

British Library Labs: Competition winners

Find out more at <http://labs.bl.uk>

The [Labs Competition](#) looks for **transformative project ideas** which use the British Library's [digital collections and data](#) in new and exciting ways. The project is supported by the [Andrew W. Mellon foundation](#).



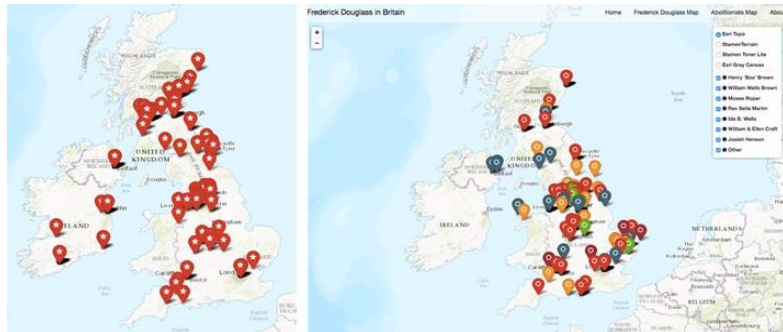
The 2016 Competition winners

Black Abolitionists and their presence in Britain

Hannah-Rose Murray, PhD student at the University of Nottingham

Maps from Hannah-Rose's website

The project focuses on African American lives, lectures and experiences in Britain between 1830–1895. By assessing Black Abolitionist speeches in the British Library's Nineteenth Century Newspaper Collection and using the British Library's Flickr Commons collection to illustrate, the project has illuminated their performances and how their lectures reached nearly every corner of Britain. For the first time, the location of these meetings has been mapped and the number and scale of the lectures given by black abolitionists in Britain has been evaluated, allowing their hidden voices to be heard and building a more complete picture of Victorian London for us. The project findings can be found on Hannah-Rose's website www.frederickdouglassinbritain.com.



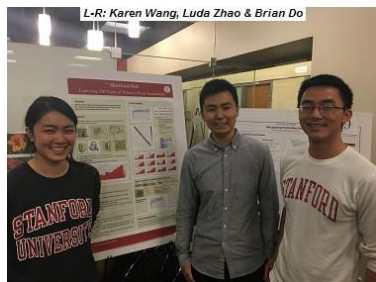
Hannah-Rose Murray is a PhD student with the Department of American and Canadian Studies, University of Nottingham. Her AHRC/M3C-funded PhD focuses on the legacy of formerly enslaved African Americans on British society and the different ways they fought British racism. Hannah-Rose received a first class Masters degree in Public History from Royal Holloway University and has a BA History degree from University College London (UCL). In Nottingham, Hannah-Rose works closely with the Centre for Research in Race and Rights and is one of the postgraduate directors of the Rights and Justice Research Priority Area, which includes the largest number of scholars (700) in the world working on rights and justice.

SherlockNet: Using Convolutional Neural Networks to automatically tag and caption the British Library Flickr collection

Karen Wang and Luda Zhao, Masters students at Stanford University, and Brian Do, Harvard Medicine MD/PhD student

SherlockNet web interface

Machine learning can extract information and insights from data on a massive scale. The project developed and optimised Convolutional Neural Networks (CNN), inspired by biological neural networks in the brain, in order to tag and caption the British Library's Flickr 1 million collection. In the first step of the project, images were classified with general categorical tags (e.g. "maps", "architecture"). This served as the basis for the development of new ways to facilitate rapid online tagging with user-defined sets of tags. In the second stage, automatically generate descriptive natural-language captions were provided for images (e.g. "A man in a meadow on a horse"). This computationally guided approach has produced automatic pattern recognition which provides a more intuitive way for researchers to discover and use images. The tags and captions are being made accessible and searchable for the public through a web-based interface and text annotations will be used to globally analyse trends in the Flickr collection over time. Visit the website: <http://bit.ly/sherlocknet>



L-R: Karen Wang, Luda Zhao & Brian Do

Karen Wang is a senior studying Computer Science at Stanford University, California and she also has an Art Practice minor. Karen is interested in the intersection of computer science and humanities research, so this project is near and dear to her heart! She will be continuing her future studies at Stanford in CS, Artificial Intelligence track.

Luda Zhao is currently a Masters student studying Computer Science at Stanford University, living in Palo Alto, California. He is interested in using machine learning and data mining to tackle tough problems in a variety of real-life contexts, and he's excited to work with the British Library to make art more discoverable for people everywhere.

