Developments in digital scholarship: at the British Library and at kitchen tables everywhere

Dr. Mia Ridge, @mia_out
Digital Curator, British Library
digitalresearch@bl.uk @BL_DigiSchol

Mitchell Library, State Library of New South Wales
http://acmssearch.sl.nsw.gov.au/search/itemDetailPaged.cgi?itemID=442696

With thanks to... teamwork

The Digital Research Team is a cross-disciplinary mix of curators, researchers, librarians and programmers supporting the creation and innovative use of British Library's digital collections.

@AquilesBrayner @ndalyrose @mia_out @miss_wisdom

@benosteen @mahendra_mahey

Supporting digital scholarship @ BL

We support new ways of exploring and accessing our collections through:

- Getting digital and digitised content available online
- Collaborative projects
- Digital scholarship support and guidance
- Events, publications and training
- Competitions, and awards (BL Labs)

https://www.flickr.com/photos/britishlibrary/11092384623/
Image taken from page 384 of 'Man, embracing his origin, ... civilization, ... mental and moral faculties. ... Illustrated'
British Library HMNTS 10007.g.36

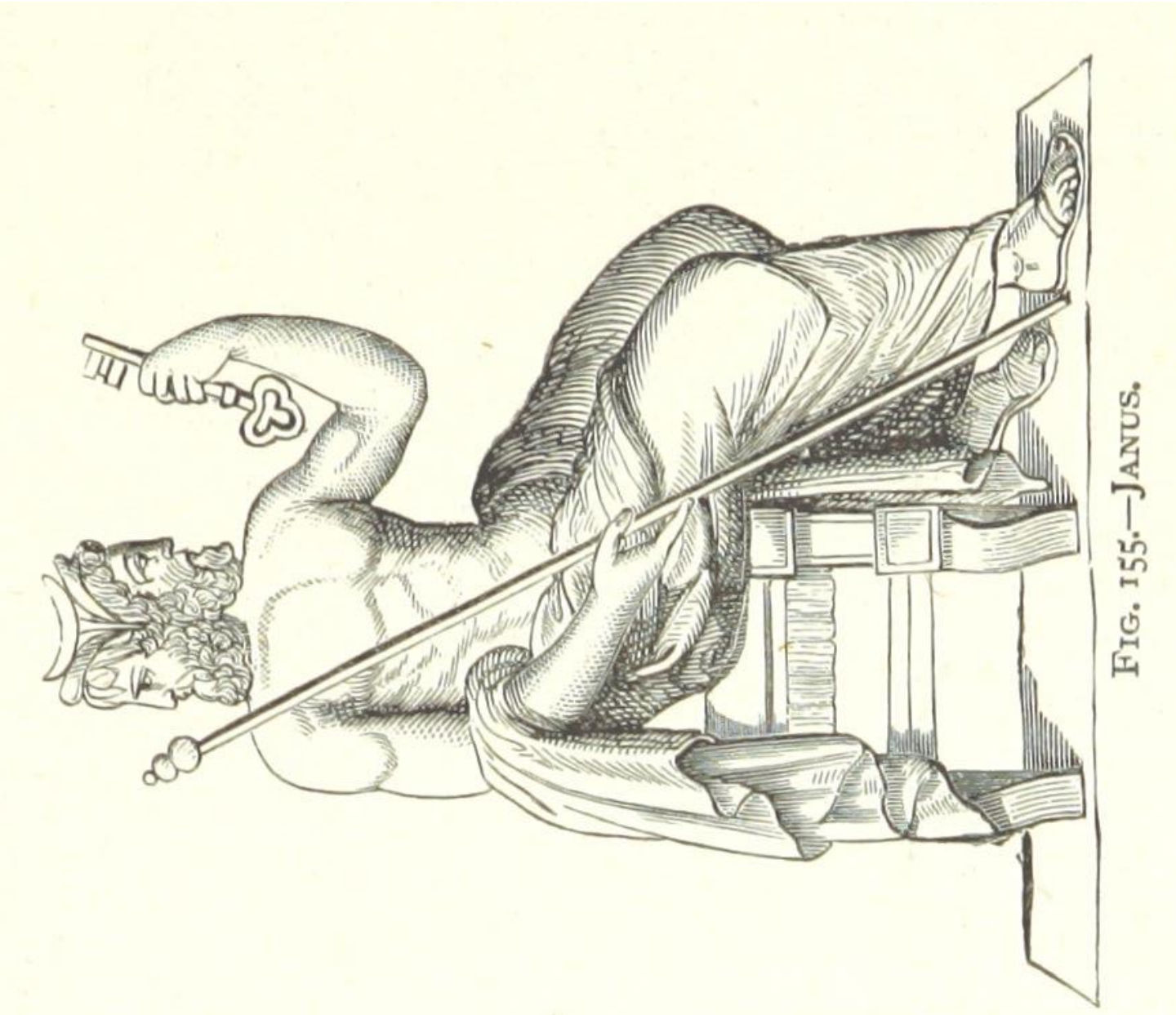


FIG. 155.—JANUS.

