Connecting Collections: Using Linked Data in Libraries, Museums and Archives

Unlock the Libraries : VÖB event, 24 May 2022

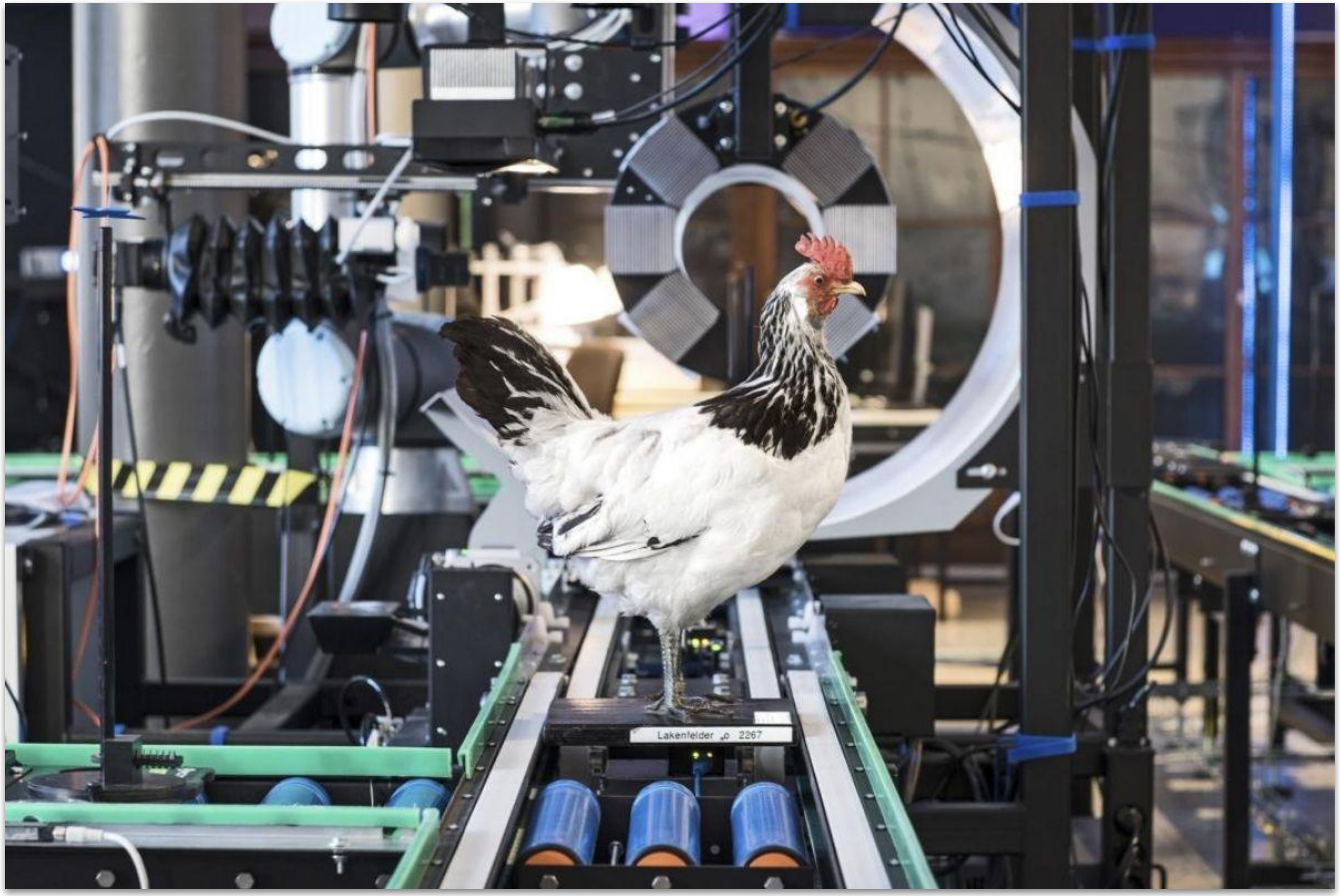
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847693

universität wien

Overview

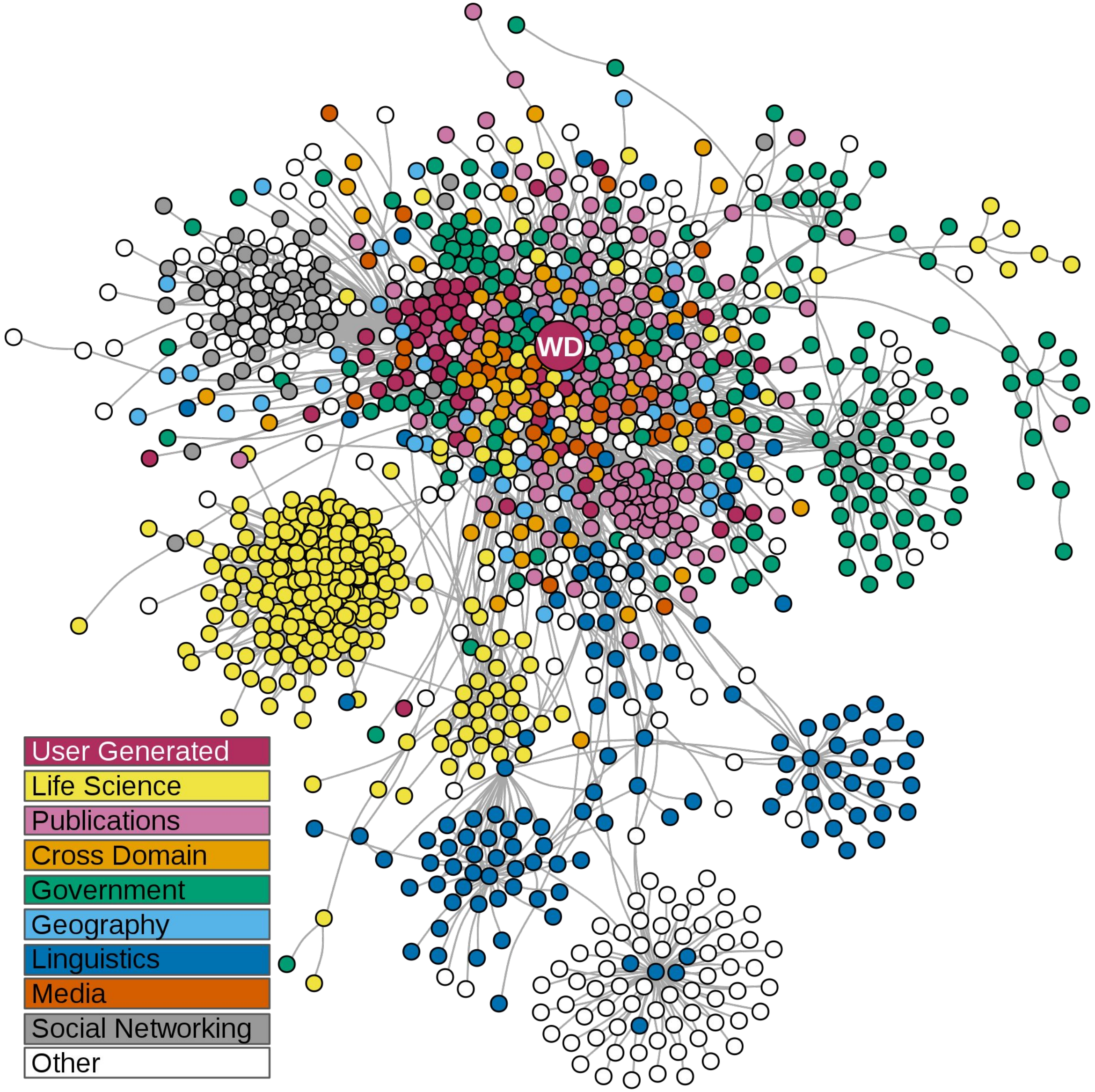
- Part 1: What is Linked Data - and how big is “Big”
- Part 2: What does Linked Data for GLAM look like?
- Part 3: Collections Data vs Collections as Data
- Data Part 4: Linked Data in the wild: the good, the bad & the ugly
- Part 5: Looking forward

Heritage digitisation has been going on for a while, with a variety of objectives.

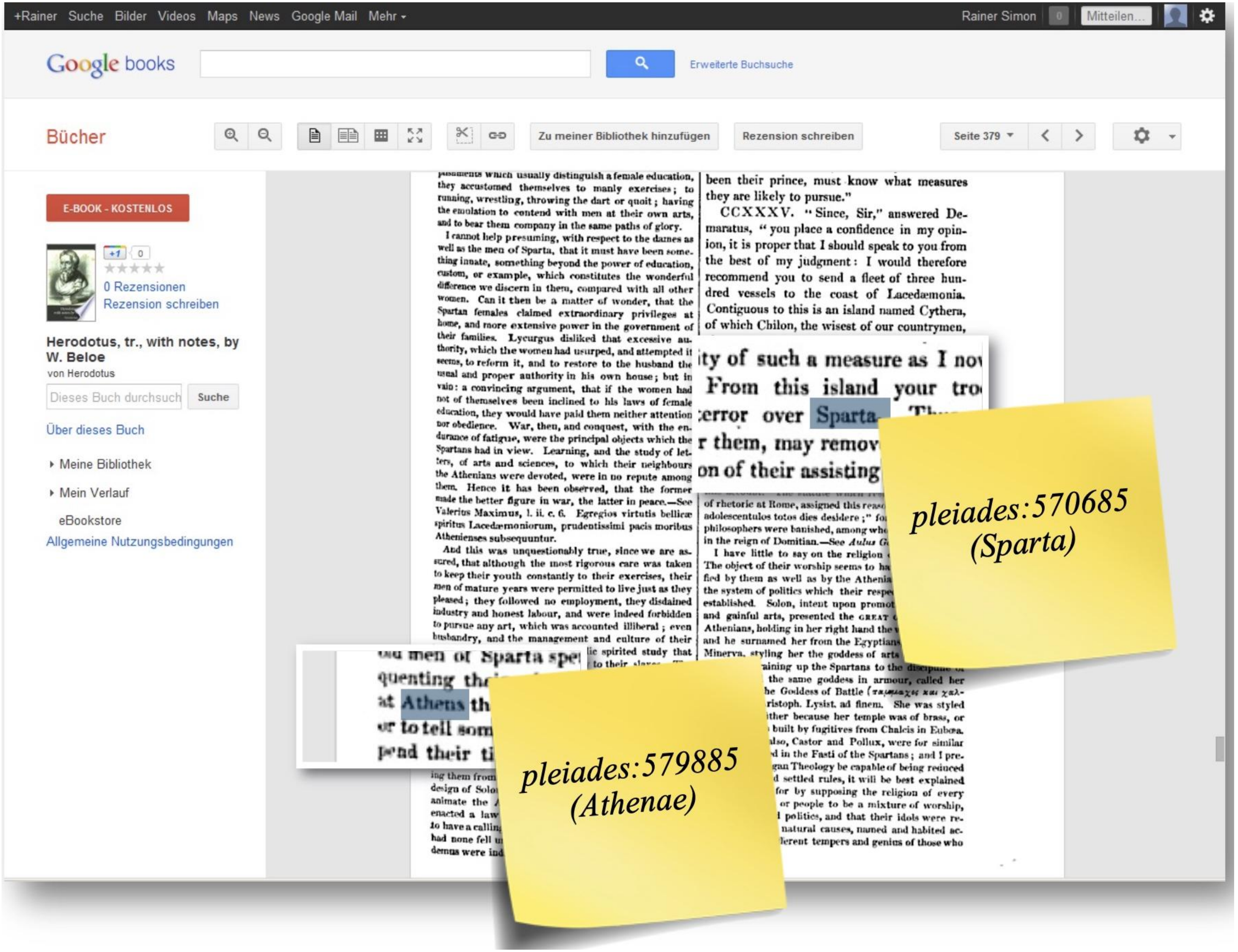


Chicken on a scanner: Museum für Naturkunde, Berlin.

