Building better ‘GLAM Labs’

Experiences and lessons learned from the British Library and around the world with Galleries, Libraries, Archives and Museums engaging with researchers, artists, educators and entrepreneurs who want to use digitised and born digital cultural heritage collections and data for innovative projects.

Wednesday 27 February 2019, 1330 – 1500 (Keynote)

Talk given on behalf of the British Columbia Research Libraries Group, in the McPherson Library/Mearns Centre for Learning, Digital Scholarship Commons, Room A308, University of Victoria, Victoria, British Columbia, Canada
The British Library or ‘BL’

St Pancras, London, UK
Many books are stored 4 stories below the building
UK Legal Deposit Library – Reference only

Inside the British Library
Space for 1200 readers, around 500,000 visitors per year

Stockton-on-Tees
Author right to payment each time their books are borrowed from public libraries

Many items stored at Document Supply and Storage centre 48 hours away
Building 37 uses low oxygen and robots
Reading room and delivery to London

https://goo.gl/tVwqHF
Roly Keating (Chief Executive Officer of the British Library)

Living Knowledge Vision (2015 – 2023)

Custodianship
We build, curate and preserve the UK’s national collection of published, written and digital content

Research
We support and stimulate research of all kinds

Business
We help businesses to innovate and grow

To make our intellectual heritage accessible to everyone, for research, inspiration and enjoyment and be the most open, creative and innovative institution of its kind by 2023 (50 year anniversary).

Culture
We engage everyone with memorable cultural experiences

Learning
We inspire young people and learners of all ages

International
We work with partners around the world to advance knowledge and mutual understanding

Document: http://goo.gl/h41wW7
Speech: https://goo.gl/Py9uHK
Collections – not just books!

> 180* million items
> 14* m books
> 60* m patents
> 8* m stamps
> 4* m maps
> 6* m sound recordings
> 1.6* m musical scores
> 0.3* m manuscripts
> 0.8* m serial titles

*Estimates

King’s Library

https://goo.gl/tVwqHF
Have you got X?

Looking for Physical Content in the British Library

https://upload.wikimedia.org/wikipedia/commons/5/50/Real_wuerzburg.jpg

https://goo.gl/tVwqHF
What percentage/proportion of our physical collections are digitised?

3%* digitised

Partnerships
Commercial & Other Organisations

Research driven digitisation
Heritage Made Digital

Digitisation / Curating Born Digital
Costs money, time, resources

Bias in digitisation
Over 720 Digital collections
15%* Openly Licensed – most online
85%* Available onsite only at the moment

What percentage/proportion of our physical collections are digitised?

* estimate

http://www.turing.ac.uk
https://www.turing.ac.uk/research/research-projects/living-machines

@BL_Labs mahendra_mahey @uvic @uviclib @britishlibrary labs@bl.uk

https://goo.gl/tVwqHF
Digital access and reuse

• All Libraries need a process for agreeing terms of access to content

• Many competing concerns
  – Re-use
  – Open research
  – Copyright
  – Licensing
  – Ethics
  – Revenue

• Large collection of books digitised by funding through Microsoft an early win for us in 2012 (More later about this collection)
The Story of the Digital Collection…

Good to know the background ‘story’ of a Digital Collection if you want to use it for projects…
Challenges of access to Digital Collections at the BL

Labs Residency Model

Competition / Digital Research Support Application

Onsite @
British Library

READING ROOM

NOT ONLINE

OPEN

£

Over 720 Digital collections
15 %* Openly Licensed – most online
85 %* Available onsite only at the moment
Have you got X?

Looking for Physical Content in the British Library

https://upload.wikimedia.org/wikipedia/commons/5/50/Real_wuerzburg.jpg

https://goo.gl/tVwqHF
Have you got X digitised / in digital form?

Looking for Digitised / Digital Content in the BL

Finding Open British Library Cultural Heritage Datasets

Collection Guides (234 as of 27/02/2019)
https://www.bl.uk/collection-guides/

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

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

https://data.bl.uk
Download collections as zips, no API
Each dataset has a Digital Object Identifier (DOI) can be referenced for research
Over 120 datasets available

https://goo.gl/tVwqHF
British Library Digital collections & Datasets

Broadcast News (TV and Radio)

Music (Recordings & Sheet) & Sounds

Playbills, Books, Newspapers (includes OCR)

Images, Manuscripts & Maps

British National Bibliography

Usage data

Qatar Digital Library

International Dunhuang Project

Hebrew Manuscripts

Images, Manuscripts & Maps

Flickr & Wikimedia Commons

http://www.qdl.qa/
http://dml.city.ac.uk/
http://idp.bl.uk/
http://www.bl.uk/maps/
http://www.bl.uk/maps/
http://dml.city.ac.uk/
http://goo.gl/frSMJt
http://goo.gl/tVwqHF
What engagement does the BL have with people wanting use our digital content?

