Bridging the collaboration gap: Europeana, DiSSCo & DARIAH through FAIR integration of cultural and natural heritage data for interdisciplinary research

Sally Chambers, Ghent Centre for Digital Humanities & KBR, Royal Library of Belgium. DARIAH-EU Senior Management team, DiSSCo Scientific Advisory Board member.

Sharif Islam: Data Architect, DiSSCo / Naturalis Biodiversity Center, Leiden, NL

Tina Loo: Project Officer, Biodiversity Data, Naturalis Biodiversity Center, Leiden, NL.

Andreas Weber: Assistant Professor, Department of Technology, Policy and Society (TPS). University of Twente, The Netherlands.

Erzsébet Tóth-Czifra: Open Science Officer, DARIAH-EU
Three things we hope our audience will take away

- **Increased understanding** of the FAIR data needs of Europeana, DiSSCo & DARIAH and ways to align the research practices of these communities
- Explore specific examples of ways to create data linkages and increase cross-domain data interoperability
- Raised awareness of the collaboration opportunities between Europeana, DiSSCo and DARIAH regarding data-intensive research using cultural & natural heritage resources.
Bridging the collaboration gap: Europeana, DiSSCo & DARIAH through FAIR integration of cultural and natural heritage data for interdisciplinary research

Menti Meter: https://www.menti.com/
Code: 7335 2893
• DiSSCo is a service-oriented infrastructure:
  ▪ Digitization, Capacity Building
  ▪ FAIR data
  ▪ Physical and digital access to Natural Science Collections
  ▪ Support & Training services

Goal: to digitally unify today’s fragmented landscape of European NSCs into a single knowledge base under common curation, access policies and practices

https://dissco.eu

Image Credit: DiSSCo Coordination and Support Office
A Digital Specimen is an online digital object, identified, organized and shaped to record provenance and to include all data created during the full life cycle of a physical specimen...

..and beyond. A true Digital Twin. -> it is persistent, updatable and FAIR by design.
A Digital Specimen is an online digital object, identified, organized and shaped to record provenance and to include all data created during the full life cycle of a physical specimen...

..and beyond.
A true Digital Twin.
-> it is persistent, updatable and FAIR by design

This makes it very suitable for extending with annotations and linkages and using it in analysis: a Digital Extended Specimen.

Image Credit: Wouter Addink, Naturalis Biodiversity Center
Example (1)

Coll. by Mr. & Mrs. Monteiro Recd. 6/73

Source: (Royal Botanic Gardens, Kew - Herbarium Specimens)
https://www.gbif.org/occurrence/912297514
https://www.europeana.eu/en/item/11614/KXHERBARIUMXK000347209
<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalogue number</td>
<td>K000347209</td>
</tr>
<tr>
<td>Occurrence ID</td>
<td><a href="http://specimens.barium/K00034729">http://specimens.barium/K00034729</a></td>
</tr>
<tr>
<td>Occurrence remarks</td>
<td>Coll. by Mr. &amp; Mrs. Monteiro Recd. 6/73.</td>
</tr>
<tr>
<td>Occurrence status</td>
<td>PRESENT Inferred</td>
</tr>
<tr>
<td>Record number</td>
<td>s.n.</td>
</tr>
<tr>
<td>Recorded by</td>
<td>Monteiro</td>
</tr>
<tr>
<td>Providing institution</td>
<td>Royal Botanic Gardens, Kew</td>
</tr>
<tr>
<td>Contributors</td>
<td>Paton, A. (identifier) ; Monteiro (collector) (field number: s.n.)</td>
</tr>
<tr>
<td>Type of object</td>
<td>Preserved Specimen</td>
</tr>
<tr>
<td>Rights statement for the media in this item (unless otherwise specified)</td>
<td><a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a></td>
</tr>
<tr>
<td>User generated content</td>
<td>false</td>
</tr>
<tr>
<td>Date</td>
<td>1843-00-00 ; 2005-11-00 (identification)</td>
</tr>
<tr>
<td>Places</td>
<td>30 miles inland from Ambriz.</td>
</tr>
<tr>
<td>Identifier</td>
<td>KOHERBARIUMXXK000347209 ; K/Herbarium/K000347209 ;</td>
</tr>
<tr>
<td></td>
<td><a href="http://specimens.kew.org/herbarium/K000347209">http://specimens.kew.org/herbarium/K000347209</a></td>
</tr>
</tbody>
</table>

**Europeana**

**GBIF**

**Wikidata**

- **Rose Monteiro**
  - Mrs Joachim John Monteiro; Miss Rose Bassett; Mrs Monteiro; R. Bassett; Mrs J. J. Monteiro; R. Monteiro
  - *May 01, 1840 – February 11, 1898*
  - British naturalist, writer, scientific illustrator and collector (1840–1898)
  - [https://www.wikidata.org/wikidata/Q28519226](https://www.wikidata.org/wikidata/Q28519226)
  - United Kingdom of Great Britain and Ireland

- **Joachim John Monteiro**
  - Jose Maria Correa Monteiro; J. J. Monteiro
  - *1833 – 1878*
  - British mining engineer, colonial administrator, naturalist and explorer
  - [https://www.wikidata.org/wiki/Q94608257](https://www.wikidata.org/wiki/Q94608257)
Example (2)

Source: https://www.gbif.org/occurrence/2434489393

GBIF

EUROPEANA

Suggested content

SOURCE:
https://classic.europeana.eu/portal/en/record/92034/GVNRC_UBA01_I2200097.html
Temminck, et al. Verhandelingen over de natuurlijke geschiedenis der Nederlandsche Overzeesche bezittingen, volume: Zoologie (1839-1844),
https://www.biodiversitylibrary.org/item/200624#page/7/mode/1up
A major challenge in Arts and Humanities data sharing workflows: the cross-sector continuum of data curation