BRITISH LIBRARY

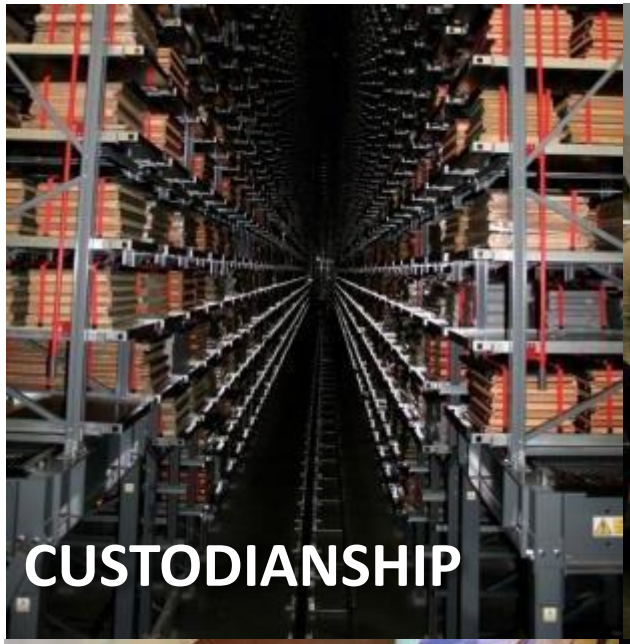
British Library Labs

Experiment with our digital collections

---

http://labs.bl.uk

Living Knowledge: The British Library 2015 – 2023



'We make our intellectual heritage accessible to everyone, for research, inspiration and enjoyment'

'We make our intellectual heritage accessible to everyone, for research, inspiration and enjoyment'

Ok, but how?

Internal digital scholarship training

Evaluating the training programme: goals

- Staff across all collection areas are familiar and conversant with the foundational concepts, methods and tools of digital scholarship.
- Staff are empowered to innovate.
- Collaborative digital initiatives flourish across subject areas within the Library as well as externally.
- Our internal capacity for training and skill-sharing in digital scholarship are a shared responsibility across the Library.

Participants appreciated…

- Hands-on, practical exercises
- Time to explore innovative digital projects
- Trying new tools, particularly with BL or similar collections
- The expertise and enthusiasm of instructors
- Meeting colleagues and learning about BL projects

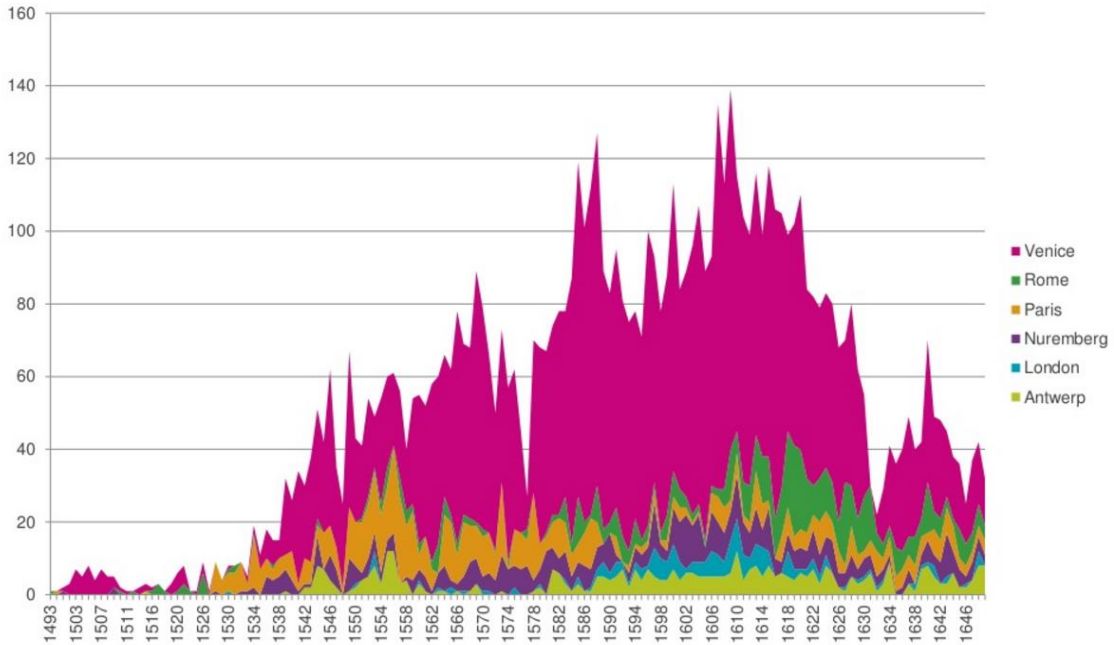
Practical digital training tips

- Case studies / real world examples help
- Articulate learning outcomes, expected results for exercises
- Provide clear, printable instructions
- No more than 15 people in hands-on courses
- Allow time for all to complete exercises
- Build in optional activities for advanced participants

Inspired... Big Data History of Music

How can vast amounts of bibliographic data held by research libraries be unlocked for music researchers to analyse?

Analyses and visualisations of these datasets exposed previously uncharted patterns in the history of music.