#bldigital

3 %* digitised

• Dialogue typically:
  – ‘You are in luck’, we have what you are looking for!
  – ‘You are in not luck’ but we have this instead…
  – Engagement **is** constantly required to maintain interest in our digital collections. **No engagement no Lab!**

• Tend to attract projects with ‘fuzzier’ boundaries

• Labs is open to more flexible, interdisciplinary / collaborative research

• Artists / Creatives often find engagement with our digital collections easier than scholars who often want a specific thing…

https://goo.gl/tVwqHF

https://goo.gl/qpCLlk

https://goo.gl/wMTS3Z
The British Library's Digital Scholarship team

Our mission is to enable the use of the British Library’s digital collections for research, inspiration, creativity, and enjoyment.

- Digital Research Team
- BL Labs
- Living with Machines

- Connect and share
- Invest in our staff
- Support digital scholars

- Agents for change
- Innovate and collaborate

https://goo.gl/tVwqHF
How do we think about Digital Scholarship?

"Digital scholarship allows research areas to be investigated in new ways, using new tools, leading to new discoveries and analysis to generate new understanding."

Dr Adam Farquhar
Head of Digital Scholarship
British Library

Combines methodologies from the humanities & social science disciplines with computational tools provided by computing disciplines

Scale
Perspective
Speed
Digital Scholarship methods

Location based searching & Geo-tagging

Using Application Programming Interfaces for datasets e.g. Metadata, Images

Visualisations

Corpus analysis, Text Mining & Natural Language Processing

Annotation

Transcribing

Crowdsourcing
Human Computation

@BL_Labs @mahendra_mahey @uvic @uviclib @britishlibrary labs@bl.uk

https://goo.gl/tVwqHF
Library Labs
– a space to experiment and innovate on-site and on-line

• Expert support and advice

• Essential equipment (software, hardware, storage, network)

• Essential ingredients (data, text, images)

• The ability to create, validate, capture, record, reproduce, archive, and share results

• Community, tutorials, examples

• Integrated into reference and research workflows
Growing GLAM Labs community…50 and counting…
Survey carried in in Sep 2018
Differences in GLAM Labs
Horses for Courses

• Variation in
  – Target users
  – Funding sources
  – Security models

• Surprises
  – Many do not facilitate access to restricted collections
  – Many do not provide dedicated physical space
  – Or simultaneous access to digital and physical

Get data here: https://goo.gl/66icov (you need a google account, you can get one here: https://goo.gl/CGdUhY)
Possible challenges GLAM Labs address

• Money spent on digitising / capturing digital – return on investment, how is it being used and what value and impact it is having, especially when opening collections for all.

• What digital collections are there that can be used openly and onsite and how do we tell people?

• How do we explore the ‘feel’ / ‘shape’ of collections at scale?

• How do we find, explore, augment discovery in often ‘messy’ cultural heritage data without public APIs?

• How do we discover, celebrate old culture & remix to create new culture?
Why are we doing this? (1)

We can learn how we are and should be supporting our users and this therefore shapes the services we build and problems and projects we work on, such as:

• Access, discovery to digital collections / data?
• Advice, guidance, technical support, training
• Services, Tools and Processes?
• Many more reasons…but learning where the gaps are

https://goo.gl/esqpRb
We help people ‘navigate’ their way through the ‘maze’ (sometimes) of the Library to what they want to do…

Requires understanding the culture of the organisation
Researchers often need a translator/advocate for successful projects.
Learn to wear the spectacles of the organisation, read their vision/strategy documents!
Our Audience and Collections

Audience research & Digital interests

Digital collections we have

This is where Labs works

It starts with making connections, engagement, talking to people!
All Labs need to do this!
Who do we work with?

Surprises of serendipity and creating luck?

- Researchers
  - https://goo.gl/WutNyI

- Artists
  - https://goo.gl/nKKnC2

- Librarians
  - https://goo.gl/9NWZUW
  - Curators
  - https://goo.gl/9NWZUW

- Software Developers
  - https://goo.gl/7QQ5Tf

- Archivists
  - https://goo.gl/x7b4tg

- Entrepreneurs
  - https://goo.gl/Fx8RG7

- Educators
  - https://goo.gl/qh01Mi
How?

