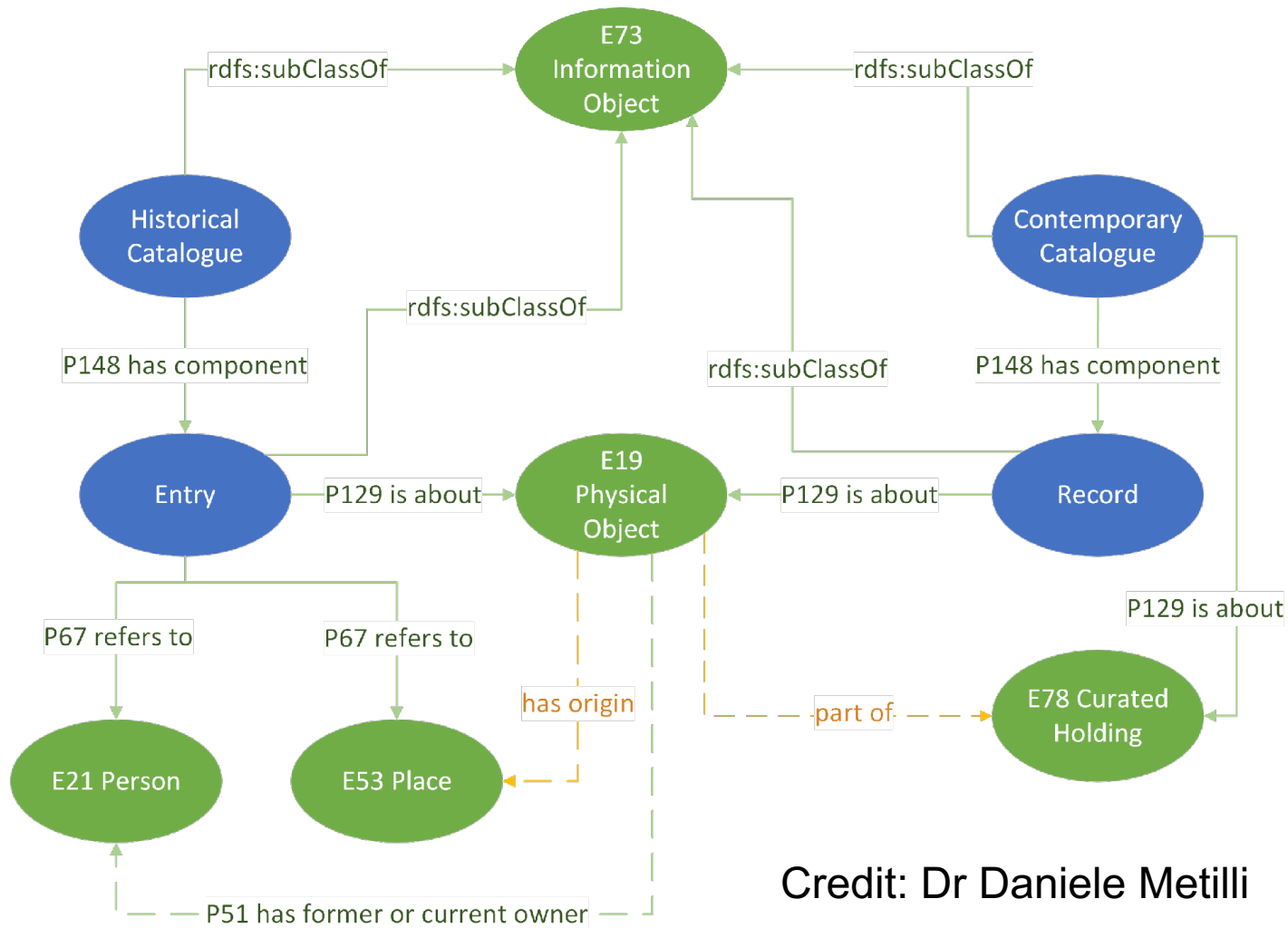
# Towards Unification



## Data Model

- The **data model** allows us to represent the **knowledge** collected in the project
- Based on **Semantic Web** standards and built on top of the **CIDOC CRM** reference model
- Centred on the **record** instead of the **object**, to support multiple perspectives over the same object