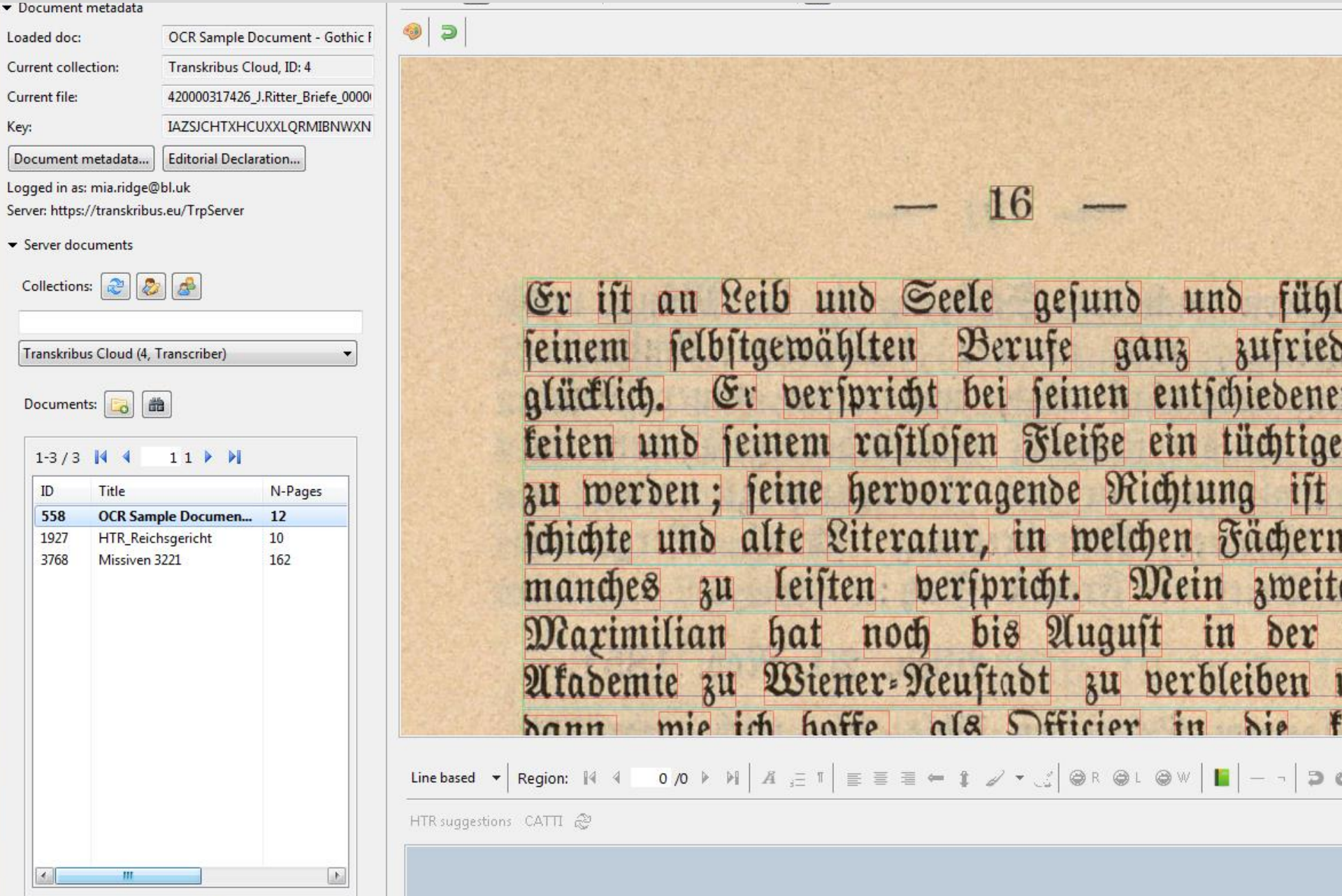
On-going challenges

- Applying new skills within the Library - more guidance and support within current Library policy and infrastructure
- Reaching staff who work on rosters and cannot attend for a whole day
- Digital tools for non-Western materials; working with non-textual digital collections such as the UK Web and Sound Archives

Challenge: the pace of change

- Software tools break/change - some courses need to be completely checked before they're run
- Digital Curators need to maintain our skills and knowledge too
  - Digital Scholarship Reading Group - 1 hour, once a month
  - Hack and Yack - 2 hours, once a month
  - Both are open to all

Transkribus for handwritten text recognition



Datasets and digital collections

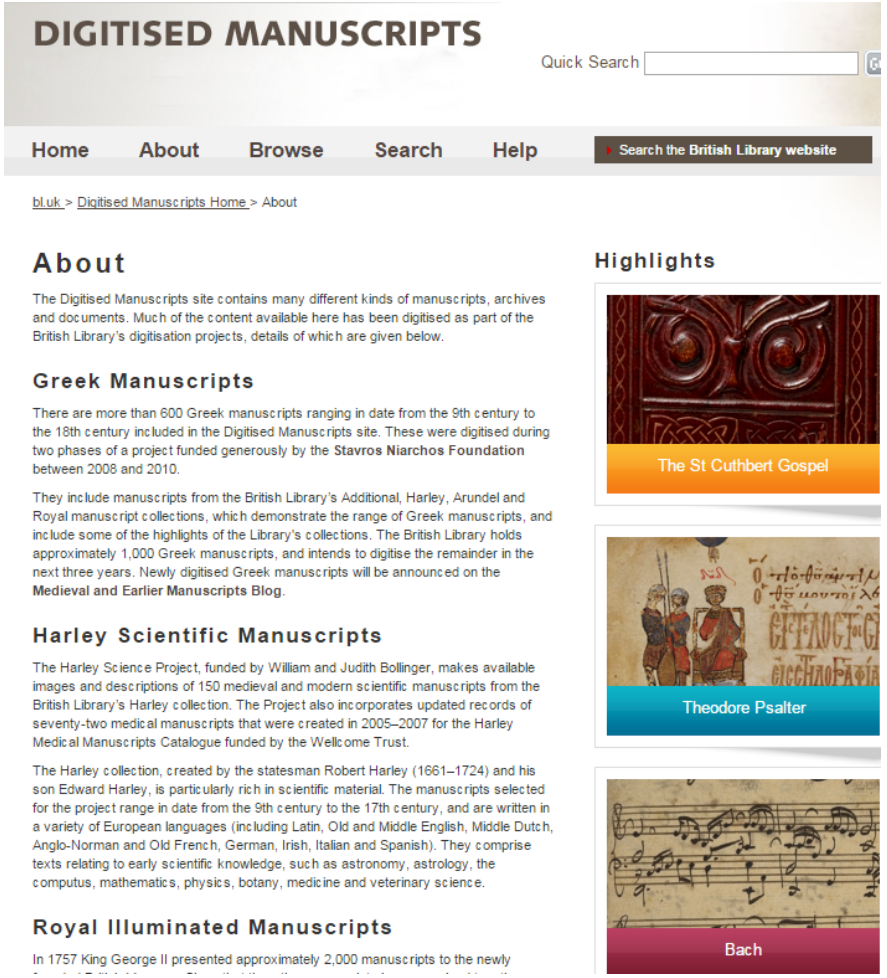
Datasets about our collections
Bibliographic datasets relating to our published and archival holdings