Linked Data is a set of principles for connecting human and machine readable structured data



Semantic Recognition (not quite yet a reality)…



Herodotus, tr., with notes, by W. Beloe
von Herodotus

plishments which usually distinguish a female education, they accustomed themselves to manly exercises; to running, wrestling, throwing the dart or quoit; having the emulation to contend with men at their own arts, and to bear them company in the same paths of glory.

I cannot help presuming, with respect to the dames as well as the men of Sparta, that it must have been something innate, something beyond the power of education, custom, or example, which constitutes the wonderful difference we discern in them, compared with all other women. Can it then be a matter of wonder, that the Spartan females claimed extraordinary privileges at home, and more extensive power in the government of their families. Lycurgus disliked that excessive authority, which the women had usurped, and attempted it seems, to reform it, and to restore to the husband the usual and proper authority in his own house; but in vain: a convincing argument, that if the women had not of themselves been inclined to his laws of female education, they would have paid them neither attention nor obedience. War, then, and conquest, with the endurance of fatigue, were the principal objects which the Spartans had in view. Learning, and the study of letters, of arts and sciences, to which their neighbours the Athenians were devoted, were in no repute among them. Hence it has been observed, that the former made the better figure in war, the latter in peace.—See Valerius Maximus, l. ii. c. 6. Egregios virtutis bellicæ spiritus Lacedæmoniorum, prudentissimi pacis moribus Athenienses subsequuntur.

And this was unquestionably true, since we are assured, that although the most rigorous care was taken to keep their youth constantly to their exercises, their men of mature years were permitted to live just as they pleased; they followed no employment, they disdained industry and honest labour, and were indeed forbidden to pursue any art, which was accounted illiberal; even husbandry, and the management and culture of their

lic spirited study that
to their slaves.

old men of Sparta spe
quenting the
at Athens th
or to tell som
pend their ti

ing them from
design of Solo
animate the
enacted a law
to have a calling
had none fell u
demus were ind

been their prince, must know what measures they are likely to pursue."

CCXXXV. "Since, Sir," answered Demaratus, "you place a confidence in my opinion, it is proper that I should speak to you from the best of my judgment: I would therefore recommend you to send a fleet of three hundred vessels to the coast of Lacedæmonia. Contiguous to this is an island named Cythera, of which Chilon, the wisest of our countrymen,

ity of such a measure as I now
From this island your tro
terror over Sparta. Th
r them, may remov
on of their assisting

of rhetoric at Rome, assigned this reaso
adolescentulos totos dies desidere;" fo
philosophers were banished, among who
in the reign of Domitian.—See *Aulus G*

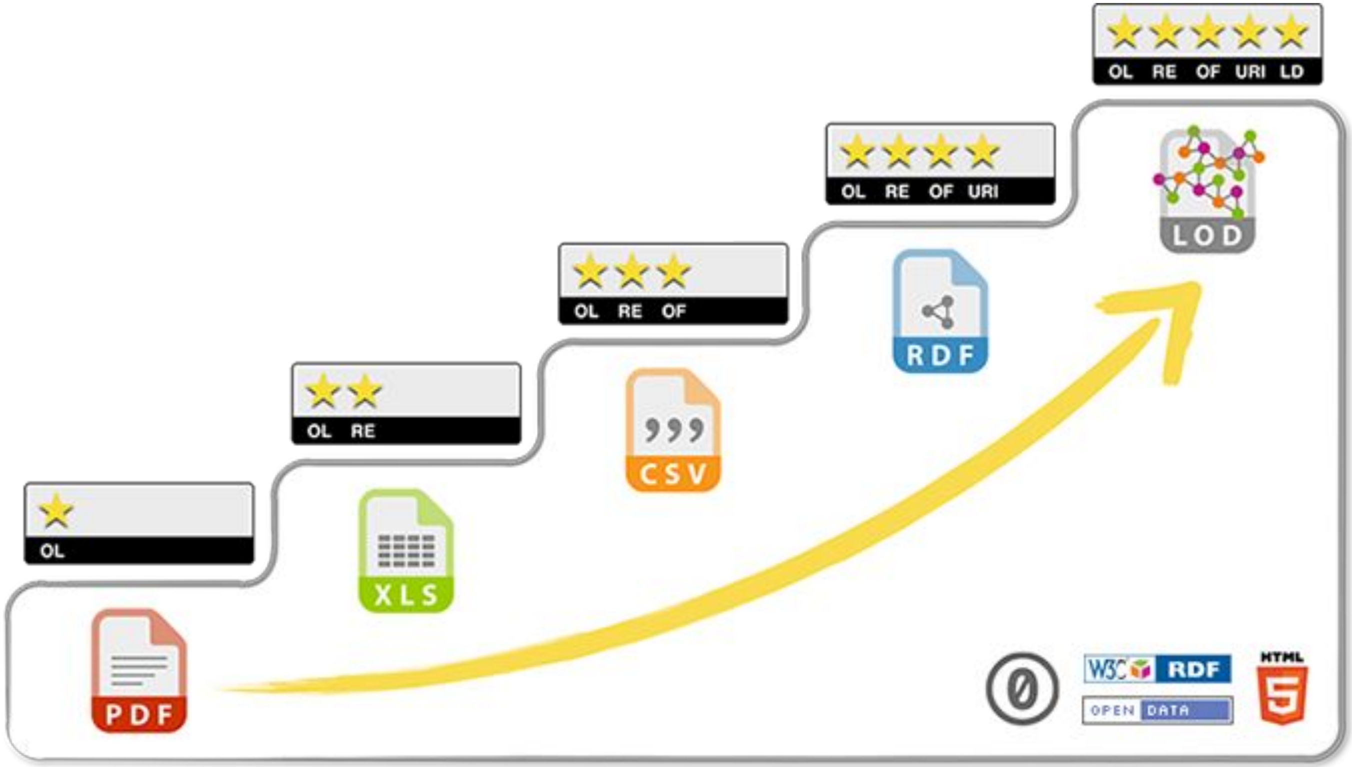
I have little to say on the religion
The object of their worship seems to ha
fied by them as well as by the Athenia
the system of politics which their respe
established. Solon, intent upon promot
and gainful arts, presented the GREAT
Athenians, holding in her right hand the
and he surnamed her from the Egyptians
Minerva, styling her the goddess of arts
aining up the Spartans to the discipline of
the same goddess in armour, called her
he Goddess of Battle (παμμαχος και χαλ-
ristoph. Lysist. ad finem. She was styled
ither because her temple was of brass, or
built by fugitives from Chalcis in Eubœa.
also, Castor and Pollux, were for similar
d in the Fasti of the Spartans; and I pre-
gan Theology be capable of being reduced
d settled rules, it will be best explained
for by supposing the religion of every
or people to be a mixture of worship,
politics, and that their idols were re-
natural causes, named and habited ac-
ferent tempers and genius of those who

pleiades:570685
(Sparta)

pleiades:579885
(Athenae)

universität wien

From 5 star LOD to…



From 5 star LOD to LOUD Data:

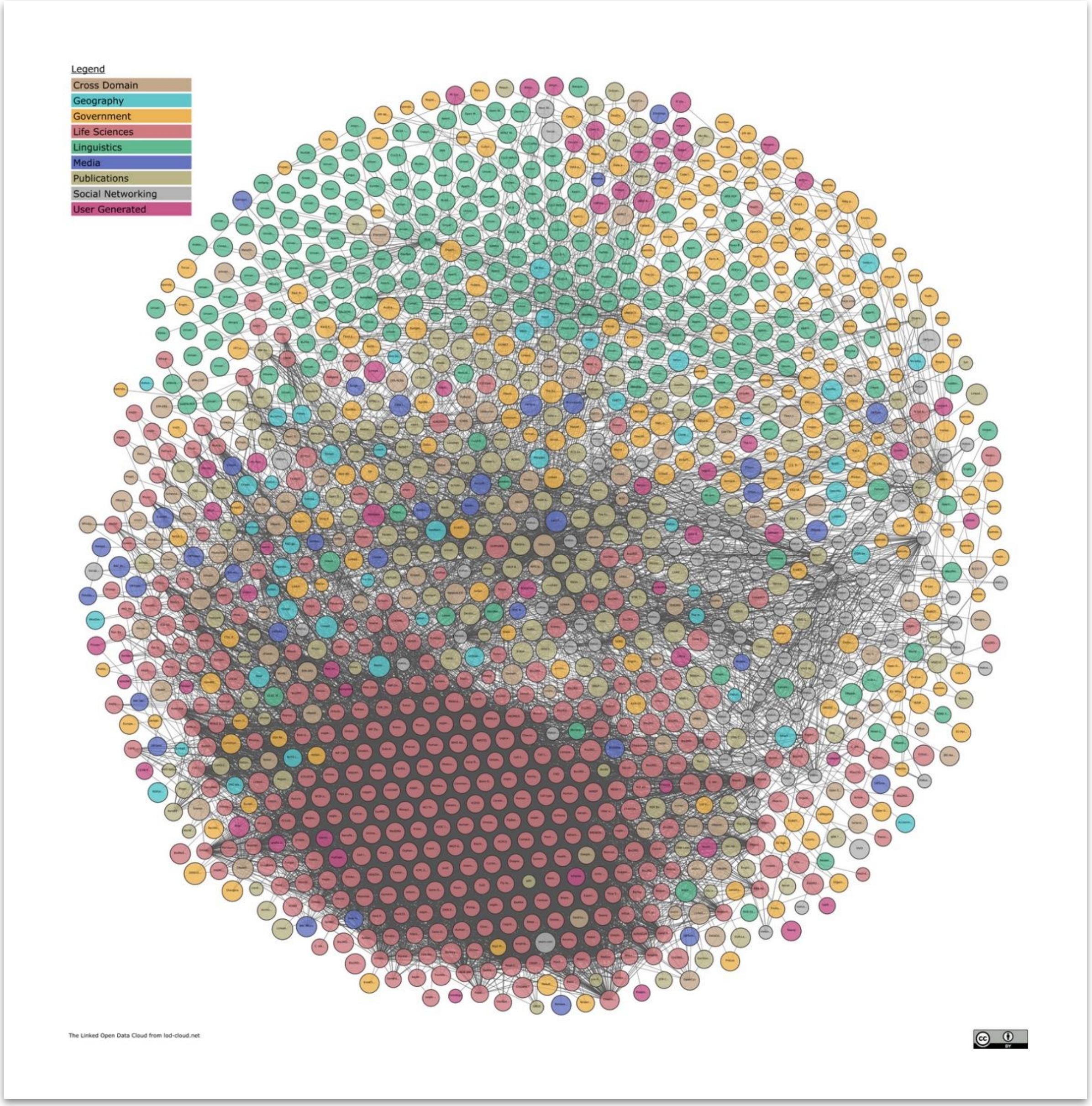
Linked
Open
Usable
Data

→

Abstracted appropriately
Barriers to entry low
Comprehensible
Documented with working examples
Exceptions few, patterns are many