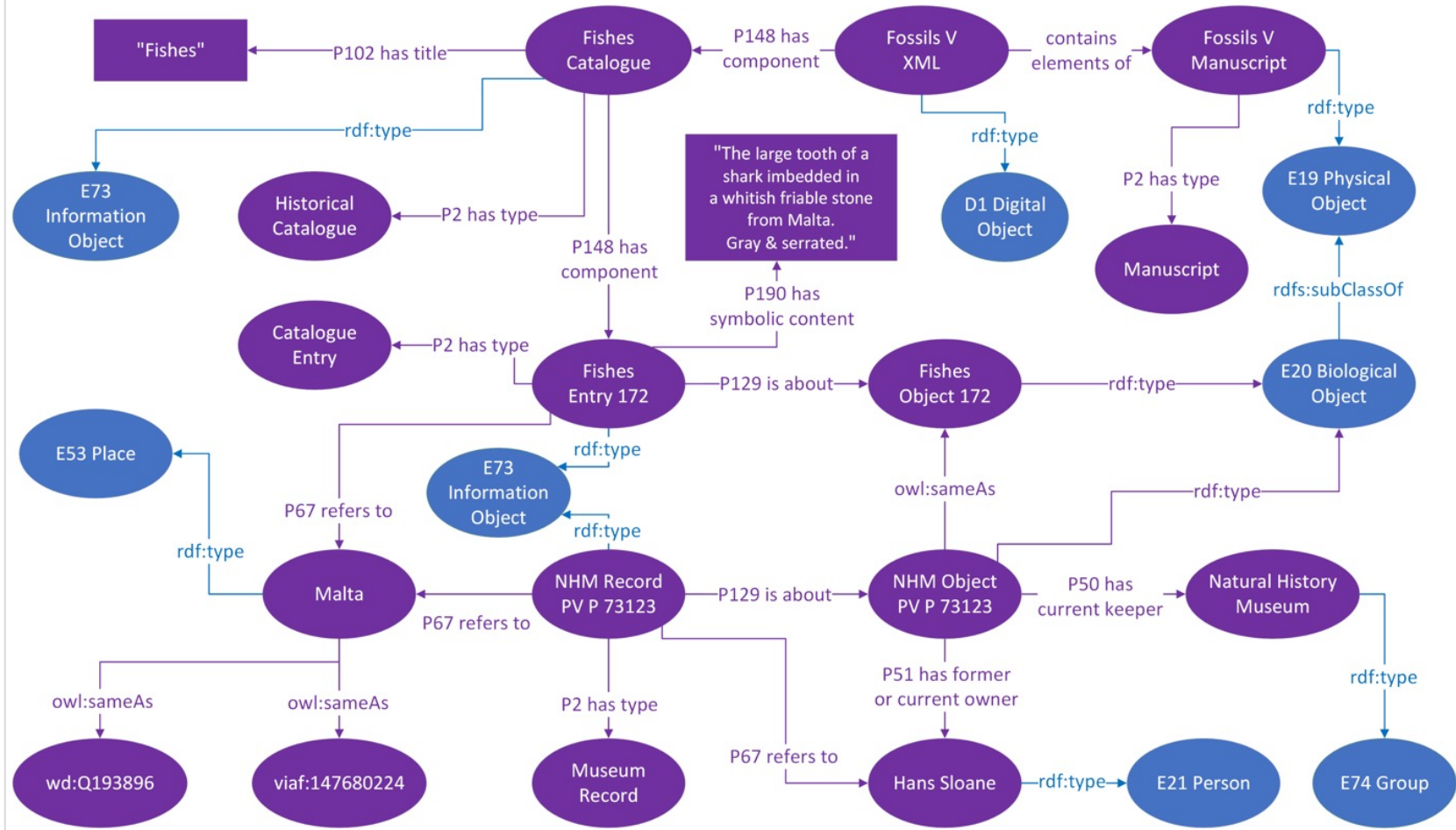
# Data Mode



Credit: Dr Daniele Metilli

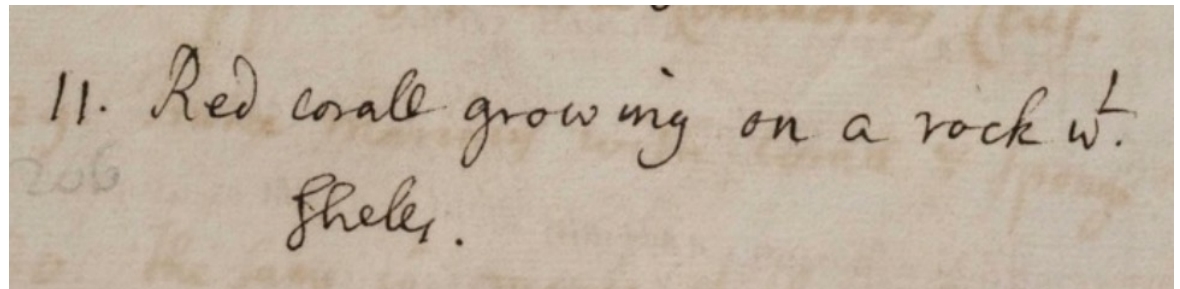
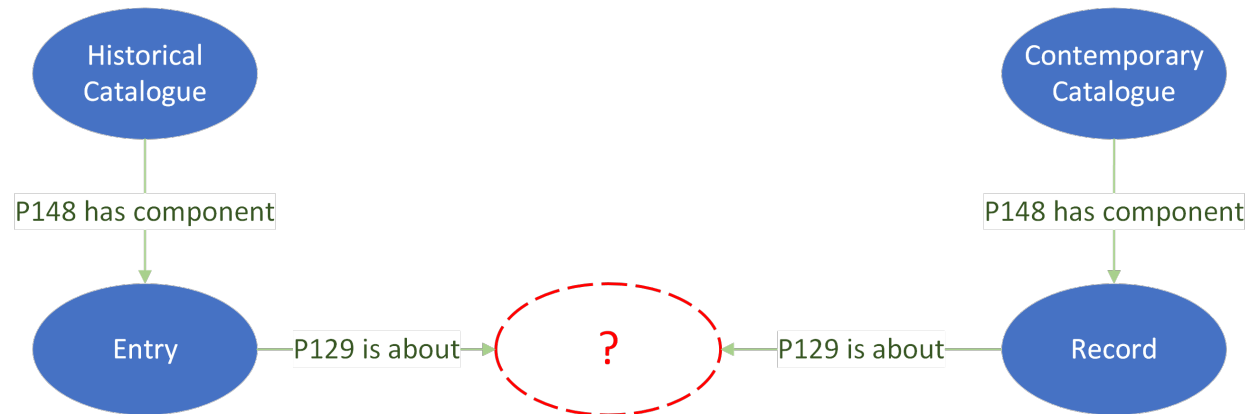


**Fishes Entry 172 — “The large tooth of a shark”**  
object with a Natural History Museum record



# The Object Problem

- Often we cannot rely on objects to connect historical and contemporary catalogues
- Moreover, it is challenging to “identify the boundaries between ‘object names’ and their qualifying descriptions”



Credit: Dr Daniele Metilli

# Challenges

- **Aggregating** records and **integrating** them into a single knowledge base
- Handling **uncertainty** in records
- Modelling **inconsistencies** among the records
- Identifying **data absences** and attempting to remedy them

-1100. A large reed pipe from the great Mogols country from M<sup>r</sup>. Dupuys.  
b ✓ 245.

## smoking-pipe

### Object Type

smoking-pipe

### Museum number

As,SLMisc.1100

### Description

Bamboo smoking pipe with hole drilled in centre, through only one side. Bands of incised geometric patterns mainly to each end, but also a floral design to one side of the drilled hole.