Datasets for content mining Content suitable for use in text and data mining research

Datasets for image analysis Image collections suitable for large-scale image-analysis-based research

Datasets from UK Web Archive Data and API services available for accessing UK Web Archive collections

Digital mapping Geospatial data, cartographic applications, digital aerial photography and scanned-in historic map materials



http://bl.uk/digital

BRITISH LIBRARY

COLLECTION METADATA
Data Services

Downloads

Bulk downloads, serialized as N-Triples, RDF/XML or CSV are available below. These may be downloaded as compressed files (ZIP format).

Files are distributed under a Creative Commons CC0 1.0 Universal Public Domain Dedication licence. For further details about our terms and conditions, please see here.

Linked Open Data

- **British National Bibliography**

Books and serials eligible for BNB. Models, schema and URI patterns available here. Updated monthly. Zipped folders include multiple files and a PDF document.

DATASET	DATE	SIZE (KB) *(full file)*	FULL FILE	SAMPLE
BNB LOD Books	2016-05	1,080,387	nt	nt
BNB LOD Serials	2016-05	41,936	nt	nt
BNB LOD Books	2016-05	1,102,362	rdfxml	rdfxml
BNB LOD Serials	2016-05	36,692	rdfxml	rdfxml
VoID Descriptions	2016-05	5	ttl	N/A

Search service: http://bnb.data.bl.uk
SPARQL editor: http://bnb.data.bl.uk/flint-sparql
SPARQL endpoint: http://bnb.data.bl.uk/sparql

Open Data: basic RDF/XML

Sections:

- **British National Bibliography**
- **Data Services**
  - Free data
  - Sample data
  - Downloads
- **Standards**
- **News**
- **Contact us**

This page contains links to Adobe PDF files. Accessibility solutions and free 'Reader' software are available from Adobe.

http://www.bl.uk/bibliographic/download.html

UKWA Open Data | Home | GitHub | UKWA

home

JISC UK WEB DOMAIN DATASET (1996-2013)

Introduction

Format Profile

Geoindex

Host-level Links

Crawled URL Index

UK SELECTIVE WEB ARCHIVE

Introduction

Website Classification Dataset

PROJECTS

Big UK Domain Data for the Arts and Humanities

Analytical Access to the Domain Dark Archive

Big Data: Demonstrating the Value of the UK Web Domain Dataset for Social Science Research

SCAPE: Scaleable Preservation Environments

RELATED WORK

Common Crawl

httparchive.org

UK Web Archive Open Data

In order to facilitate research, and so we might better understand and preserve the UK's web history, the UK Web Archive has decided to make a number of data and API services available for general use. We also make a few example tools available, showing how the open data might be used, and these are hosted in this GitHub repository.

We hope that by making these datasets available, the broader community will find interesting ways to re-use, explore and visualise the contents of our web archive. We are keen to work with any interested parties to exploit these datasets, and understand what other derivative or summary data would be of interest.

UK WEB ARCHIVE preserving uk websites

Datasets, Tools & APIs

Open Datasets

In general, we can't provide remote bulk access to the primary datasets listed above (although bulk access can be arranged for particular projects - see below). This is mainly because the conditions underwhich we hold the content do not permit it, but also because the data sets are very large and so providing bulk downloads is not practical.

However, secondary datasets, composed of metadata that describes simple facts about the content, can be made available under open terms.

For all the details, follow the links on the left, or look through the README files and code in the GitHub repository.

Tools

We also make a few tools available, which illustrate how the open datasets might be used. These are hosted on GitHub, and you should feel free to fork our repository or download our tools using the links above. Pull requests containing new or improved tools are welcome!

http://data.webarchive.org.uk/

By using this site, you agree we can set and use cookies. OK For more details of these cookies and how to disable them, see our cookie policy.

data.bl.uk

data.bl.uk

Welcome to the British Library's alpha open data portal.

As part of its work to open its data to wider use, the British Library is making copies of some of its datasets available for research and creative purposes.

We aim to describe collections in terms of their data format (images, full text, metadata, etc), licences, temporal and geographic scope, originating purpose (e.g. specific digitisation projects or exhibitions) and collection, and related subjects or themes.

This site is in the early stages of development. If you have questions or feedback about this site or our open data work, please email digitalresearch@bl.uk. We'd also love to hear what you've done or made with the data.

Each dataset may have its own licencing conditions and re-use stipulations, so please check their pages for more information.

Ask me for a demo!

A collection of datasets released by the British Library

portraitsviewsplaybills - comprises 2 datasets.

india-office-records - comprises 1 dataset.

taverns-signs - comprises 3 datasets.

portraits-views-playbills - comprises 1 dataset.

advertisements - comprises 6 datasets.

dig19cbooks - comprises 13 datasets.