Digital Research Support
Tell us your ideas of what to do with our digital content

Awards
Show us what you have already done with our digital content in research, artistic, commercial, learning and teaching, staff categories

Projects
Talk to us about working on collaborative projects

Engagement
• Roadshows
• Events
• Meetings
• Conversations

New!
Phases of interaction at BL Labs

**Phase 1: Exploration**

- Rethink

**Phase 2: Query-Focussed**

- Iterate

**Phase 3: Wrap-up**

Submit idea for support

Ideas always change

Once people experience the data and culture of the organisation
Labs Engagement 2013 - current

- Over 100 institutions visited
- Over 70,000 miles travelled around UK, USA, Canada, Australia, Europe, Middle East and Asia!
- 100s presentations & over 100 workshops
- 1500 researchers / artists / entrepreneurs / educators / public
- Over 1000 expressions of interest to use collections
- 150 researchers, artists, entrepreneurs & educators supported – potential case studies
- 200 TB of data via post
- 9 TB of data on data.bl.uk
- Over a billion views through Labs projects

© 2009 www.outline-world-map.com
Hard work, no magic formula!
A dozen BL Labs Lessons!

https://www.flickr.com/photos/26765085@N08/3365589331
Early ‘BL Labs’ lessons 1

https://www.theodysseyonline.com/one-size-does-not-fit-all
Engagement starts with people not technology!

Start a conversation, generate positive energy, encourage fun/play/experimentation and try to support as many ideas as is humanly possible, be kind, nice, want to share and genuinely want to help people!
Run Competitions

Good way to kick start engagement. Spread risk by having more than one finalist. Ensure entrants own their own IP, but all ideas are published. Good way to generate ideas for use.
Start small but think big!

Start with small experiments, digital use can be really simple, but OK to think big!
Early ‘BL Labs’ lessons 5

Keep it open, simple and don’t overcomplicate
Policies and processes for digital re-use are critical
Early ‘BL Labs’ lessons 7

Be brave! Fail fast!
Reject perfectionism, enemy of rapid progress!

Good enough is sometimes... good enough!

(This can be difficult message for Libraries... metadata will never be perfect!)
Services that allow *useful* exploration of cultural heritage data are rare!
Training or Collaboration?

Exploring data is difficult to do with large datasets

Often requires specific skills and capabilities that many of our users don’t have.
Celebrate the uses of digital collections!

Run Awards for those already using your digital materials, great way to find who is doing what with your digital content.
Success is rare, failure is common!

Success is sometimes all about the right people, place & right time...so it won’t always happen...
embrace failure, learn from it!
Example of useful pattern of research for GLAM Labs

- Finding invisible / well hidden things in ‘messy’ historical data
- Unearthing / unlocking hidden histories & data to stimulate new research
- Celebrating hidden histories / data creatively through events, art & performance
Finding things in ‘messy’
Optical Character Recognised (OCR) text

An example pattern of research

Mrs Folly
Code: Machine Learning / Reading

- Labs sometimes use Machine Learning / Reading techniques often called AI
- Analogies to how humans read / learn
- Machines acquire ‘knowledge’ / data, use that knowledge / data to make sense / identify patterns
- Labs doing this on a case by case basis so methods can vary but need computational AND human effort
- Legalities of Text and Data mining being ‘ironed’ out with publishers, on-going…Often a misunderstood …AI for good not evil?
- Perhaps we need a metaphor from history…
Smell of soup & Machine Learning
Who pays?

Thanks to Memo Akten (@memotv on twitter) for the inspiration!

Nasreddin, 13th Century Turkish Sufi
Victorian Meme Machine (2014)

Bob Nicholson interviewed on BBC Radio 4 Making History Programme:
http://goo.gl/4mV9ep

And telling jokes to the public:
http://tubmblr.com

Bob obtained further funding from his university
Looking for more collaborations

Learnt about access paths to digital collections

Rob Walker, Victorian Mother-in-law Jokes

https://www.youtube.com/watch?v=GRg7Q5OM0
Victorian Meme Machine (2014)

Rob Walker, Victorian Mother-in-law Jokes

https://www.youtube.com/watch?v=GRgj7Q5OM0
Katrina Navickas (2015)
Political Meetings Mapper

Learnt that domain knowledge reduces noise
Bringing History to Life to engage a wider audience!