### Production date

18thC

### Production place

Made in: India (pre-1947) (?)

Asia: South Asia: India (pre-1947)



# Identifiers

Multiple identifiers, Not persistent Not always unique

## abacus

### Object Type

abacus

### Museum number

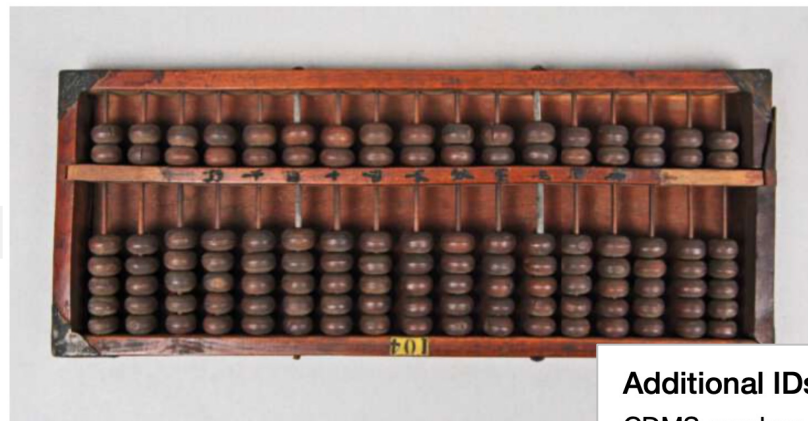
As,SLMisc.594

### Description

Abacus (with 17 rows of beads, divided by cross-bar carrying Chinese characters) made of wood, metal.

### Findspot

Found/Acquired: China



### Additional IDs

CDMS number: CDMS number:

As1753D10.594 (old CDMS no.) (old CDMS no.)

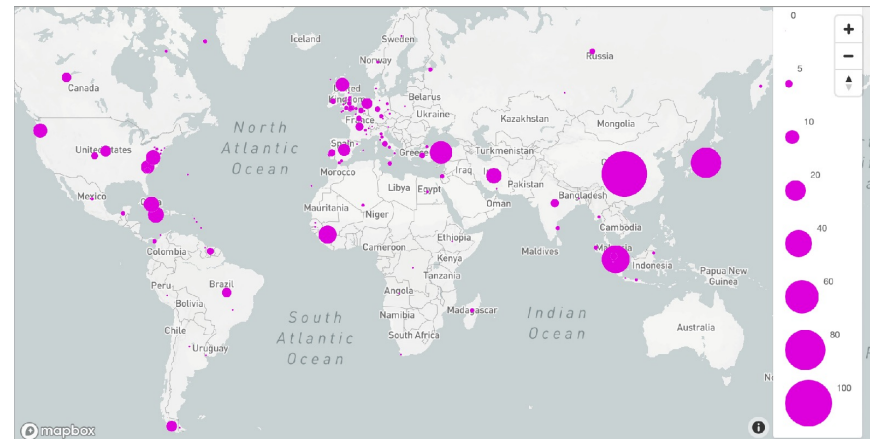
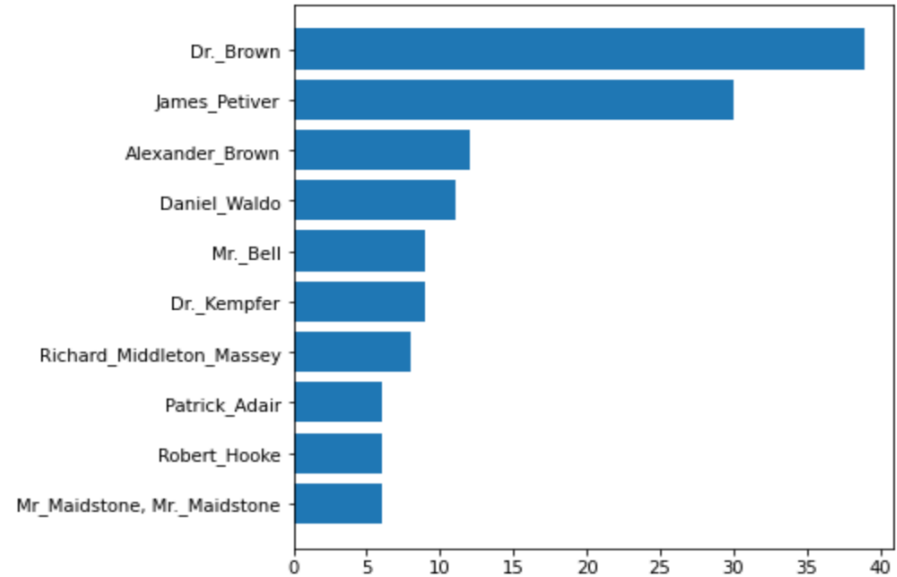
Miscellaneous number: Miscellaneous number: 104 (paper label attached) (paper label attached)