“Big” is relative…

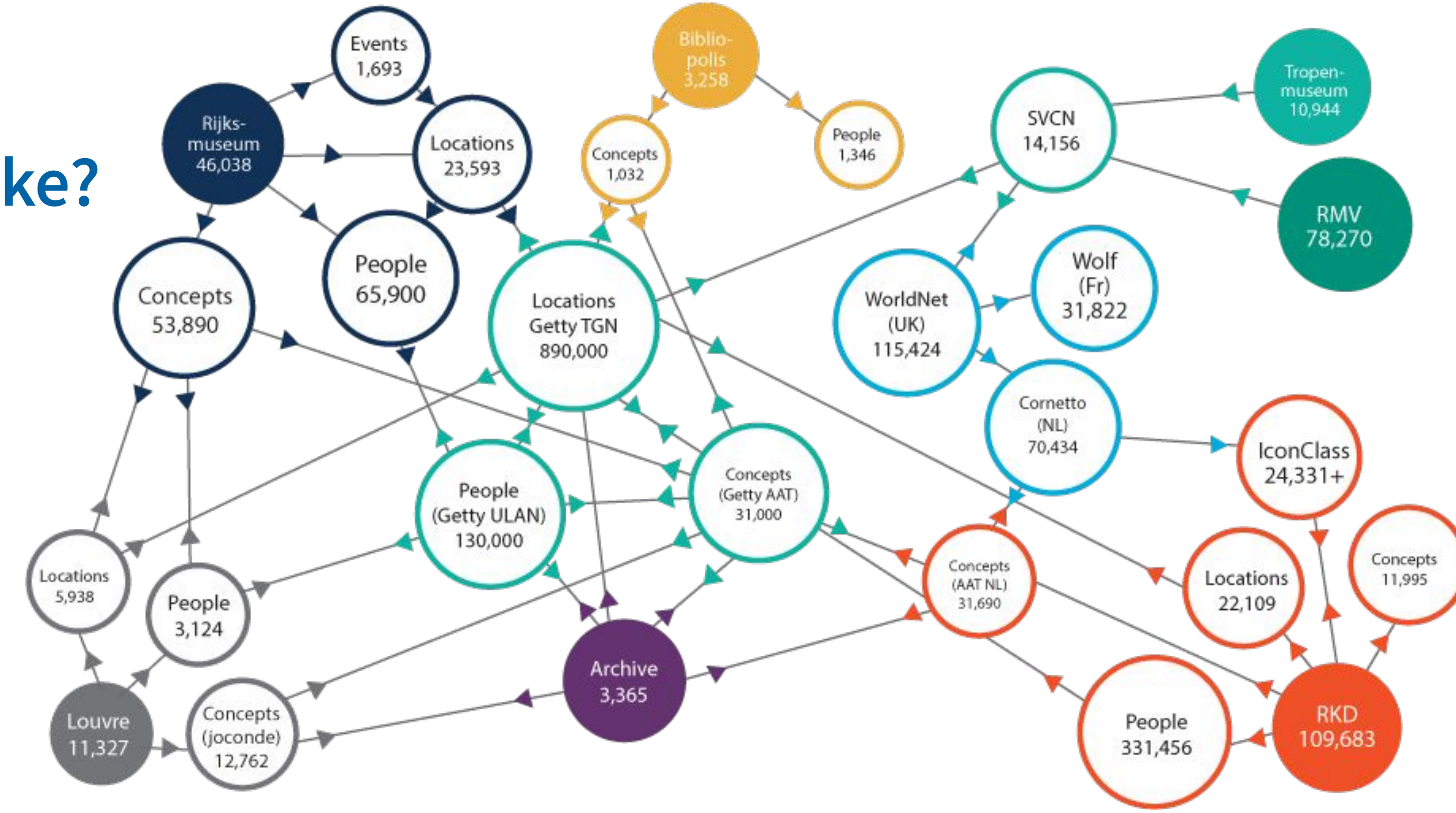
- Compared to other sectors, the heritage contribution is relatively small:
  - British museum: 4 million objects
  - Europeana: 50 million
  - Wikidata 98 million
- Data arrives in many different formats.
- Interoperable systems are hard to build.



https://lod-cloud.net/

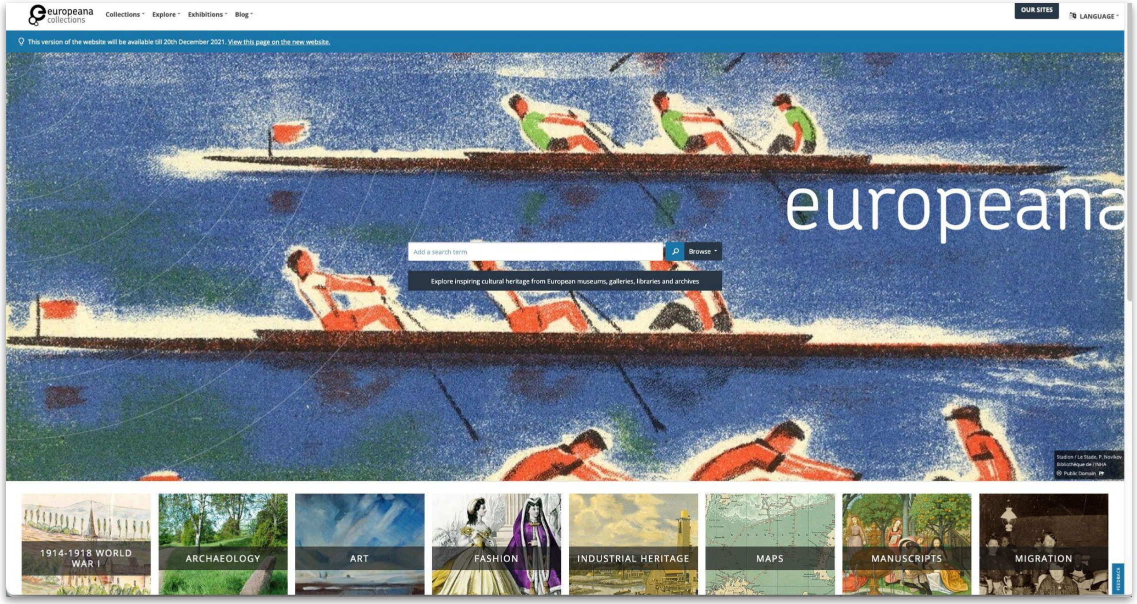
Part 2:
What Does LD for GLAM look like?

- Heritage institutions were early adopters
- Saw the value in linking collections
- Data is heterogenous
- Too many standards!



Large Scale Infrastructures

Europeana



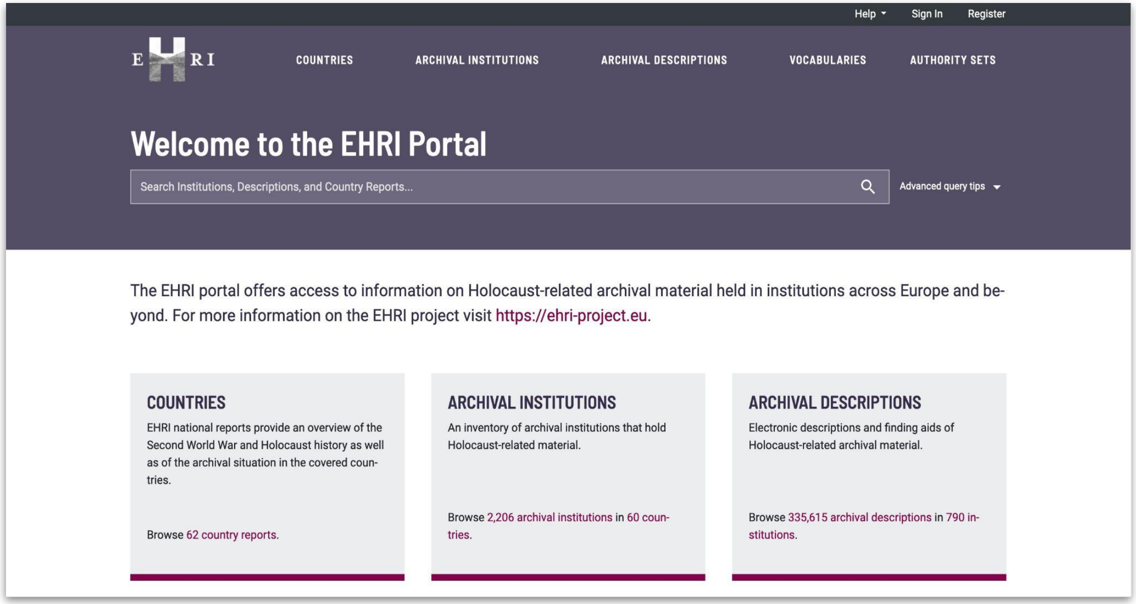
This version of the website will be available till 20th December 2021. View this page on the new website.

Explore inspiring cultural heritage from European museums, galleries, libraries and archives

Stadion / Le Stade, P. Novikov
Bibliothèque de l'INHA
Public Domain

https://www.europeana.eu

European Holocaust Research Infrastructure



Welcome to the EHRI Portal

The EHRI portal offers access to information on Holocaust-related archival material held in institutions across Europe and beyond. For more information on the EHRI project visit https://ehri-project.eu.

COUNTRIES

EHRI national reports provide an overview of the Second World War and Holocaust history as well as of the archival situation in the covered countries.

Browse 62 country reports.

ARCHIVAL INSTITUTIONS

An inventory of archival institutions that hold Holocaust-related material.

Browse 2,206 archival institutions in 60 countries.

ARCHIVAL DESCRIPTIONS

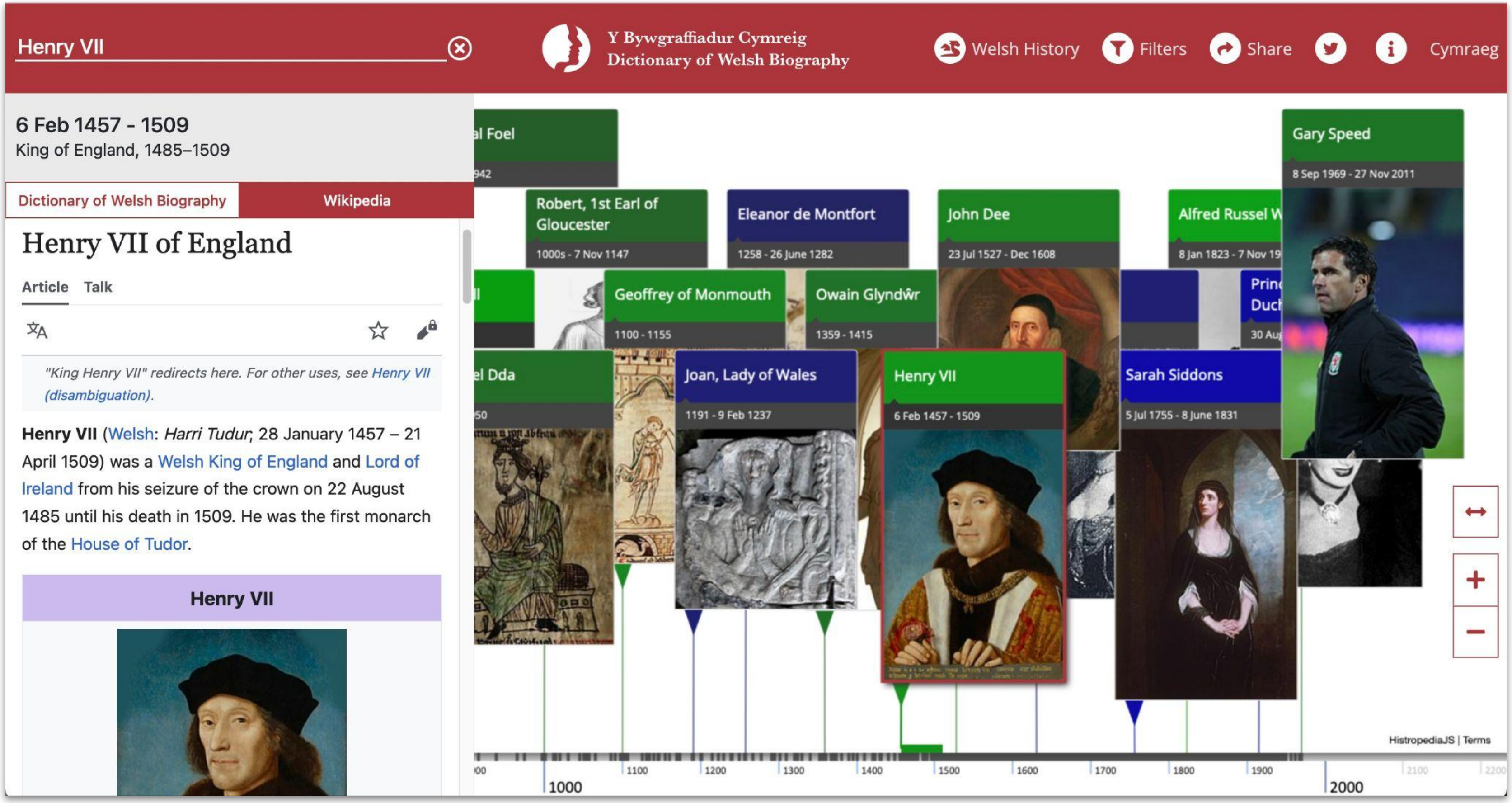
Electronic descriptions and finding aids of Holocaust-related archival material.