Contact us: digitalresearch@bl.uk @BL_DigiSchol @mia_out @Miss_Wisdom

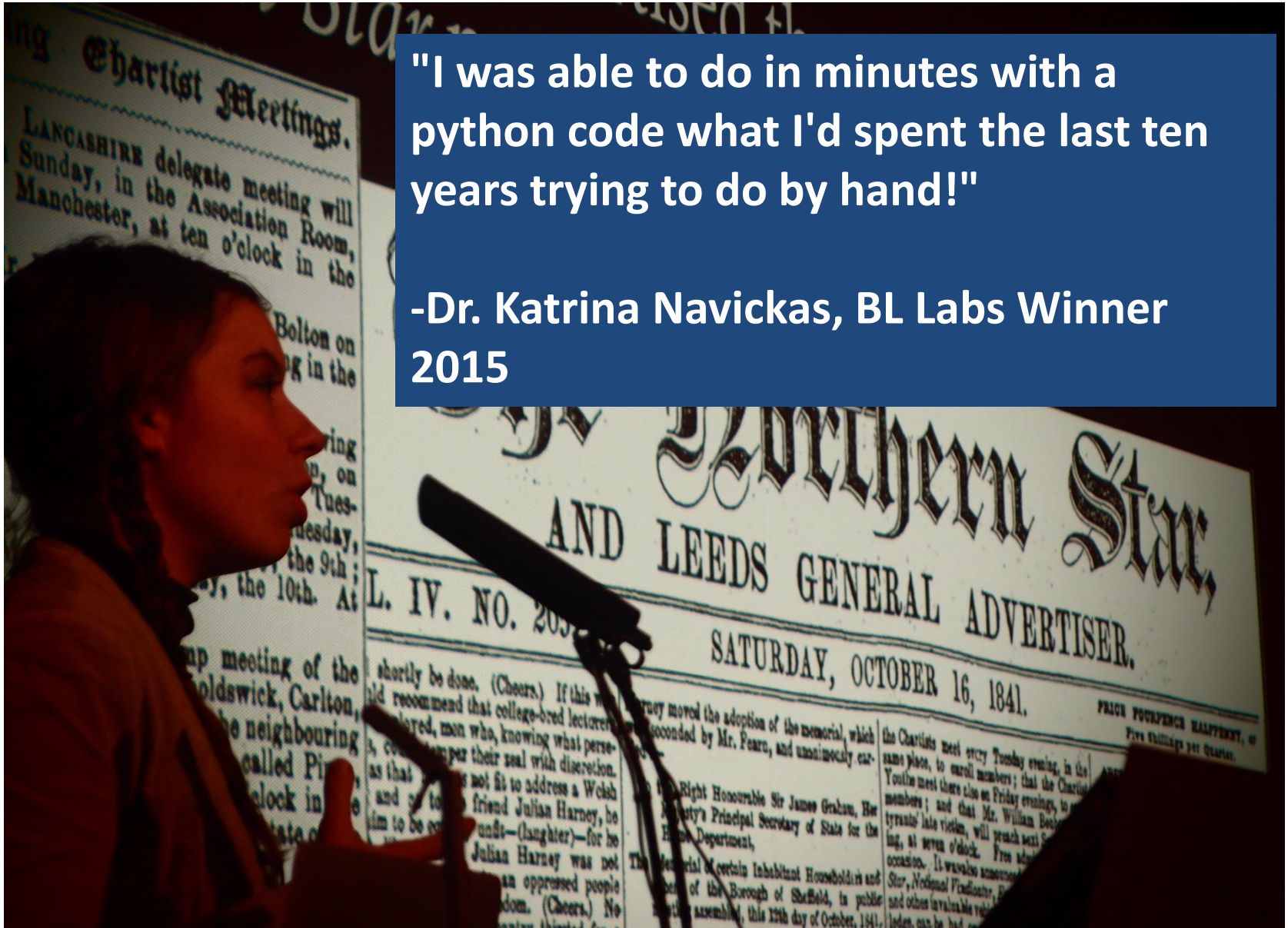
BL Labs Competition

BL Labs Competition finalists for 2016

- **Black Abolitionist Performances and their Presence in Britain**
- *Hannah-Rose Murray, PhD student at the University of Nottingham*