Sloane Miscellanea Catalogue number: Sloane Miscellanea Catalogue number: 594 (Miscellanies) (Miscellanies)

Credit: Dr Daniele Metilli

# Opportunities – People and Places

- If we successfully identify the people and places, we can map their distribution in the catalogues
- We can also try to reconstruct Sloane's social network
- Then we can use computational methods to analyse the data







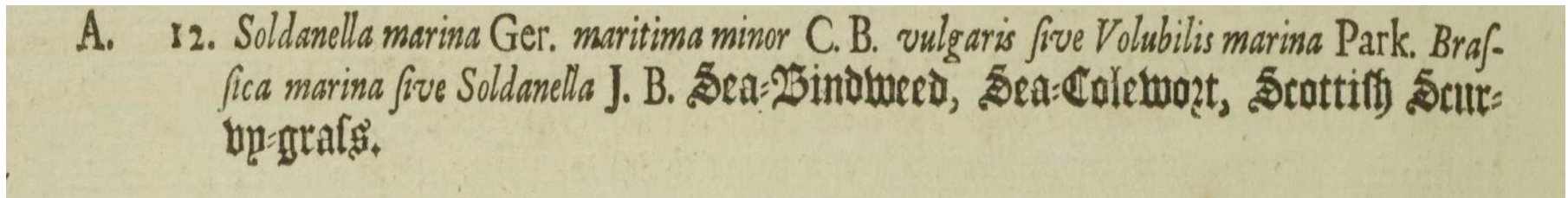




# Challenges

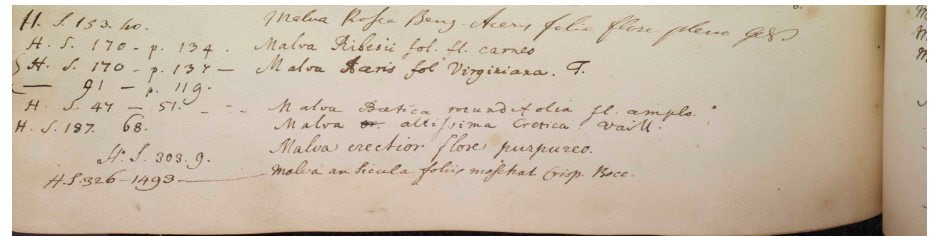
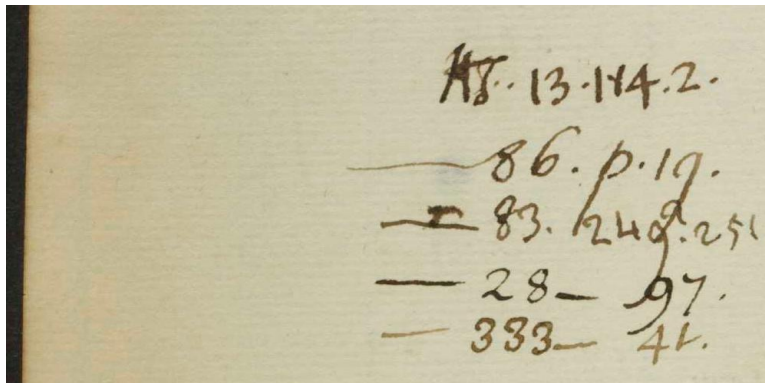
## Mix Latin and Gothic scripts:

Latin Script Accuracy > 90%, Gothic Script ≈ 75% up to 85% when combined with Transcribus

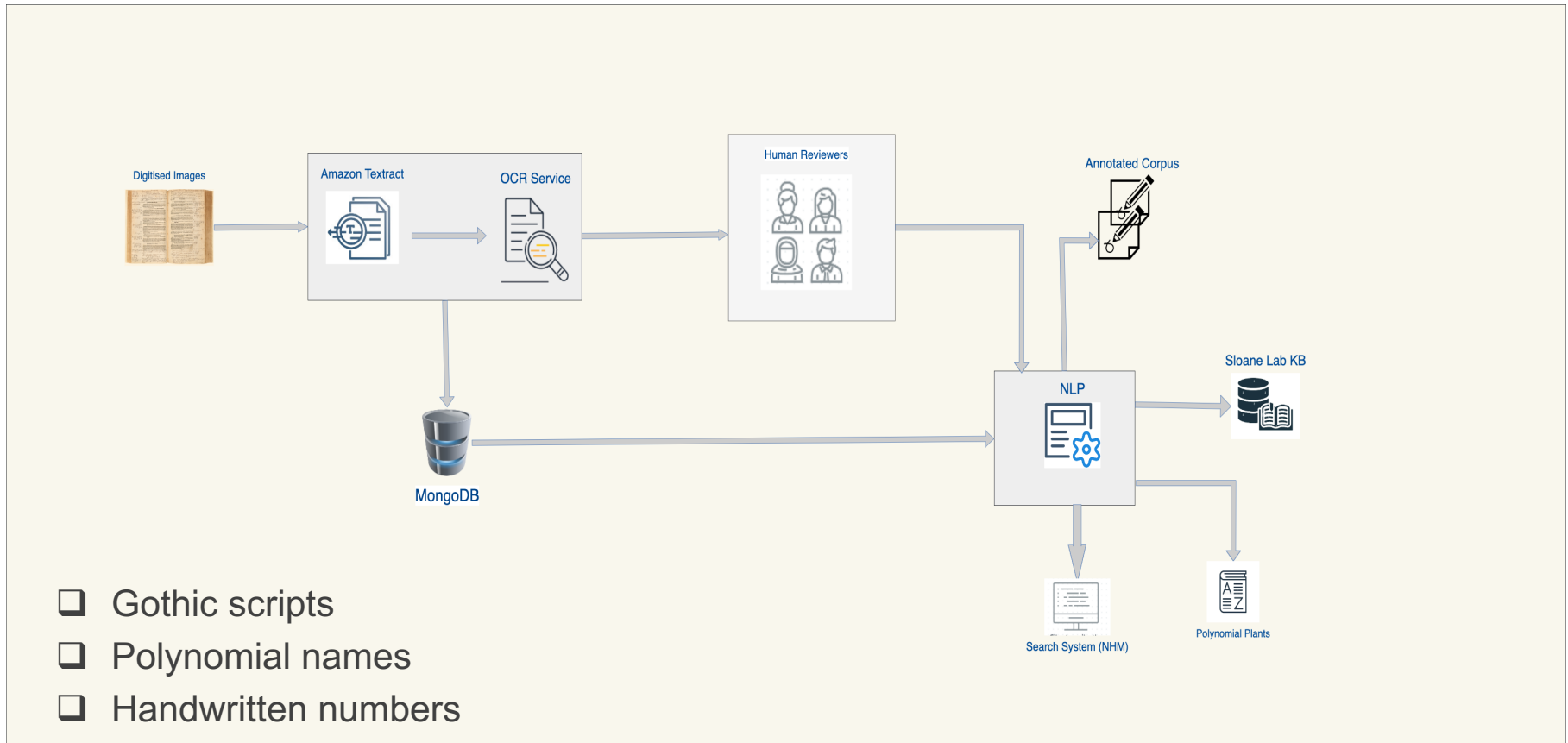


## Hand-written numbers:

Accuracy ≈ 75 % - Required 100%



# Workflow



- Gothic scripts
- Polynomial names
- Handwritten numbers

# Zooniverse – Human Transcription

Trained Human reviewers will transcribe hand-written numbers on left/right margins