Browse 335,615 archival descriptions in 790 institutions.

https://www.ehri-project.eu/

Wikidata

National Library of Wales in Wikidata



Henry VII

Y Bywgraffiadur Cymreig
Dictionary of Welsh Biography

6 Feb 1457 - 1509
King of England, 1485–1509

Dictionary of Welsh Biography | Wikipedia

Henry VII of England

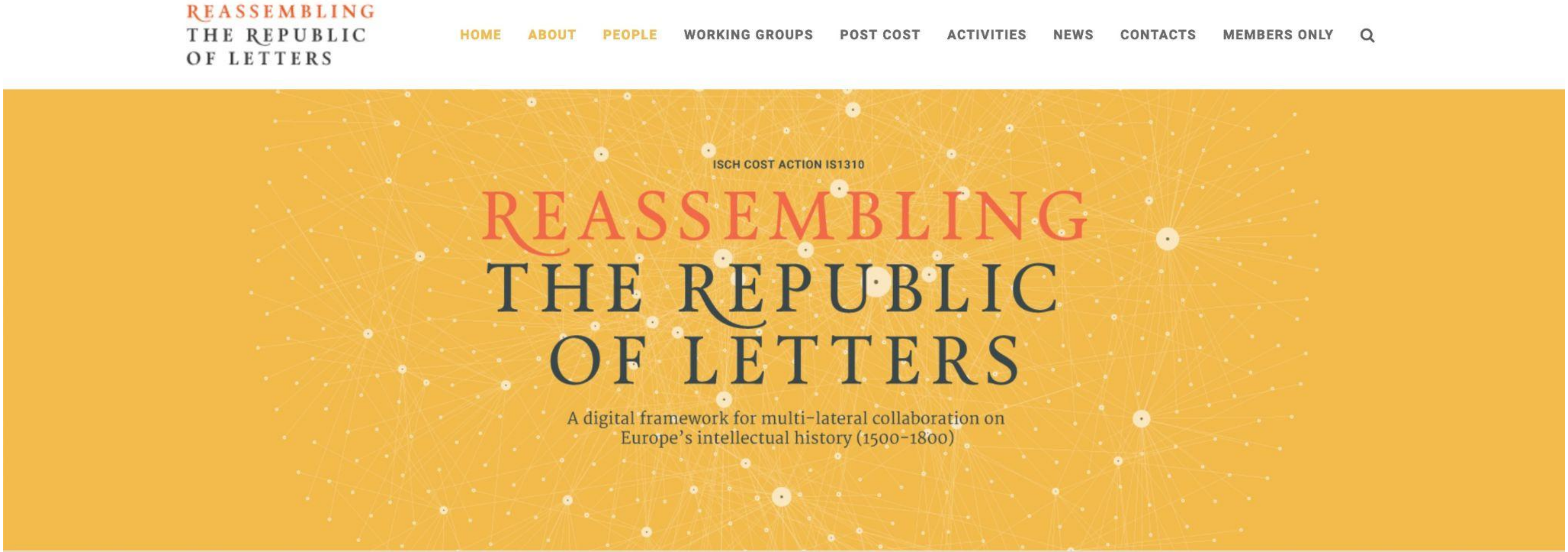
Article Talk

"King Henry VII" redirects here. For other uses, see Henry VII (disambiguation).

Henry VII (Welsh: *Harri Tudur*, 28 January 1457 – 21 April 1509) was a Welsh King of England and Lord of Ireland from his seizure of the crown on 22 August 1485 until his death in 1509. He was the first monarch of the House of Tudor.

Henry VII

Reassembling the Republic of Letters



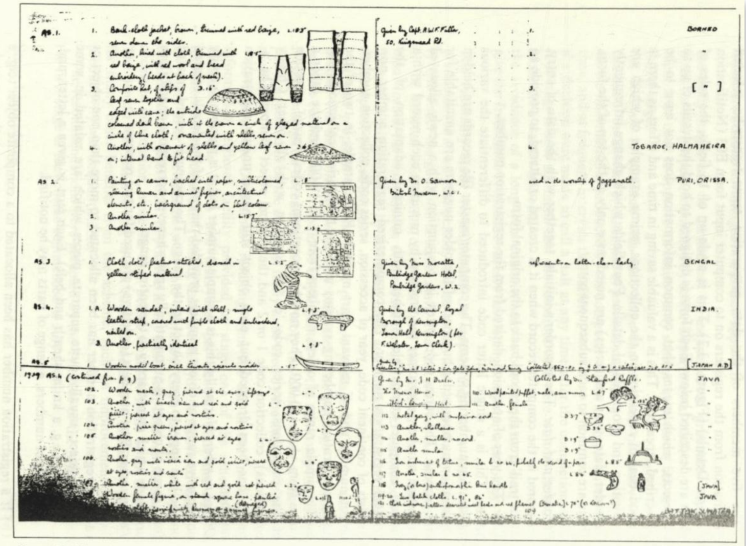
REASSEMBLING THE REPUBLIC OF LETTERS

ISCH COST ACTION IS1310

REASSEMBLING THE REPUBLIC OF LETTERS

A digital framework for multi-lateral collaboration on Europe's intellectual history (1500–1800)

Challenges: Technical…



As.1. 1. Bark-cloth jacket, brown, trimmed with red baize, sewn down the sides. 2. Another, lined with cloth, trimmed with red baize, with red wool and bead embroidery (beads at back of neck). 3. Conical hat, of strips of leaf sewn together and edged with cane; the outside coloured dark brown, with at the crown a circle of glazed material on a circle of blue cloth; ornamented with shells, sewn on. 4. Another, with ornament of shells and yellow leaf sewn on; internal band to fit head. — Given by Capt. A.W.F. Fuller, 50, Kingsmead Rd. — 1. BORNEO 2. " 3. ["] 4. TOBAROE, HALMAHEIRA

As.2. 1. Painting on canvas, backed with paper, multicoloured, showing human and animal figures, architectural elements, etc., background of dots on flat colour. 2. Another similar. 3. Another similar. — Given by Dr. O. Samson, British Museum, W.C.1. — used in the worship of Jaggenath. — PURI, ORISSA.

As.3. 1. Cloth doll, feather-stitched, dressed in yellow striped material. — Given by Miss Mocatta, Pembridge Gardens Hotel, Pembridge Gardens, W.2. — [illegible] — BENGAL

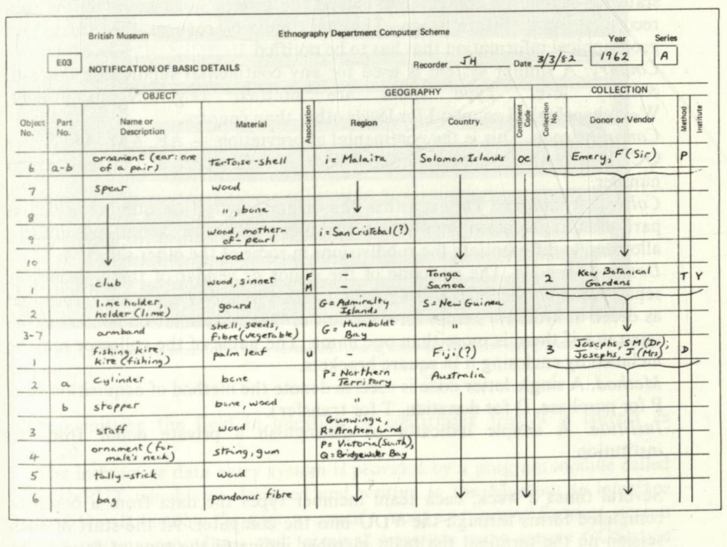
As.4. 1.A. Wooden sandal, inlaid with shell; single leather strap, covered with purple cloth and embroidered, nailed on. B. Another, practically identical. — Given by the Council, Royal Borough of Kensington, Town Hall, Kensington (for F. Webster, Town Clerk). — INDIA.

As.5 1 Wooden model boat, [illegible] — [illegible] — [JAPAN A.D]

1939 As.4 (continued from p 9)
102. Wooden mask, grey, pierced at the eyes, lifesize.
103. Another, with black hair and red and gold [illegible]; pierced at eyes and nostrils.
104. Another, pale green, pierced at eyes and nostrils.
105. Another, smaller, brown, pierced at eyes, nostrils and mouth.
106. Another, grey with black hair and gold [illegible], pierced at eyes, nostrils and mouth.
107. Another, smaller, white with red and gold, not pierced.
108. Wooden female figure, on stand, square base, [illegible] (damaged) [illegible]

Given by Mr. J H Drake, The Manor House, [illegible] — Collected by Sir Stamford Raffles. — JAVA
110. Wood painted puppet, male, arms moving
111. Another, female
112. Metal gong, with suspension cord
113. Another, shallower
114. Another, smaller, no cord
115. Another similar
116. [illegible]
117. Another, similar [illegible]
118. [illegible] — [JAVA]
119-20 Two batik cloths [illegible] — JAVA
121. Cloth [illegible]

Figure 13.1 Sample original register page



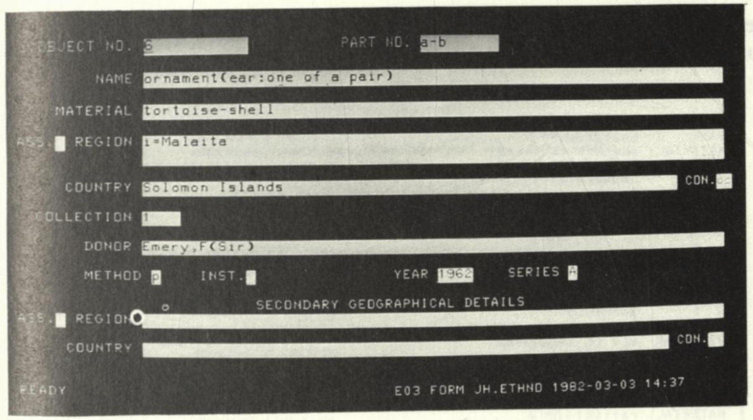
British Museum Ethnography Department Computer Scheme