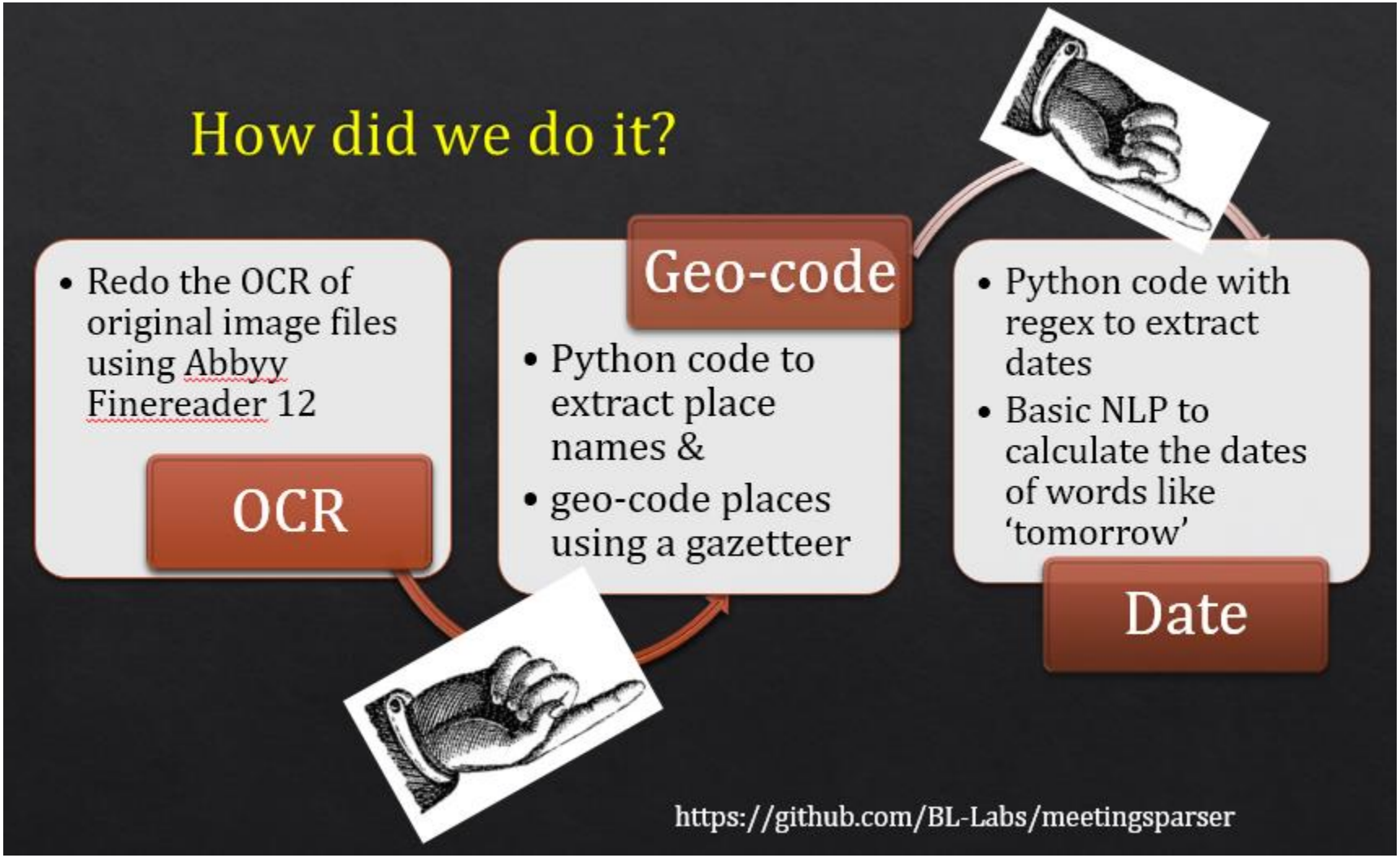
- **SherlockNet: Using Convolutional Neural Networks to automatically tag and caption the British Library Flickr collection**
- *Luda Zhao, Brian Do and Karen Wang, students from Stanford University, California*

Political Meetings Mapper



Political Meetings Mapper

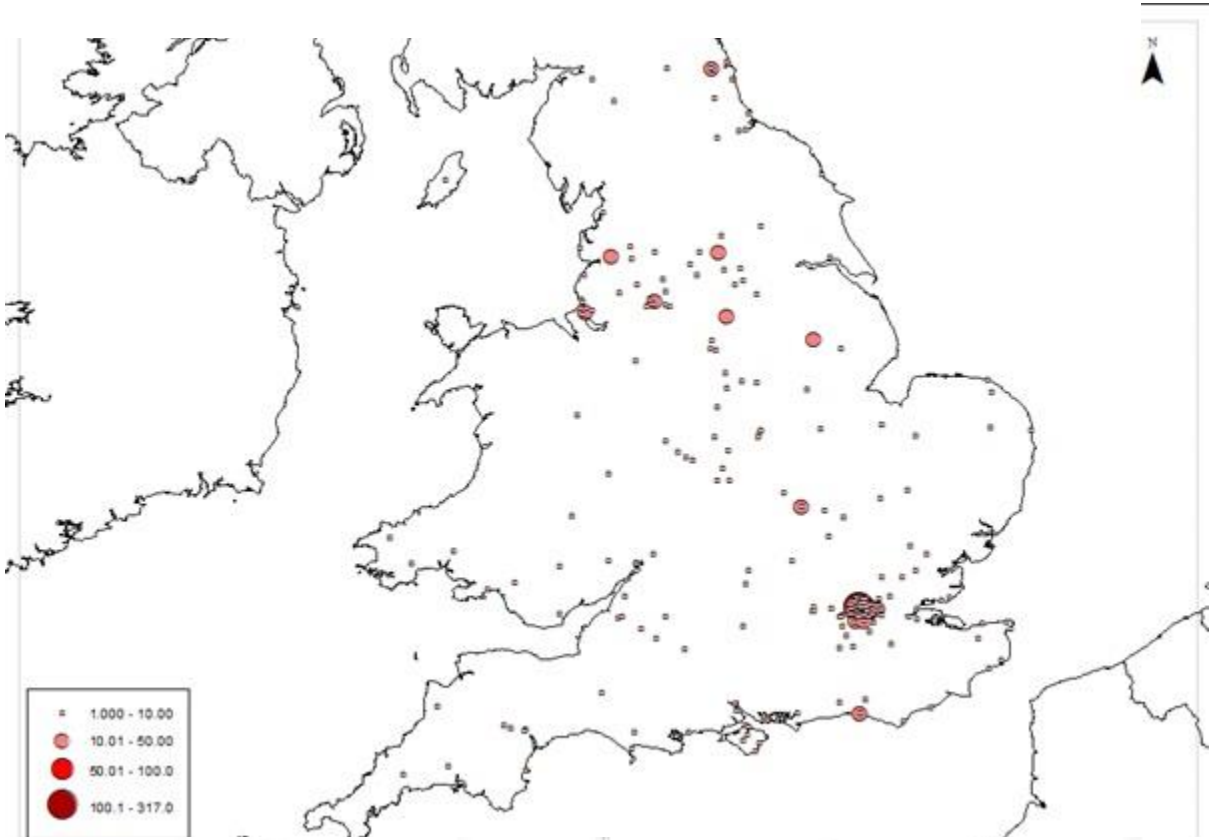
Video: https://youtu.be/XabsuyNkD5s



Spatial Humanities

‘Being able to link the map and the underlying text allows us to understand how patterns vary from place to place.’