Produce accurate references to Herbarium

The screenshot shows the Zooniverse transcription interface. On the left, a manuscript page is displayed with a red box highlighting a handwritten number '139' in the right margin. A green arrow points from this number to a transcription dialog box. The dialog box contains the text 'Transcribe the handwriting linked to this segment' and a text input field with 'H.S. 139'. Below the input field are buttons for 'deletion', 'unclear', and 'insertion', and a large yellow 'OK' button at the bottom. On the right side of the interface, there are two tabs: 'TASK' and 'TUTORIAL'. Below the tabs, the text 'Mark and transcribe' is displayed. There are two rows of buttons: a teal button with a green checkmark and the text 'Mark and transcribe the correct 1 pairs drawn', and a grey button with a pink square and the text 'Draw and transcribe the incorrect 0 pairs drawn'. Below these buttons is a yellow 'Next →' button with a gear icon to its right. At the bottom of the interface, there is a navigation bar with various icons including a play button, a blue circle, a right arrow, and a series of circles representing a progress bar.

# Conclusions

## The Promise

- Foster Connection and Contextualisation
- Define Concepts and Facilitate Multifaceted Retrieval
- Harmonised View, Reasoning, Recommendation
- Understand Perceptions, Recollections, Viewpoints

## The Complexity

- Data integration and knowledge reuse
- Level of Digitization
- Institutional Legacies and Curational Practice
- Policies and Copyrights

## The Absence

- *When we attend to the who, what and how of data it is likely that we will also need to attend to the who-not, what-not and how-not, or absences of that data*  
(Prof Julianne Nyhan)

# Semantic Web Modelling: Challenges and Opportunities in Small and Large Museum Collections

**Dr Andreas Vlachidis – UCL Department of Information Studies**

[a.vlachidis@ucl.ac.uk](mailto:a.vlachidis@ucl.ac.uk)





# Evaluation

- Conducted in two rounds between March and July 2018
  - 4 Humanities Experts
  - 81 other users
  - non-expert users from among students of diverse degrees in the University of Vigo in Spain and the Arab Academy in Egypt.
- The Evaluation Task
  - tag precomputed associations for three (synthetic) personae
  - experts received a total of 136 associations
  - Non- experts (potential visitors) received 323 associations



Sample Association Description used in Evaluation

<p><b>Association</b></p> <p>Due to your interest in <b>Women and Mythology</b>, let us tell you about an event that happened exactly 104 years ago:</p>		<p><b>10 March – 1914:</b> Suffragettes in London damage Rokeby's painting Venus of Velasquez</p>
<p><b>Local exhibit</b></p> <p>This date, linked to <b>Venus</b>, takes us to this item from our collection:</p>		<p><i>Marble tombstone with a representation of a woman and a young athlete (mother and son). The female figure bears a reaching to the feet sleeved chiton and a cloak covering her head and the lower part of her body. The young man is represented nude, in an enface position and holding a strigil in his right hand. The heads of the figures are broken. Originated from Mantinea and dated in the end of the 5th century B.C.</i></p> <p><i>Height: 0.95m Width: 0.47m Location: Room 15, 1st floor.</i></p>
<p><b>Reflective narrative</b></p> <p>On 10 March 1914, the suffragette Mary Richardson walked into the National Gallery and attacked Velázquez's canvas with a meat cleaver and left 7 slashes on the painting, the most notable between the figure's shoulders. The incident has come to symbolize a particular perception of feminist attitudes towards the female nude (and, in a sense, a stereotypical image of feminism as a whole).</p> <p>Do you think the picture above does portray women in a disadvantageous position? Or is it liberating them? What about our marble statue, that depicts women in the opposite way, and that was also vandalized in the past?</p>		