We're now at the Red Lion pub in Soho, and we've just walked all the way from the British

https://youtu.be/0Ix0CL_dsQs?t=132

https://goo.gl/tVwqHF
Black Abolitionist Performances & their Presence in Britain (2016) – Hannah-Rose Murray

Frederick Douglass
Ellen Craft
Josiah Henson
Ida B Wells

http://frederickdouglassinbritain.com/

A Performance by JoeWilliams & Martelle Edinborough

Aberdeen Journal, 14 April 1847
“Frederick Douglass, The Emancipated Slave”

Started to implement Machine Learning Techniques
Microsoft Books…Our Dream Collection!!!

What can 65,000 books tell us?

Image: Artwork by Alicia Martin
Collection guide by Nora McGregor

Digitised printed books (18th-19th century)

https://www.bl.uk/collection-guides/digitised-printed-books
This is where true magic happens!

30 August 2012

https://goo.gl/tVwqHF
Practice what you preach!
Creating our dream example to inspire others

• The Labs team needed to run our own experiments...to understand our users
  – ‘Eating own dog food’

• One idea from hack event in June 2013 from Matt Prior looks promising…
Scissors and books – a match made in heaven?

RELAX Librarians!
It’s the digital version…
Done algorithmically via OCR process, details here:
https://goo.gl/jke4sy
Worked better for female faces than men's

Press

Creative uses of images

Internal IT/Wikimedia
Flickr Commons
Individual URL & API
1,020,418 images
need tagging!

Work @ BL by Ben O'Steen, Labs & Digital Research Team

Since Dec 2013
>1,000,000,000* views
>17,500,000* tags
>More demand to see physical items

Snipping out images from 65,000 Digitised Books*

Face recognition
Algorithms based on photos

Mechanical Curator
with an algorithmic brain
(Circles, Squares and Slanty etc)

Tumblr

http://mechanicalcurator.tumblr.com
Posts image every 30 minutes

http://www.flickr.com/photos/britishlibrary/

http://goo.gl/tVwqHF
https://goo.gl/FgZ4HM

*Matt Prior - http://goo.gl/j29Tnx

*Estimates

Worked better for female faces than men's
Tagging a million images

Iterative Crowdsourcing

Use computational methods

Top British Library Flickr Commons Taggers
18 hard core taggers
How to reward and keep motivated?

Mobile games for ‘Ships’, ‘Covers’ and ‘Portraits’

Cardiff University’s Lost Visions Project
Interface for tagging

http://www.metadatagames.org/

Average for ‘crowd’ is 1 tag per person

http://goo.gl/j6fxac

http://goo.gl/tVwqHF
Adam Crymble: Crowdsourcere Arcade

Using Arcade Games to help Tag images

Game Jam

Tag Attack

30 mins talk
Labs Symposium (2015)
Results of a Game Jam

https://youtu.be/xoCgHo2rwN4?t=82

The process of building these video games now is a lot easier than it was a few years ago.
Special Jury’s Prize (2015)

James Heald – Wikimedia and Map work

50,000 Maps Found in Flickr 1 million
Human & Computational Tagging & Community engagement
Geo-referencing work

Labs Symposium (2015)

https://commons.wikimedia.org/
https://goo.gl/tVwqHF
SherlockNet: Karen Wang, Luda Zhao and Brian Do

Using Convolutional Neural Networks to Automatically Tag and Caption the British Library Flickr Commons 1 million Image Collection

bit.ly/sherlocknet

>15.5 million tags added
>100,000 captions

Tags

Captions

Tags

Captions

12 categories

Pooled surrounding OCR text on page from similar images

Used Microsoft COCO (photographs) & British Museum Prints and Drawings collections as training sets.

“a man standing in a field with a cow”

“a close up of a bird on a tree branch”
Visibility – What happened to our Flickr images?

Understanding value / impact of making the BL’s data open / in the public domain

Peter Balman developed an analytics dashboard for the Library showing what is happening to our open Images

Number one use was?

Challenge details: http://goo.gl/Hb6I4A
David Normal - Artist

‘It was beyond my wildest dreams’

https://youtu.be/Q3SBxO34Zlc
Conflamingulation by David Normal
Curioscillotropy by David Normal

https://goo.gl/tVwqHF
Purcogitoresque by David Normal

@BL_Labs @mahendra_mahey @uvic @uviclib @britishlibrary labs@bl.uk
https://goo.gl/tVwqHF
Late August / Early September 2014

Four of these surrounded the Burning Man in Nevada Desert

Crossroads of Curiosity @ Burning Man

https://goo.gl/tVwqHF
Let's have a party!