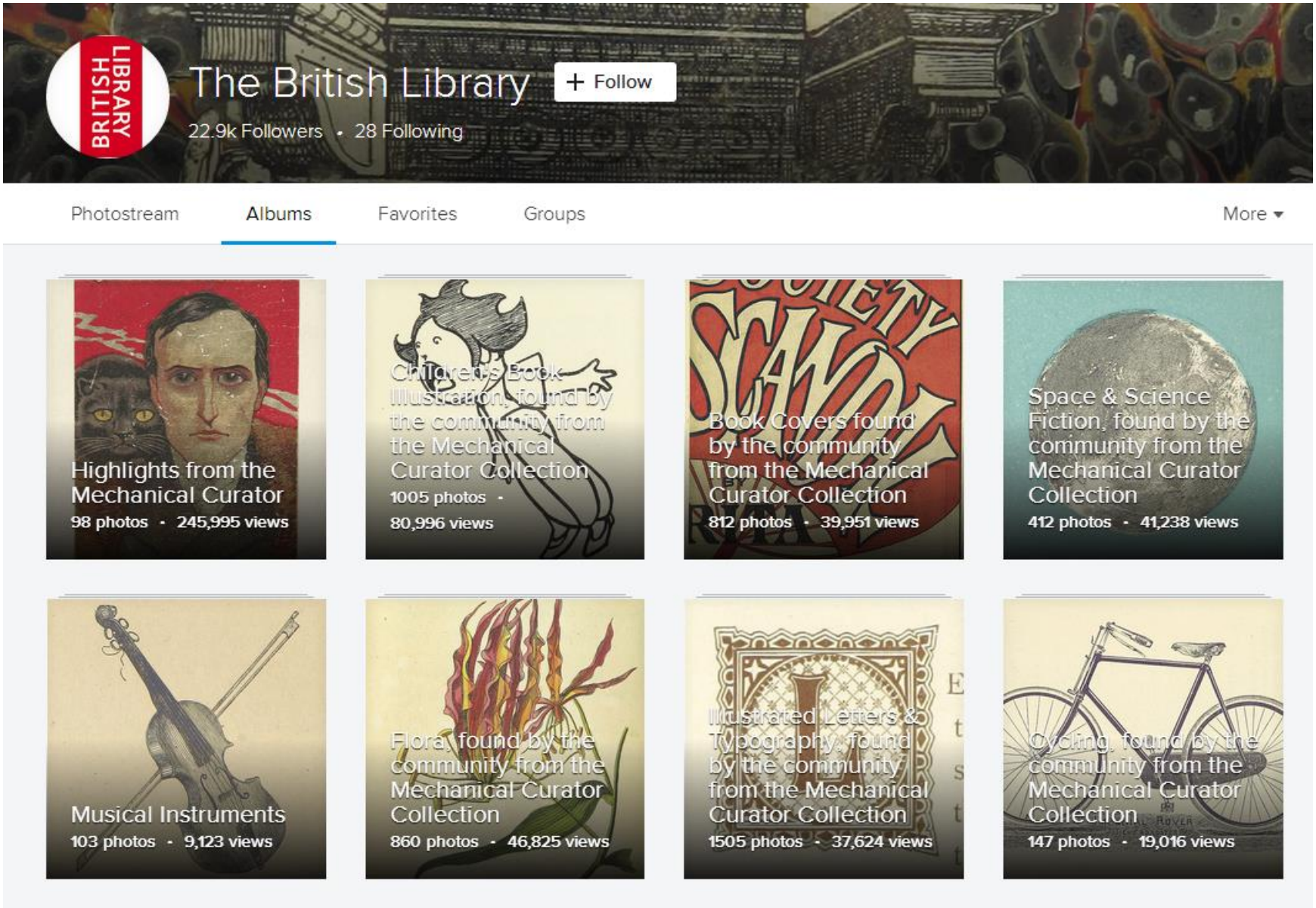
Combining Text Analysis and Geographic Information Systems to investigate the representation of disease in nineteenth-century newspapers, Lancaster



BL Labs Bi-Annual Competition & Awards

The 2016 British Library Labs Awards deadline is **5 September**.

'A million images' on Flickr Commons



Mario Klingemann
'556 Minerals'

'Grassroots' projects

Welcome to FreeBMD.

FreeBMD is an ongoing project, the aim of which is to transcribe the Civil Registration index of births, marriages and deaths for England and Wales, and to provide free Internet access to the transcribed records. It is a part of the FreeUKGEN family, which also includes FreeCEN (Census data) and FreeREG (Parish Registers). To search the records that have so far been transcribed by FreeBMD click on the **Search** button below.

The recording of births, marriages and deaths was started in 1837 and is one of the most significant resources for genealogical research. The transcribing of the records is carried out by teams of dedicated volunteers and contains index information for the period 1837-1983, **BUT WE HAVE NOT YET TRANSCRIBED THE WHOLE PERIOD.** A breakdown by event and year can be viewed here.

FreeBMD is exactly that - FREE. We do not make any charge whatsoever for use of the site. FreeBMD is a registered charity and our objective is to provide free online access to the GRO Index. However, FreeBMD costs money to run and if you would like to make a donation by PayPal click on the button to the left. For other methods of making donations to FreeBMD please see here.