E03 NOTIFICATION OF BASIC DETAILS

Recorder JH Date 3/3/82 Year 1962 Series A

OBJECT				GEOGRAPHY				COLLECTION			
Object No.	Part No.	Name or Description	Material	Association	Region	Country	Continent Code	Collection No.	Donor or Vendor	Method	Institute
6	a-b	ornament (ear: one of a pair)	TorToise-shell		i = Malaita	Solomon Islands	OC	1	Emery, F (Sir)	P	
7		spear	wood		↓	↓	↓	↓	↓		
8		↓	", bone								
9			wood, mother-of-pearl		i = SanCristobal(?)						
10			wood		"						
1		club	wood, sinnet	F M	– –	Tonga Samoa		2	Kew Botanical Gardens	T	Y
2		lime holder, holder (lime)	gourd		G = Admiralty Islands	S = New Guinea			↓		
3-7		armband	shell, seeds, fibre (vegetable)		G = Humboldt Bay	"					
1		fishing kite, kite (fishing)	palm leaf	u	–	Fiji (?)		3	Josephs, SM (Dr); Josephs, J (Mrs)	D	
2	a	cylinder	bone		P = Northern Territory	Australia			↓		
	b	stopper	bone, wood		"	↓					
3		staff	wood		Gunwingu, R = Arnhem Land						
4		ornament (for male's neck)	string, gum		P = Victoria (South), Q = Bridgewater Bay						
5		tally-stick	wood		↓						
6		bag	pandanus fibre								

Figure 13.2 An E03 input form, with information from the Oceania register for 1962

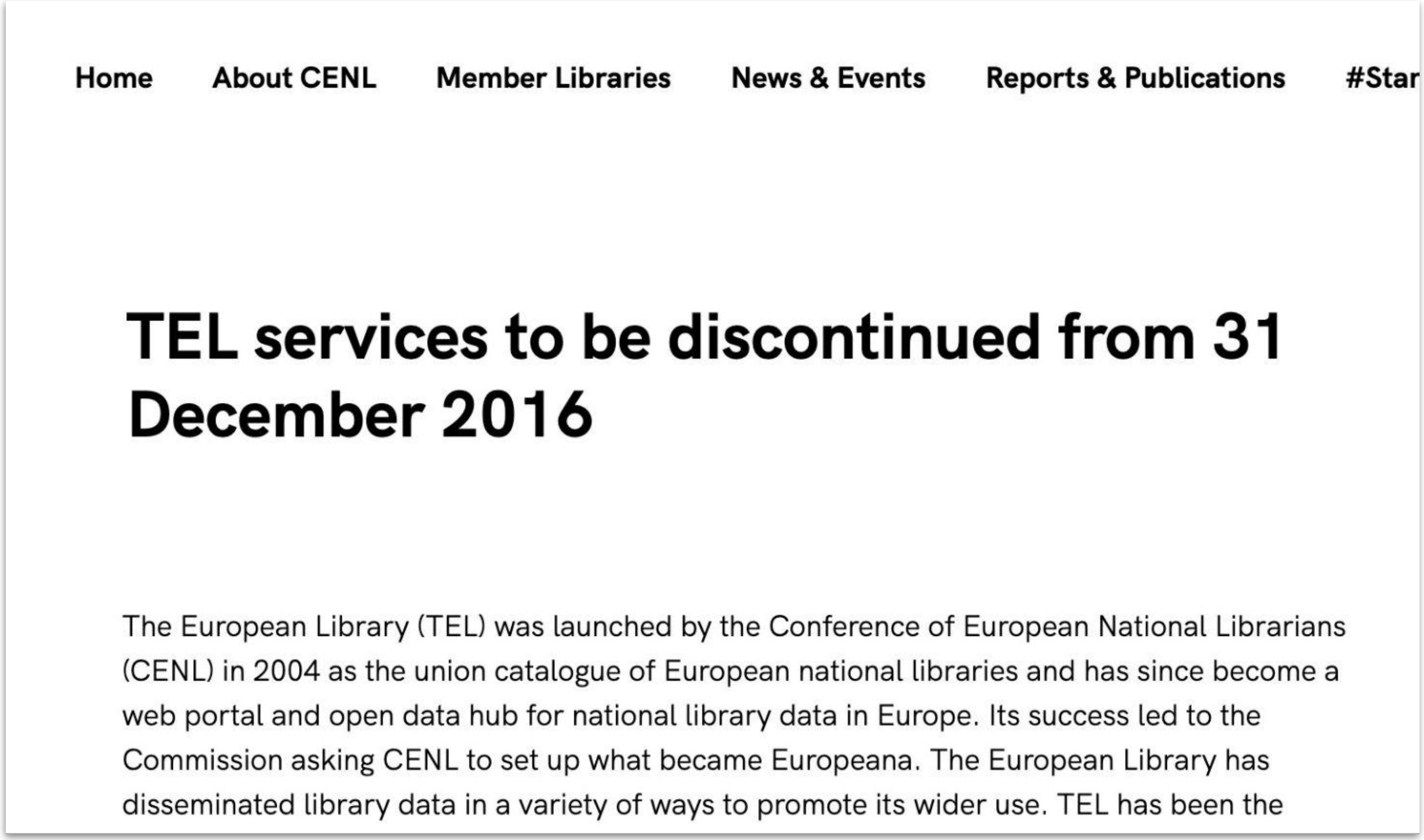


```
OBJECT NO. 6                    PART NO. a-b
      NAME ornament(ear:one of a pair)
  MATERIAL tortoise-shell
ASS.  REGION i=Malaita

   COUNTRY Solomon Islands                         CON.
COLLECTION 1
     DONOR Emery,F(Sir)
    METHOD p    INST.            YEAR 1962    SERIES A
                 SECONDARY GEOGRAPHICAL DETAILS
ASS.  REGION
   COUNTRY                                         CON.

READY                        E03 FORM JH.ETHND 1982-03-03 14:37
```

Figure 13.3 The equivalent form on a terminal screen, showing the first entry from the form in Figure 13.2

McCutcheon, D. (1986) The British Museum. In Light, R. B., Roberts, D., Stewart, J. D. (Eds.), *Museum documentation systems: Developments and applications*

Challenges: Often Political



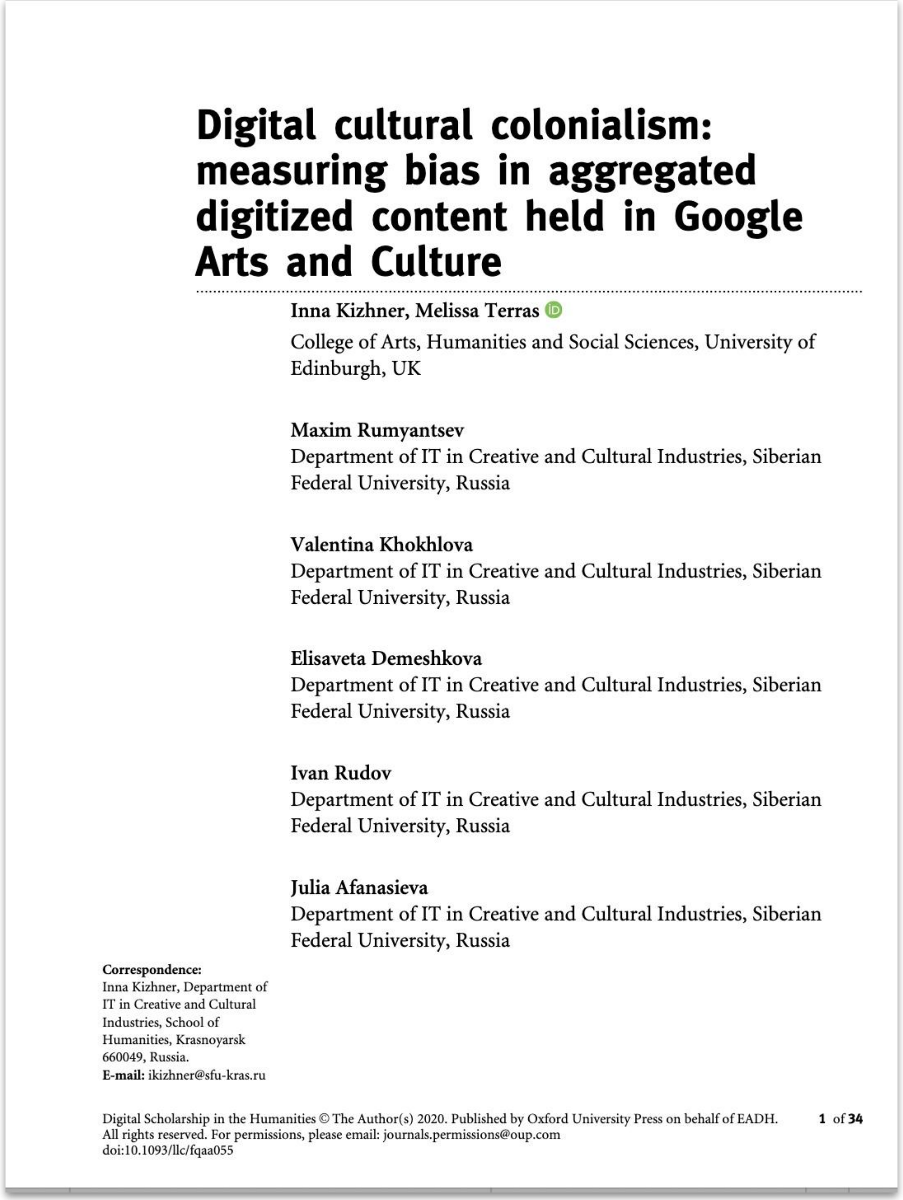
TEL services to be discontinued from 31 December 2016

The European Library (TEL) was launched by the Conference of European National Librarians (CENL) in 2004 as the union catalogue of European national libraries and has since become a web portal and open data hub for national library data in Europe. Its success led to the Commission asking CENL to set up what became Europeana. The European Library has disseminated library data in a variety of ways to promote its wider use. TEL has been the

CENL News, December 13, 2016.
https://www.cenl.org/tel-services-to-be-discontinued-from-31-december-2016/

Part 3: Collections data vs…

- Heritage institutions shape knowledge,
- Heritage professionals know how selection happens,
- What is included has been valourised



Digital cultural colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture

Inna Kizhner, Melissa Terras
College of Arts, Humanities and Social Sciences, University of Edinburgh, UK

Maxim Rumyantsev
Department of IT in Creative and Cultural Industries, Siberian Federal University, Russia

Valentina Khokhlova
Department of IT in Creative and Cultural Industries, Siberian Federal University, Russia

Elisaveta Demeshkova
Department of IT in Creative and Cultural Industries, Siberian Federal University, Russia

Ivan Rudov
Department of IT in Creative and Cultural Industries, Siberian Federal University, Russia

Julia Afanasieva
Department of IT in Creative and Cultural Industries, Siberian Federal University, Russia

Correspondence:
Inna Kizhner, Department of IT in Creative and Cultural Industries, School of Humanities, Krasnoyarsk 660049, Russia.
E-mail: ikizhner@sfu-kras.ru