**Cultural Heritage Institutions**
- Selecting, hosting, curating, preserving cultural artefacts
- Digitizing them and making them available online
- Rights management: opening or closing away collections

**Digital Humanities Labs**
- Cleaning, standardizing and enriching these materials
- Opening up existing shared resources for large-scale, computational analysis in ways that align with researcher requirements

**Researcher Communities**
- Analyzing, enriching, reusing, interpreting collections
- Both large-scale and small-scale analyses
- New computational methods change the ways in which we interact with our heritage, enable us to ask new questions
- New insights
- Digital storytelling to a broader audience

**Research Infrastructures, Publishers**
- Hosting, preserving, publishing, sustaining content generated by scholars
- Interlinking different content types
- Discovery environments

---

A frequent case:

The website does not give terms and conditions for the use of images. Source does neither name author nor give copyright information or contact for publication permission.

The problem was not openness but the lack of transparency. Re-usability = sustainability. The longevity of CHI data is defined by its presence in scientific, cultural and social discourses.

How to avoid such cases?
Cultural Heritage Collections as Data

Collections as Data
Providing data-level access to digital collections is a primary challenge for undertaking digital humanities research. In the United States, the flagship initiatives, ‘Always Already Computational: Collections as Data’ and ‘Collections as Data: Part to Whole’, define ‘Collections as Data’ as a “conceptual orientation to collections that renders them as ordered information, stored digitally, so that they are inherently amenable to computation”. The initiative was established to document, exchange experience and share knowledge to encourage cultural heritage institutions to implement ‘collections as data’ in their own institutions. DATA-KBR-BE will kick-start the implementation of ‘Collections As Data’ in Belgium.

https://collectionsasdata.github.io

Providing access to the underlying files of digitised cultural heritage resources to facilitate data analysis by means of tools and methods developed in the field of digital humanities

Madoc - IIIF annotation and crowdsourcing platform

Madoc is an Omeka-based Open Source platform for the display, enrichment, and curation of IIIF-based digital objects. The platform itself is a combination of crowdsourcing services and technology, bound together to provide a single management interface for IIIF collections.

Madoc allows researchers and institutions to:
- Assemble digital collections with IIIF objects from multiple archives, libraries, and museums and build an interactive web site around them.
- Import and enhance existing metadata and image annotations (such as fax, OCR, or segmentation) with your own editorial content.
- Annotate digital objects and image regions with new transcriptions, translations, entities, tags, commentaries, and other material, using a fully interactive 'capture model'.
- Run a crowdsourcing project and invite the public, researchers, or students to contribute.
- Showcase and export the enriched digital collection.

Interlinking with Europeana

Cuneiform tablet with cylinder seal impression

The tablet mentions the purchase or offering of an axe in the third year of Ibbi-Sîn (2028-2004). Seal: a person sits on a cubic chair. He holds his left hand near his belt and lifts the right one up as a sign of homage. Another figure, not visible thanks to the incorrect use of the seal, leads his client to the god. The client, head shaved, is led by the left wrist and lifts up his right hand as a sign of homage. Three-line framed inscription. This seal is well attested in Umma. Same seal on O.114, ur-dîâra dub sar dumu šēkāliš-ta “Ur-Šara, scribe, son of Šeš-kaia”.

A European data space for cultural heritage

Commission proposes a common European data space for cultural heritage

Today, the European Commission publishes a recommendation on a common European data space for cultural heritage. The aim is to accelerate the digitisation of cultural heritage assets.

The objectives

The Commission recommends Member States to accelerate the digitisation of all cultural heritage monuments and sites, objects and artefacts for future generations, to protect and preserve those at risk, and boost their reuse in domains such as education, sustainable tourism and cultural creative sectors.

Bridging the collaboration gap: Europeana, DiSSCo & DARIAH through FAIR integration of cultural and natural heritage data for interdisciplinary research

Menti Meter: https://www.menti.com/
Code: 7335 2893
Discussion

- **Data Challenges:** (e.g. Persistent Identifiers: Workflows; Metadata usability, standardisation, granularity, contextualisation; Digitised Materials: interoperability, e.g. IIIF, annotation, OCR/HTR)

- **Social Challenges:** informal platform to discuss cross-domain data challenges (e.g. Cross-Domain Heritage Working Group, Europeana Research? Network? Aggregators Forum?; Follow-up workshop? Hackathon?)

- **Systemic Challenges:** (Heritage Data Reuse Charter: supporting heritage for research; infrastructural challenges, e.g. changing standards, adoption, if we build it will they come?)
Next Steps

- **Data Challenges:** (e.g. Persistent Identifiers: Workflows; Metadata usability, standardisation, granularity, contextualisation; Digitised Materials: interoperability, e.g. IIIF, annotation, OCR/HTR)

- **Social Challenges:** informal platform to discuss cross-domain data challenges (e.g. Cross-Domain Heritage Working Group, Europeana Research? Network? Aggregators Forum?; Follow-up workshop? Hackathon?)

- **Systemic Challenges:** (Heritage Data Reuse Charter: supporting heritage for research; infrastructural challenges, e.g. changing standards, adoption, if we build it will they come?)
Bridging the collaboration gap: Europeana, DiSSCo & DARIAH through FAIR integration of cultural and natural heritage data for interdisciplinary research

Menti Meter: https://www.menti.com/
Code: 7335 2893
Would you like to continue the conversation?

- Add your ideas here: https://tinyurl.com/CultNatHeritage
- Get in touch! info@dissco.eu

We look forward to hearing from you!!!!