Brian Do grew up in sunny California and is a first-year MD/PhD student at Harvard Medical School. Previously he studied Computer Science and biology at Stanford. Brian loves using data visualisation and cutting edge tools to reveal unexpected things about sports, finance and even his own text message history.

The 2015 Competition winners



Crowdsourcing Arcade: Repurposing the 1980s arcade console for image classification

<http://goo.gl/Nfg9D5>

Dr Adam Crymble, Lecturer of Digital History at the University of Hertfordshire

The '[Crowdsourcing Arcade](#)' experimented with [crowdsourcing](#) the tagging of the British Library's [Flickr Commons images](#), through the use of a specially built 1980s-style [arcade game machine](#) (see image to the left) installed with [games](#) (see image to the right) to help with the tagging. Inspired by the 'maker' community and physical computing, this project takes the crowdsourcing experience off the web and puts it into a replica machine, replete with joysticks and plastic shiny buttons. This old interface put to new uses acknowledges that people increasingly associate their computers with work, and by providing a digital experience that doesn't feel like a computer, we can tap into energy currently reserved for play.



The 'Art Treachery' image tagging game developed by Janus Druz. The 'Art thief' has to use their torch to find an art piece they have been asked to steal from an art gallery whilst being chased by robot guards.



Adam Crymble specialises in how technology can change the way we research and present the past. His research focuses on 18th and 19th century social and cultural history, especially on the ways migrants and minority groups were represented in various forms of media. In 2015, he was shortlisted by the British Broadcasting Corporation (BBC) and Arts and Humanities Research Council (AHRC) for the New Generation Thinkers programme. He is a founding editor of *The Programming Historian* (<http://programminghistorian.org>), a suite of tutorials teaching humanities scholars to take advantage of digital methods in their research.

Political Meetings Mapper: Bringing British Library maps to life with the history of popular protest

<http://politicalmeetingsmapper.co.uk>

Dr Katrina Navickas, Senior Lecturer of History at the University of Hertfordshire

[Political Meetings Mapper](#) used [text mining and geo-location](#) to find records of when and where political meetings took place in the Library's digitised newspapers, enabling anyone to access them on maps and explore the data on an [interactive website](#). The tool visualises the locations of political events in the crucial era of the 1830s and 1840s, when Chartism, the first and largest movement for democracy in Britain, held thousands of meetings and demonstrations to campaign for the vote. By plotting the meetings listed in the Chartist newspaper, *The Northern Star*, from 1838 to 1844, it discovered new spatial patterns in which popular politics happened, and in so doing, has helped to answer the questions of how and why it happened. This project showcases the British Library's collections and combines them in a way not done before: the geo-referenced maps in the BL geo-referencer and Flickr Commons, the Ordnance Surveyors' drawings, and the 19th century newspaper collection.



Political Meetings Mapper. Red Flags show Chartists' meetings that took place in Bradford.



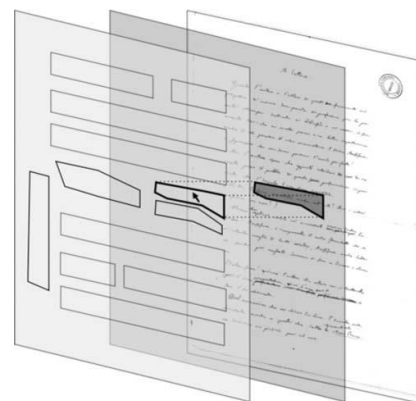
Katrina Navickas has research interests in the history of popular protest and democratic movements in 18th and 19th century Britain. She is currently developing her digital skills in applying Geographic Information Systems (GIS) and Space Syntax methods to historical research. Katrina published a book about her research in 2015: *Chartists, Protest and the Politics of Space and Place, 1789–1848*, Manchester University Press (<http://goo.gl/OSsfkO>).

The 2014 Competition winners

Text to Image Linking Tool (TILT) <http://bltilt.blogspot.co.uk/>

Desmond Schmidt and Anna Gerber

TILT is a tool linking digitised handwritten manuscripts to transcribed texts. In order to make old printed books and manuscripts accessible to a Web audience, it is essential to display the page image / facsimile of the original document next to its transcription. This allows the user to comment on the text, and to read it clearly, but because original documents are often hard to read, or have different line-breaks than text on a computer screen, it is easy to get lost trying to match up words in the document with words in the transcription. To overcome this, the team are developing semi-automatic methods to generate links that highlight corresponding parts of the page image and the text.