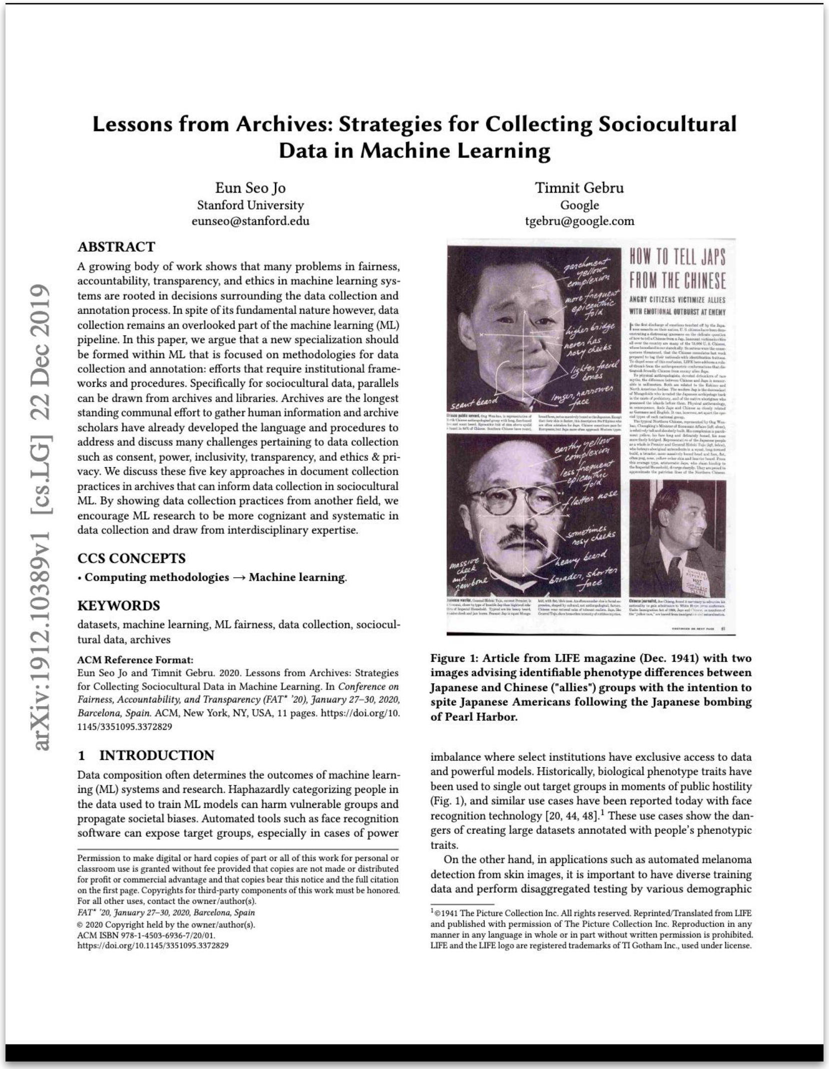
Digital Scholarship in the Humanities © The Author(s) 2020. Published by Oxford University Press on behalf of EADH. All rights reserved. For permissions, please email: journals.permissions@oup.com
doi:10.1093/llc/fqaa055

Collections as Data:
Santa Barbara Statement on Collections as Data

- Aims to encourage computational use of digitized and born digital collections.
- Guided by ongoing ethical commitments.
- Aim to lower barriers to use.
- Designed for everyone, serve no one.
- Shared documentation helps others find a path to doing the work.
- Default open, except in cases where ethical or legal obligations preclude it.
- Data development values interoperability.
- Data stewards work transparently in order to develop trustworthy, long-lived collections.
- Data, as well as the data that describe those data are considered in scope.
- The development of collections as data is an ongoing process and does not necessarily conclude with a final version.
- **https://collectionsasdata.github.io/statement/**

Lessons for AI research

- Pre-curated sources of data are attractive…
- But also complex
  - Jo & Gebru, *Lessons from Archives* (2020): **https://arxiv.org/pdf/1912.10389.pdf**
- “Haphazardly categorizing people in the data used to train ML models can harm vulnerable groups and propagate societal biases”



Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning

Eun Seo Jo
Stanford University
eunseo@stanford.edu

Timnit Gebru
Google
tgebru@google.com

arXiv:1912.10389v1 [cs.LG] 22 Dec 2019

ABSTRACT

A growing body of work shows that many problems in fairness, accountability, transparency, and ethics in machine learning systems are rooted in decisions surrounding the data collection and annotation process. In spite of its fundamental nature however, data collection remains an overlooked part of the machine learning (ML) pipeline. In this paper, we argue that a new specialization should be formed within ML that is focused on methodologies for data collection and annotation: efforts that require institutional frameworks and procedures. Specifically for sociocultural data, parallels can be drawn from archives and libraries. Archives are the longest standing communal effort to gather human information and archive scholars have already developed the language and procedures to address and discuss many challenges pertaining to data collection such as consent, power, inclusivity, transparency, and ethics & privacy. We discuss these five key approaches in document collection practices in archives that can inform data collection in sociocultural ML. By showing data collection practices from another field, we encourage ML research to be more cognizant and systematic in data collection and draw from interdisciplinary expertise.

CCS CONCEPTS

• **Computing methodologies → Machine learning**.

KEYWORDS

datasets, machine learning, ML fairness, data collection, sociocultural data, archives

ACM Reference Format:
Eun Seo Jo and Timnit Gebru. 2020. Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning. In *Conference on Fairness, Accountability, and Transparency (FAT* '20), January 27–30, 2020, Barcelona, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3351095.3372829

HOW TO TELL JAPS FROM THE CHINESE
ANGRY CITIZENS VICTIMIZE ALLIES WITH EMOTIONAL OUTBURST AT ENEMY

Figure 1: Article from LIFE magazine (Dec. 1941) with two images advising identifiable phenotype differences between Japanese and Chinese ("allies") groups with the intention to spite Japanese Americans following the Japanese bombing of Pearl Harbor.

1 INTRODUCTION

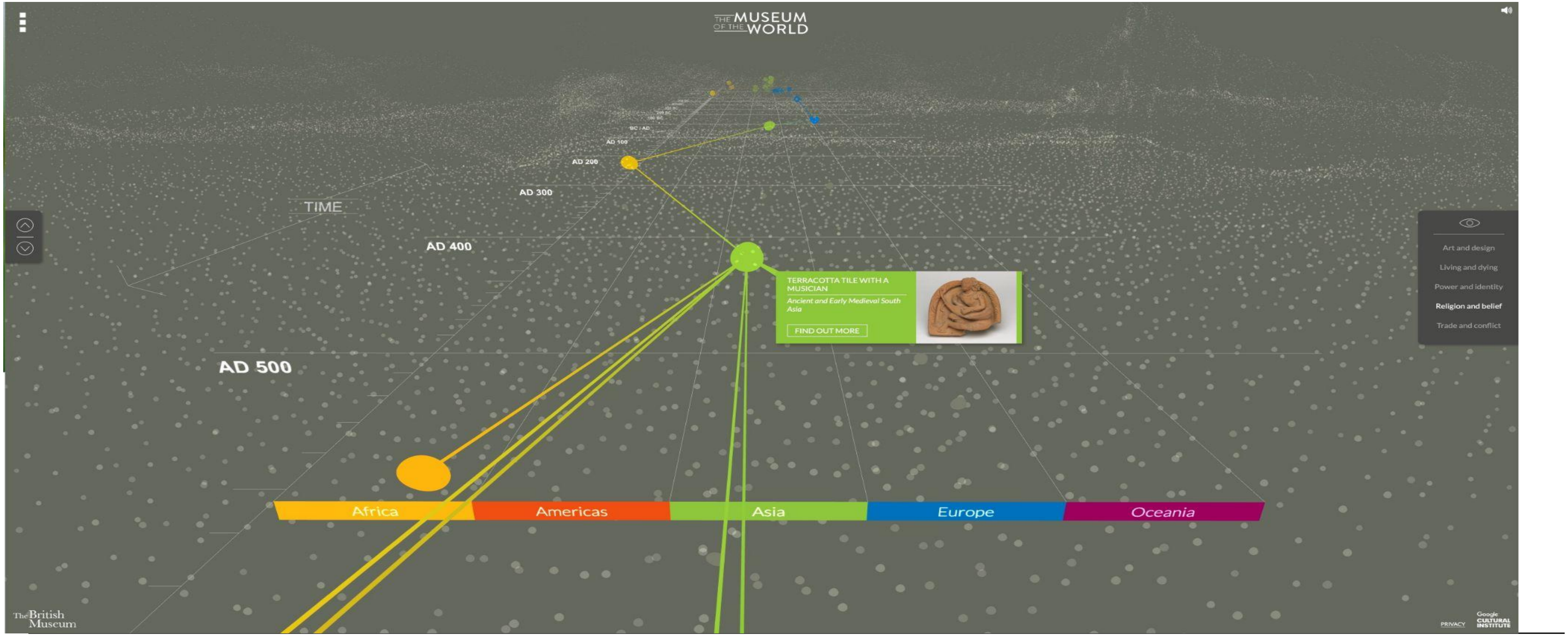
Data composition often determines the outcomes of machine learning (ML) systems and research. Haphazardly categorizing people in the data used to train ML models can harm vulnerable groups and propagate societal biases. Automated tools such as face recognition software can expose target groups, especially in cases of power imbalance where select institutions have exclusive access to data and powerful models. Historically, biological phenotype traits have been used to single out target groups in moments of public hostility (Fig. 1), and similar use cases have been reported today with face recognition technology [20, 44, 48].[1] These use cases show the dangers of creating large datasets annotated with people's phenotypic traits.

On the other hand, in applications such as automated melanoma detection from skin images, it is important to have diverse training data and perform disaggregated testing by various demographic

FAT '20, January 27–30, 2020, Barcelona, Spain*
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6936-7/20/01.
https://doi.org/10.1145/3351095.3372829

[1] ©1941 The Picture Collection Inc. All rights reserved. Reprinted/Translated from LIFE and published with permission of The Picture Collection Inc. Reproduction in any manner in any language in whole or in part without written permission is prohibited. LIFE and the LIFE logo are registered trademarks of TI Gotham Inc., used under license.

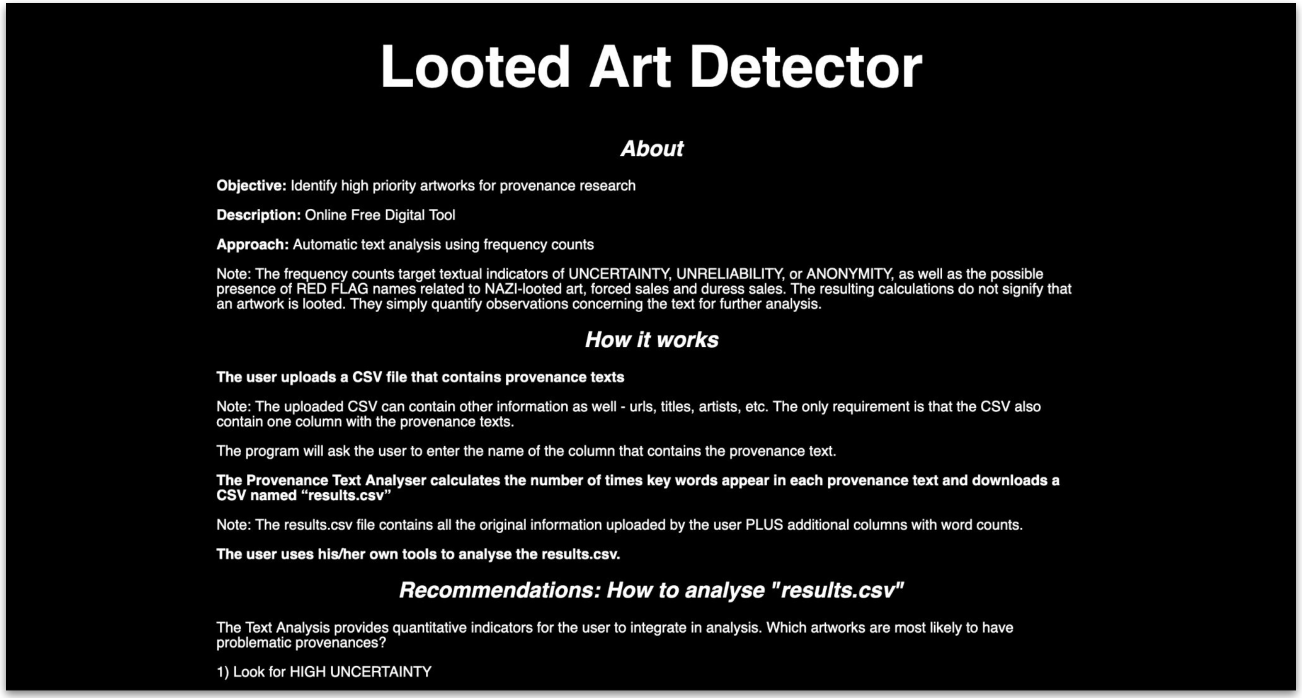
Part 4: GLAM data projects: The good, the bad and the ugly



TERRACOTTA TILE WITH A MUSICIAN

Ancient and Early Medieval South Asia

Linked Data projects: the good



Looted Art Detector

About

Objective: Identify high priority artworks for provenance research

Description: Online Free Digital Tool

Approach: Automatic text analysis using frequency counts

Note: The frequency counts target textual indicators of UNCERTAINTY, UNRELIABILITY, or ANONYMITY, as well as the possible presence of RED FLAG names related to NAZI-looted art, forced sales and duress sales. The resulting calculations do not signify that an artwork is looted. They simply quantify observations concerning the text for further analysis.