Exhibited from June to Nov 2015

20th June 2015

Music mix by DJ Yoda using British Library Sounds: https://goo.gl/z3k4JT

Images from Burning Man and Flickr brought into the Poet’s Circle

Physical

Digital

Physical

Digital

https://goo.gl/tVwqHF
Mario Klingemann – Code Artist
Our first Artistic Award winner!

44 Men who Look 44
(Notice the direction faces)

Tragic Looking Women

A Hat on the Ground
Spells trouble

Mario Klingemann
German artist

Mario Klingemann is a German artist and Google Arts and Culture resident known for his work involving neural networks, code, and algorithms. He is considered a pioneer in the use of computer learning in the arts. (Wikipedia)

Known for: Digital art, Conceptual art

http://goo.gl/dM8ieA

http://goo.gl/tVwqHF
Mario Klingemann – AI Portraits

The Butcher’s Son
2018 LUMEN Prize winner
https://lumenprize.com/a-i-portrait-scoops-gold/

The future of AI art goes up for auction at Sotheby’s for $50,000
Mario Klingemann’s “Memories of Passersby I” generates portraits using neural networks. But that’s not what makes it art.

https://goo.gl/tVwqHF
Hey there Young Sailor – from Malaysia – Ling Low

Ling Low 2016 – Hey there Young Sailor

Hey there Young Sailor – from Malaysia – Ling Low

https://www.youtube.com/watch?v=bcOP1E5bRE0

https://goo.gl/tVwqHF
Hey there young sailor

https://youtu.be/bcOP1E5bRE0?t=99
Imaginary Cities
Exhibition 2019 (Michael Takeo Magruder)

Virtual Reality with Unity 3D

An artistic exploration seeking to create provocative fictional cityscapes for the Information Age from the British Library’s digital collection of historic urban maps

Exhibition: 4 April to 14 July 2019

https://goo.gl/tVwqHF
Michael Takeo Magruder – Artist Residency

Over the course of the project, I brought in two collaborators to help me realise some

https://youtu.be/hUWhc_nLyZU?t=208
Building GLAM Labs
The Cookbook and Showcase!
Coming soon!
Why this presentation?

• GLAM Labs are emerging around the world
• We share common goals to
  – Understand the value of a digital Lab for GLAMs
  – Share what we know and learn from others!
  – Explore differences in approaches
  – Build a support network
  – Build better GLAM Labs
  – Join our mailing list: http://www.jiscmail.ac.uk/lists/LIBRARYLABS.html
We want you to join us if you work in GLAM Labs!

Next event is in Copenhagen, 4-5 March 2019, see programme: [https://goo.gl/xNRxaS](https://goo.gl/xNRxaS)
Outcomes of GLAM Labs network

• Make existing or planned GLAM Labs be the best they can be
• Increase our joint understanding
• Build a supportive, kind, generous caring network
• Tell the world what we do and what you do!
Many miles to go…

Thanks to:
Eleanor Cooper (0.5) BL Labs
Adam Farquhar (Principal Investigator)

Alumini
Ben O’Steen – Tech Lead
Hana Lewis (0.5) BL Labs Project Officer

Digital Scholarship team at the BL
British Library Labs
Experiment with our digital collections

http://www.bl.uk/projects/british-library-labs
Funded by the Andrew W. Mellon Foundation and the British Library
Thank you!

British Columbia Research Libraries for inviting me and supporting this trip, University of Victoria for hosting and supporting this trip and especially Scott Johnston from the McPherson Library, University of Victoria for helping organising this trip!
In honour of…

my Canadian Indian/Punjabi family (past and present)…who settled in Vancouver and Victoria …from the early 1960s via India & Fiji
## Questions?

<table>
<thead>
<tr>
<th>Prompt Question</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I didn’t understand….</td>
<td>Can you tell me more about…</td>
</tr>
<tr>
<td>Why did you…</td>
<td>I am not sure about…</td>
</tr>
<tr>
<td>What if…</td>
<td>Why didn’t you…</td>
</tr>
<tr>
<td>What’s the best thing about…</td>
<td>What was the worst thing…</td>
</tr>
<tr>
<td>If you could have your time again, …</td>
<td>How did you…</td>
</tr>
<tr>
<td>I am not sure I agree about…</td>
<td>What was the biggest challenge…</td>
</tr>
<tr>
<td>What was the most successful thing about…</td>
<td>Who did…</td>
</tr>
</tbody>
</table>