Visualising manuscript regions to enable linking to transcriptions



Desmond Schmidt has degrees in classical Greek papyrology from the University of Cambridge, UK, and in Information Technology from the University of Queensland, Australia. He has worked in the software industry, in information security, on the Vienna Edition of Ludwig Wittgenstein, on Leximancer, a concept-mining tool and on the AustESE (Australian Electronic Scholarly Editing) project. He is currently a Research Scientist at the Institute for Future Environments, Queensland University of Technology.



Anna Gerber is a software developer. She was a technical project manager specialising in Digital Humanities projects at the University of Queensland's ITEE (Information Technology and Electrical Engineering) eResearch group. Anna was the senior software engineer for the AustESE project, developing eResearch tools to support the collaborative authoring and management of electronic scholarly editions. She is a contributor to the W3C (World Wide Web) Community Group for Open Annotation and was a co-principal investigator on the Open Annotation Collaboration project.

Victorian Meme Machine <http://victorianhumour.tumblr.com>

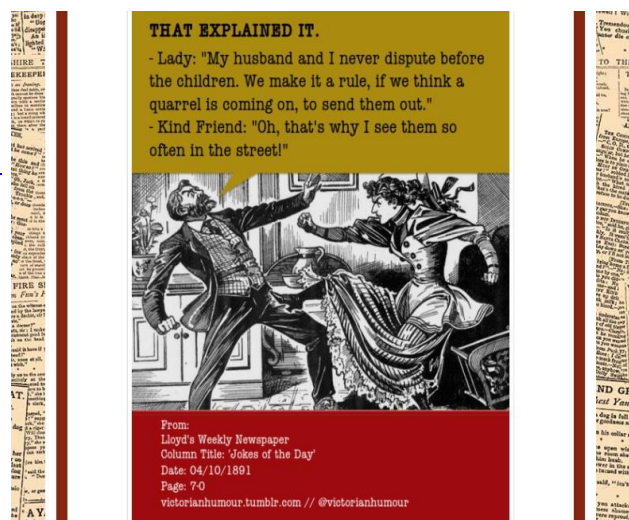
Dr Bob Nicholson, Senior Lecturer at Edge Hill University

Victorian Meme Machine (2014) created a [database of Victorian jokes](#) from the Library's digital archives and attempted to revive this humour through [social media such as twitter](#) and via the ['Mechanical Comedian'](#).

What would it take to make a Victorian joke funny again? While the great works of Victorian art and literature have been preserved and celebrated by successive generations, even the period's most popular jokes have now been lost or forgotten. Fortunately, thousands of these endangered jests have been preserved within the British Library's digital collections.

Original Joke	OCR
<p>Mrs. Talbot: "I see by the paper, May, that Mrs. Folly is still dangerously ill, but insists upon having two wings added to her house before she dies." May Talbot: "Good idea, ma, judging from all reports, they are the only pair she will ever have."</p>	<p>.e-Mrs. Talbot: ".I see by the paper, May, thatal ;r r,Folly is still dangerously ill, but <u>insist</u>s no my upon having two wings added to her house before' Fr she dies." let ard May Talbov: "Good idea, ma, judging from, <u>hal</u> ilug all reports, they are the only pair sile will ever <u>pu uns</u>, have."</p>

An example of joke text extracted from the OCR of a digitised scan of a newspaper



The 'Mechanical Comedian' posts a Joke every day

This project aimed to find these forgotten jokes and bring them back to life. Bob created an online transcription platform which cleaned up text about jokes, extracted originally through the digitisation of newspapers and books and Optical Character Recognition (OCR). These were then put into a database. Bob published a detailed [article](#) (2015) in *The Journal of Interdisciplinary Studies in the Long Nineteenth Century* about his work.



Bob Nicholson is a senior lecturer in history specialising in nineteenth-century Britain and America, focusing on journalism, popular culture, jokes, and transatlantic relations. Bob has been exploring representations of the United States, and the circulation of its popular culture in Victorian newspapers and periodicals. He is a keen proponent of the Digital Humanities and has written for The Guardian, had his research covered by The Times, and was shortlisted by the British Broadcasting Corporation (BBC) and Arts and Humanities Research Council (AHRC) in their first search for New Generation Thinkers (2011).