How it works

The user uploads a CSV file that contains provenance texts

Note: The uploaded CSV can contain other information as well - urls, titles, artists, etc. The only requirement is that the CSV also contain one column with the provenance texts.

The program will ask the user to enter the name of the column that contains the provenance text.

The Provenance Text Analyser calculates the number of times key words appear in each provenance text and downloads a CSV named "results.csv"

Note: The results.csv file contains all the original information uploaded by the user PLUS additional columns with word counts.

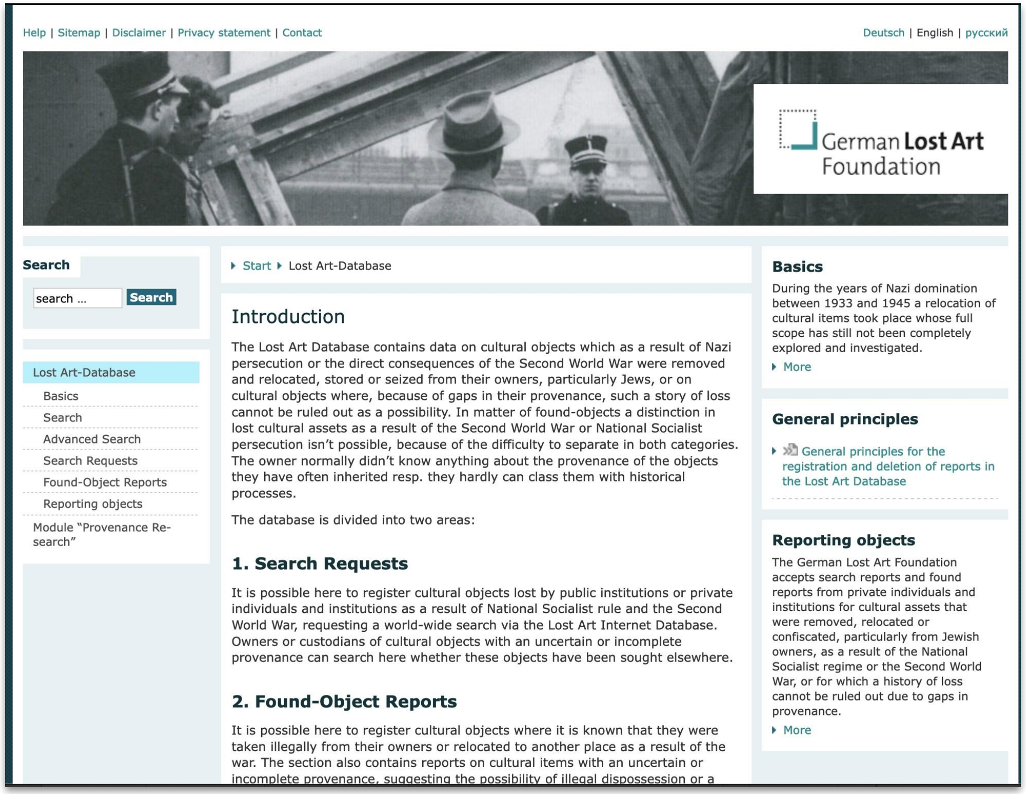
The user uses his/her own tools to analyse the results.csv.

Recommendations: How to analyse "results.csv"

The Text Analysis provides quantitative indicators for the user to integrate in analysis. Which artworks are most likely to have problematic provenances?

1) Look for HIGH UNCERTAINTY

https://artdata.pythonanywhere.com/about/



Introduction

The Lost Art Database contains data on cultural objects which as a result of Nazi persecution or the direct consequences of the Second World War were removed and relocated, stored or seized from their owners, particularly Jews, or on cultural objects where, because of gaps in their provenance, such a story of loss cannot be ruled out as a possibility. In matter of found-objects a distinction in lost cultural assets as a result of the Second World War or National Socialist persecution isn't possible, because of the difficulty to separate in both categories. The owner normally didn't know anything about the provenance of the objects they have often inherited resp. they hardly can class them with historical processes.

The database is divided into two areas:

1. Search Requests

It is possible here to register cultural objects lost by public institutions or private individuals and institutions as a result of National Socialist rule and the Second World War, requesting a world-wide search via the Lost Art Internet Database. Owners or custodians of cultural objects with an uncertain or incomplete provenance can search here whether these objects have been sought elsewhere.

2. Found-Object Reports

It is possible here to register cultural objects where it is known that they were taken illegally from their owners or relocated to another place as a result of the war. The section also contains reports on cultural items with an uncertain or incomplete provenance, suggesting the possibility of illegal dispossession or a

Basics

During the years of Nazi domination between 1933 and 1945 a relocation of cultural items took place whose full scope has still not been completely explored and investigated.

▸ More

General principles

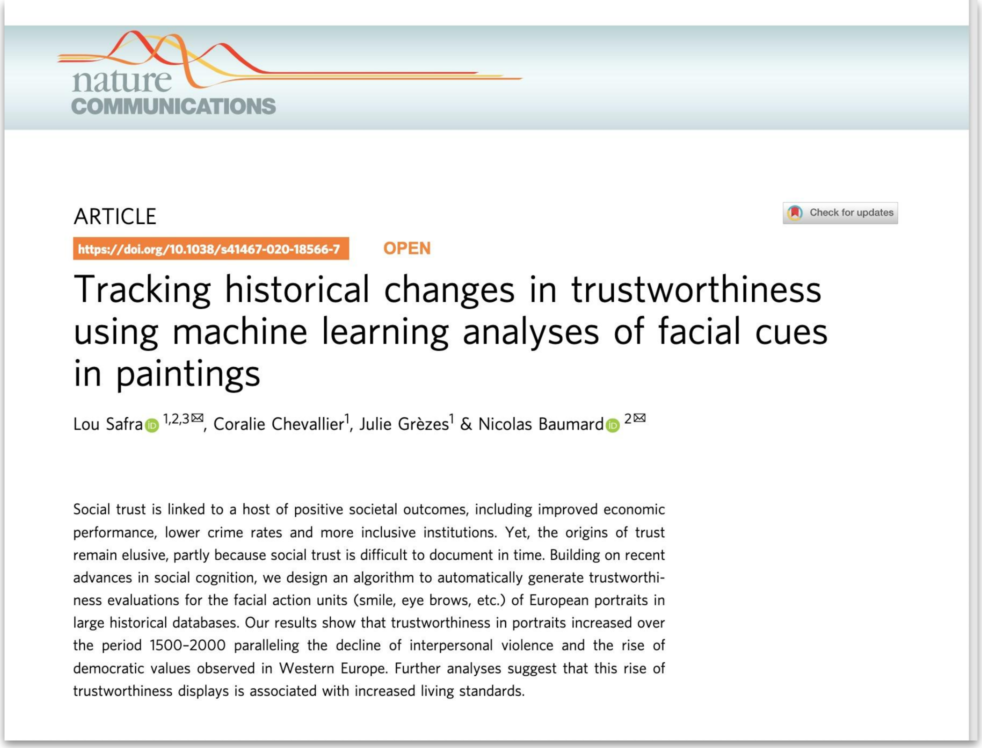
▸ General principles for the registration and deletion of reports in the Lost Art Database

Reporting objects

The German Lost Art Foundation accepts search reports and found reports from private individuals and institutions for cultural assets that were removed, relocated or confiscated, particularly from Jewish owners, as a result of the National Socialist regime or the Second World War, or for which a history of loss cannot be ruled out due to gaps in provenance.

▸ More

Linked Data Projects: the bad



nature COMMUNICATIONS

ARTICLE

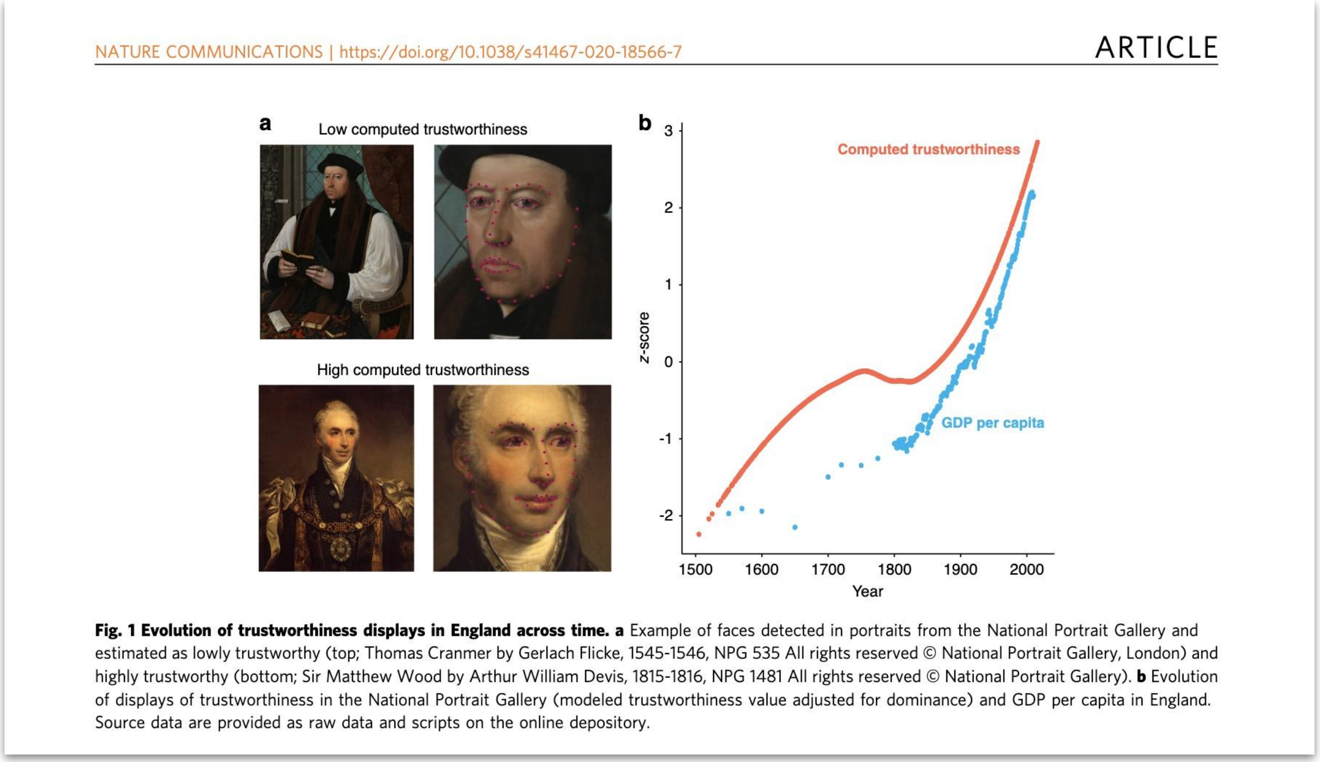
Check for updates

https://doi.org/10.1038/s41467-020-18566-7 OPEN

Tracking historical changes in trustworthiness using machine learning analyses of facial cues in paintings

Lou Safra [1,2,3]✉, Coralie Chevallier[1], Julie Grèzes[1] & Nicolas Baumard [2]✉

Social trust is linked to a host of positive societal outcomes, including improved economic performance, lower crime rates and more inclusive institutions. Yet, the origins of trust remain elusive, partly because social trust is difficult to document in time. Building on recent advances in social cognition, we design an algorithm to automatically generate trustworthiness evaluations for the facial action units (smile, eye brows, etc.) of European portraits in large historical databases. Our results show that trustworthiness in portraits increased over the period 1500–2000 paralleling the decline of interpersonal violence and the rise of democratic values observed in Western Europe. Further analyses suggest that this rise of trustworthiness displays is associated with increased living standards.