Search View Images Information Join FreeBMD Transcribers' Page

The FreeBMD Database was last updated on Thu 30 Oct 2014 and currently contains 242,129,588 distinct records (308,090,800 total records).
On Mon 17 Nov 2014 FreeBMD users did 175,397 searches. (More information)

Search engine, layout and database Copyright © 1998-2013 The Trustees of Free UK Genealogy (Ben Laurie, Graham Hart, Camilla von Massenbach, David Mayall and Jim Killock), a charity registered in England and Wales, Number 1096940.
We make no warranty whatsoever as to the accuracy or completeness of the FreeBMD data.

Oxfordshire Family History Society

Transcribed Wills

This area of the OFHS web site contains the transcribed texts of wills, administrations and other testamentary documents, relating to Oxfordshire, which have been generously donated by their transcribers. Do you have any that you would be prepared to donate? If so, please click **here** for contact details.

Testator Index	Index of Persons Mentioned
A - Z	A B C D - F G H I - K L M - O P Q - R S T U - W X - Z ?

Search for Surnames matching a Specific Pattern Search

The search pattern may be a single surname, or may use a **Regular Expression** to match a set of alternative spellings for the surname. If you are not familiar with the concept of Regular Expressions, click here for an explanation.

For Help Using These Pages

- See Notes for Users if you are a first time visitor to these pages.
- See Notes for Transcribers if you have any transcripts which you would be prepared to donate.
- See Glossary for an explanation of some unusual words found in wills.
- See Regular Expressions to learn about patterns which may be used to search for surnames.
- See Probate on the GENUKI web site for general background information on finding wills.
- See Oxford Wills before 1858 for specific information about Oxfordshire wills.

Location of Original Documents

The majority of the transcripts presented here will be from originals held in one of two places, as indicated in the heading of the transcript:

- **ORO** or **OHC** - designates Oxfordshire Record Office (ORO) or Oxfordshire History Centre (OHC) where you will find wills proved locally, when the deceased's property was all within the diocese. During 2011, Oxfordshire Record Office was renamed Oxfordshire History Centre and the web address changed.
  The link here will direct you to their new web site at: http://www.oxfordshire.gov.uk/oxfordshirehistory
  Scans of the original probate documents held at Oxfordshire History Centre may be purchased directly from OHC via the Staff-operated copying service on their web site.

- Scans of the original probate documents held at Oxfordshire History Centre, are also available from Findmypast as part of their collection of England & Wales Published Wills & Probate

Participatory history as learning

- Constructive environments for learning and mastering skills
  - Palaeography
  - Tagging / classification
  - Source familiarity
  - Biographical research
  - Database and web design
  - Data collation, analysis, interpretation
  - Extended research projects

MarineLives digital pop up lab

Opens on July 4th 2016, two hours a week for ten weeks

Spaces for twelve volunteers

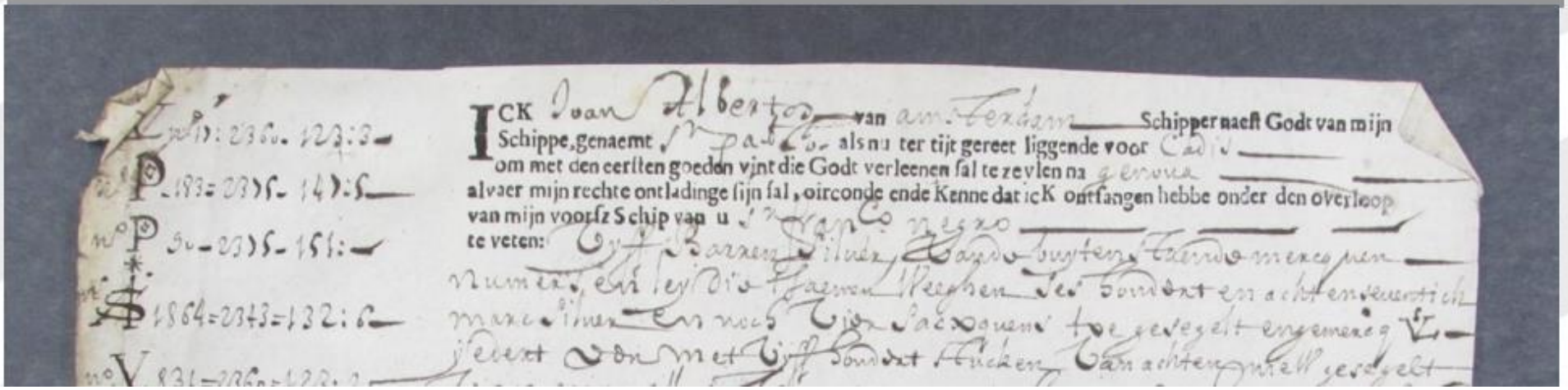
Collaborative experience for

historians, coders and computer scientists

Team one: working on semi-automated recognition of handwritten manuscripts

Team two: working on tailored search of historical documents

Team three: working on visualisation of historical data



@marinelivesorg http://digitalpopuplab.org/

Current most popular features

Each of the following three features (*a feature comprises several web pages*) can be accessed directly by clicking the text or images below:

Liz Knight's
latest researches:
A series featuring lacemaking

Tour of the
Market Place: 1800-1950s
Check out the video

Tom Garner's
WW1 Diary of his Service
on the Eastern Front

Link to full content of Features

Current most popular articles

Each of the following three articles can be accessed directly by clicking the text or images below:

LMS Railway at Olney
(Northampton-Bedford)
includes a video

Orchard House Nos 67 & 69 High Street
converted in 1904
for Joseph William Mann

Oliver Ratcliff's 1907
Book 'Olney, Bucks'
('Turning page' version)

Conclusion

Kiitos / Thank you!

Mia Ridge @mia_out
Digital Curator, British Library
digitalresearch@bl.uk @BL_DigiSchol