The 2013 Competition winners

Mixing the Library: The Disc Jockey and the Digital Collection <http://www.ablab.org/>

Dan Norton, Artist and researcher at The Computer Vision Centre (CVC) group on Visual Interaction (VI-CVC) in Barcelona, Spain

The project developed a prototype tool which applied the intuitions of a DJ on working with and mixing the British Library's digital collections. The work was built on two interface features of: "continual visual presence" of the collection as a writeable menu system; and a "mixing screen" (see the diagram on the right) that allows two (or more) data fragments from the collection to be presented, combined, and linked and later evaluated and extended with semantic links. It is hoped that the tool will operate as a powerful scholarly interface for learning in digital collections, developing material, and sharing annotated semantically rich data.



Dan Norton's art practice integrates interface technologies with digital and non-digital archives, and his work has been shown in film festivals, technology conferences, fine art and media art forums, as well as music festival. Currently he is working with Research Arts (Berlin) and Computer Vision Centre (Barcelona) to further develop the platform he worked on at the British Library. The project integrates art practice

(installations and workshops), software engineering, and architectural development of the library space, exploring the transforming use of the public library in the face of digital content.

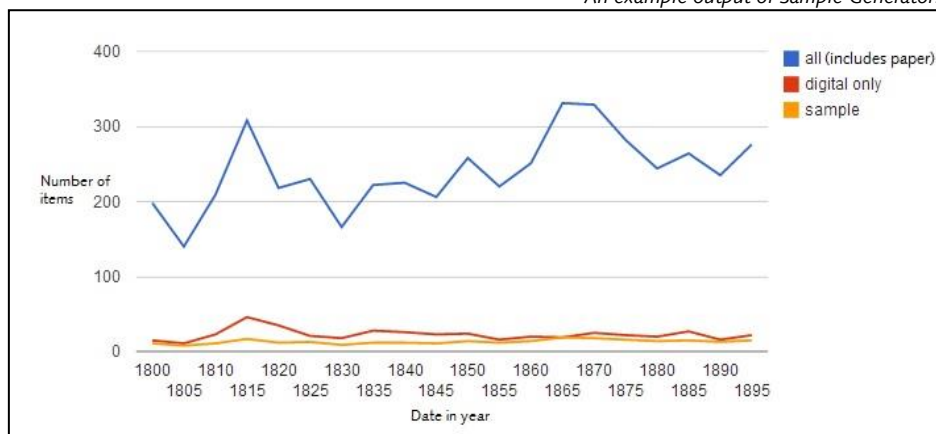


a Library's digital collections, entitled 'Memory Fields'.

The Sample Generator for Digitised Texts <http://goo.gl/bR9UJL>

Pieter Francois, Senior Lecturer at the University of Hertfordshire and Postdoctoral Researcher at the University of Oxford

The Sample Generator provides users with a way of exploring large (digitised) text collections through the use of unbiased structured sampling. It creates samples from large digital collections which are representative of even larger or 'complete' paper collections. It works on the metadata of a large digital / paper collection (e.g. year and place of publication, name of author, title words, gender of author, etc.). The larger paper collection is usually considered to be as representative as possible of the overall publication landscape.



The Sample Generator attempts to remove the 'bias' introduced during the digitisation process and replaces it with the bias of the overall paper collection which is usually much better understood. Claims made on the basis of analysing the sample can be linked to the paper collection/overall publication landscape and are no longer linked to the digital collection. For example, it is possible to create a sample of British travel accounts which takes into account the rises and declines of the genre as is shown in the overall paper collection. All samples are stored, can be shared with peers and linked to in publications. The tool allows you to explore your data quickly and also functions as a hypothesis generating and testing tool.



Pieter Francois specialises in the longitudinal analysis of archaeological and historical data. The main focus of his research is on how ritual, warfare and religion contributed in different ways to the evolution of social complexity. Together with Professor Harvey Whitehouse (University of Oxford), Professor Peter Turchin (University of Connecticut) and Professor Edward Slingerland (University of British Columbia), he co-founded the international research project 'Seshat-Global History Databank' (<http://evolution-institute.org/seshat>) which is currently funded by two multi-million pound grants from the ESRC (UK) and SSHRC (Canada). He trained as a nineteenth-century historian and has researched and published extensively on the various forms of travel and migration between Britain and the Continent.