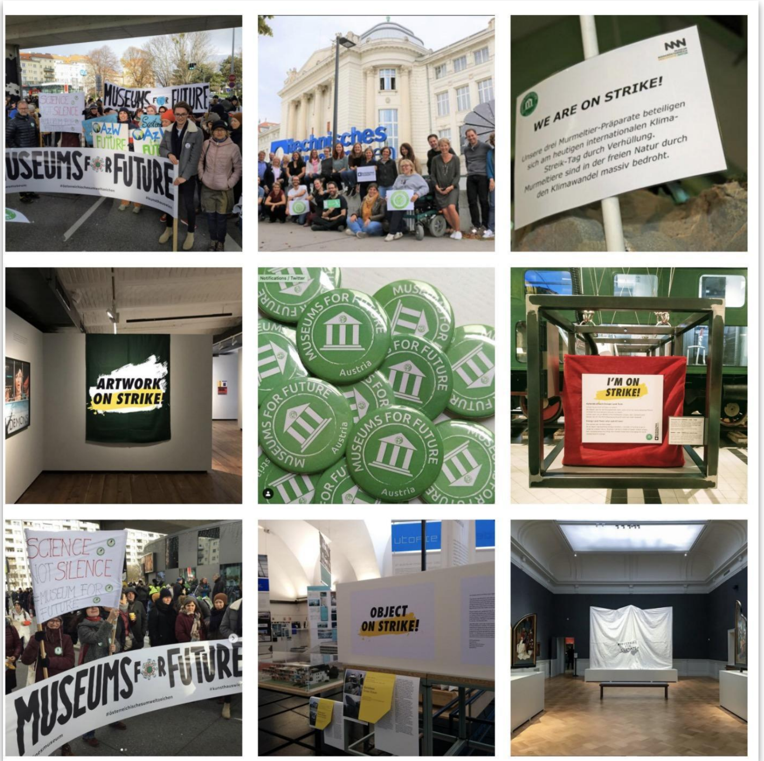


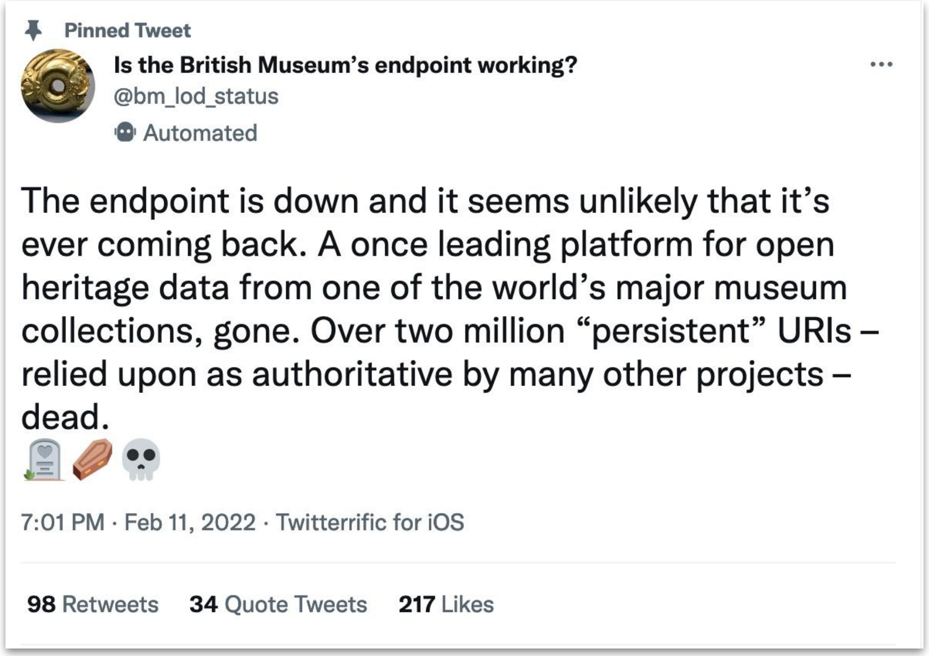
NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-18566-7 ARTICLE

Fig. 1 Evolution of trustworthiness displays in England across time. a Example of faces detected in portraits from the National Portrait Gallery and estimated as lowly trustworthy (top; Thomas Cranmer by Gerlach Flicke, 1545-1546, NPG 535 All rights reserved © National Portrait Gallery, London) and highly trustworthy (bottom; Sir Matthew Wood by Arthur William Devis, 1815-1816, NPG 1481 All rights reserved © National Portrait Gallery). **b** Evolution of displays of trustworthiness in the National Portrait Gallery (modeled trustworthiness value adjusted for dominance) and GDP per capita in England. Source data are provided as raw data and scripts on the online depository.

Screenshot of cover page and figures from the trustworthiness study "Tracking historical changes in trustworthiness using machine learning analyses of facial cues in paintings" conducted by Baumard, et al on European portraits (courtesy and via Nature.com open access) Full paper at: https://rdcu.be/b8PAF

Linked Data projects: the ugly little secrets


SCIENCE NOT SILENCE
MUSEUMS FOR FUTURE
Technisches
WE ARE ON STRIKE!
Unsere drei Murmeltier-Präparate beteiligen sich am heutigen internationalen Klima-Streik-Tag durch Verhüllung. Murmeltiere sind in der freien Natur durch den Klimawandel massiv bedroht.
ARTWORK ON STRIKE!
MUSEUMS FOR FUTURE Austria
I'M ON STRIKE!
SCIENCE NOT SILENCE MUSEUM FOR FUTURE
MUSEUMS FOR FUTURE
OBJECT ON STRIKE!


Pinned Tweet

Is the British Museum's endpoint working?
@bm_lod_status
Automated

The endpoint is down and it seems unlikely that it's ever coming back. A once leading platform for open heritage data from one of the world's major museum collections, gone. Over two million "persistent" URIs – relied upon as authoritative by many other projects – dead.
🪦⚰️💀

7:01 PM · Feb 11, 2022 · Twitterrific for iOS

98 Retweets **34** Quote Tweets **217** Likes

Why this matters:

> “Museums are seen as a beacon of trust… The level of criticality museums have when considering collections is the same level they need to have when it comes to developing digital applications.”

Dr Oonah Murphy, Goldsmiths University, Museums + AI Network.

Pratt | Goldsmiths UNIVERSITY OF LONDON

THE MUSEUMS + AI NETWORK

AI: A Museum Planning Toolkit

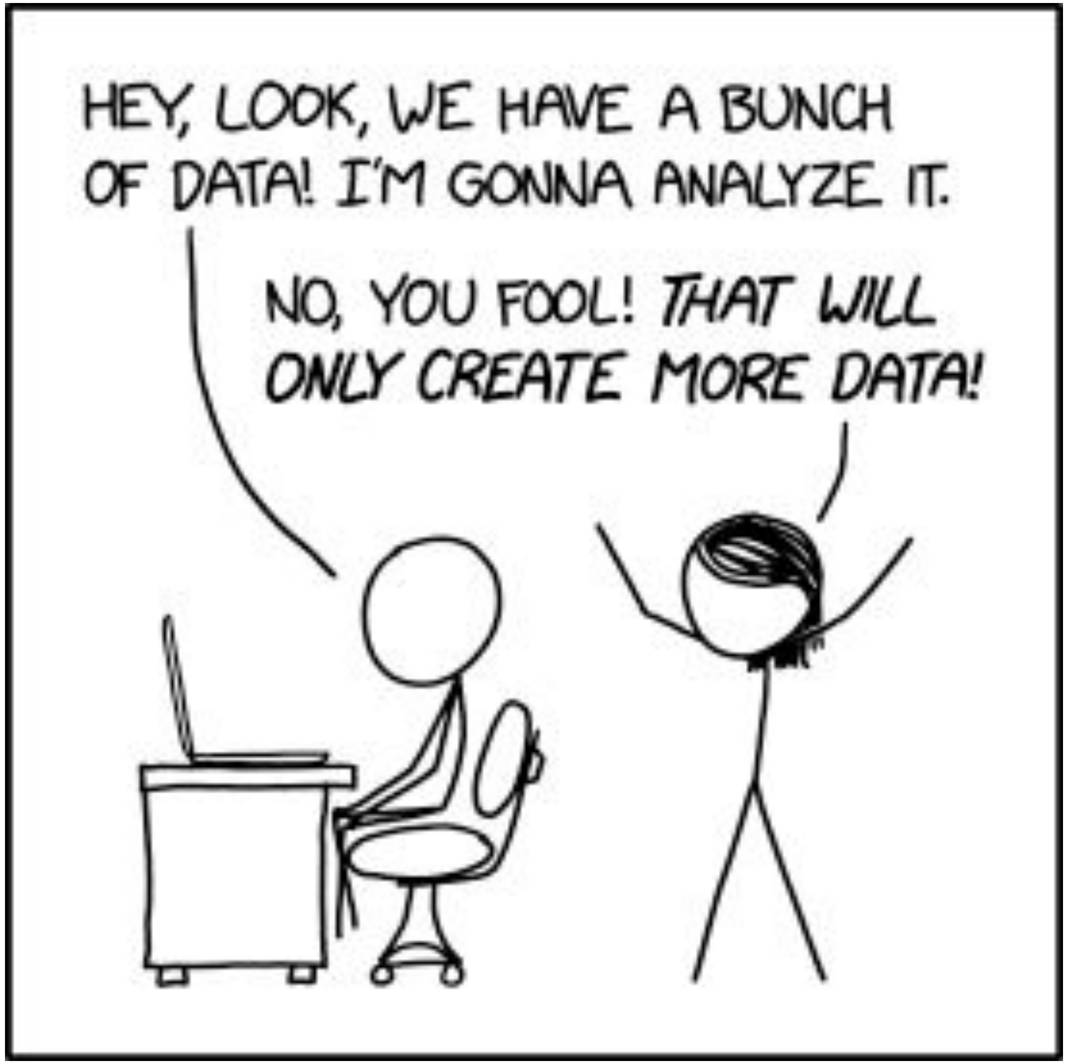
Dr Oonagh Murphy
Dr Elena Villaespesa

www.themuseumsai.network/toolkit/

Thank you!

rebecca.kahn@univie.ac.at

@rebamex


HEY, LOOK, WE HAVE A BUNCH OF DATA! I'M GONNA ANALYZE IT.
NO, YOU FOOL! THAT WILL ONLY CREATE MORE DATA!

Data Trap: xkcd, CC BY-SA https://xkcd.com/2582/