Morning everyone.

I’m Mahendra Mahey, from the British Library in London, England, ‘Hello’. I am here to tell you my personal story about the experiences and lessons I’ve had working for my institution as well as with other Galleries, Libraries, Archives and Museums or ‘GLAMS’ at National, State, Public, University organisations and charitable and commercial organisations around the world. My story will focus particularly on how we have engaged with researchers, artists, educators and entrepreneurs from school children to adults who have used digitised and born digital cultural heritage collections and data to inspire them to create innovative, fun and inspiring projects. I would love it if my experiences can help other organisations build better ‘GLAM’ Labs, but I am also here to learn too from you.

For the last 6 years, I have been running ‘British Library Labs’ a digital Laboratory to encourage anyone to experiment with our vast, incredible, sometimes totally unique and mind blowing digital collections and data. Our work has been generously funded over these years by the Andrew W. Mellon foundation and the British Library. We are in fact waiting for news any day now to see if the Library will be moving our work into its core business on a long term basis.

During and after my presentation, please feel free to use twitter to amplify anything that resonates with you. My slides include much more information than I will have time to talk about, including links for you to delve deeper into the subject. On the right hand side in the footer there is a link to download all my slides, this
will appear on all my slides. Please feel free to reuse them but it would great if you could attribute me and the Library when you do.

My presentation will last about an hour. I will do my best to keep your attention. If there are any questions, something springs to mind, please make a note as. I will take questions at the end of my presentation, though I may ask you some quick ones along the way.

So like all stories, lets start at the beginning and let me take you on my personal journey.

Slide 2

I have always loved Libraries. They are places to time travel, to get lost and captivated. They take us on adventures in our minds and give us an experience that the web simply cannot give. The digital cannot replicate the tactile and physical experience of touching a book or looking at a manuscript for instance.

They allow us explore vast stores of information that are not simply available on the web and national libraries in some way have a responsibility to capture our nations’ memories when we are long gone and turned to dust. They hold the creative potential to inspire us. They can change us, and they can change our world as well as future generations to come.

The British Library’s mission is to support research, to inspire and enable enjoyment for everyone in the world.
The origins of the British Library stem back to the beginnings of the British museum in 1753, and it’s Library forms the foundations of our vast collection today. It would take another 200 years and an act of parliament before the British Library was born legally in 1973 and then another 24 years before our London collection was transferred from the British Museum to its current building at St Pancras which you can see, where around 20% of our physical items are stored.

<CLICK>

Designed by Colin St John Wilson, part of it is designed to look like a ship. In fact, many have said the Library represents a ship of knowledge sailing though a gothic landscape perhaps provided by the backdrop of the St Pancras Renaissance hotel.

<CLICK>

Incidentally, the hotel is a place of history itself. This where the Spice girls, ‘wannabe video’ was filmed, a song incidentally that is one of the most instantly recognisable studies have shown, back then it was called the St Pancras Grand.

The Library is one of 5 legal deposit reference libraries in the UK, it’s not a lending library. Whilst we acquire items through purchase or gifts, much of the collection has been built up through legal deposit. That is, by law, a copy of every UK and Ireland publication must be given to the British Library automatically with around 3 million physical items added each year. In 2013, legal deposit was extended to
cover non-print material which means by law we take in digitally published items as well. This includes regular mass crawls of the entire UK web domain as well as ebooks, ejournals etc which means our digital content is rapidly beginning to outgrow our physical collections as our digital items number billions of webpages for example.

<CLICK>
The building in London can sit 1,200 researchers at any one time across 5 reading rooms, we get around half a million visitors per year.
<CLICK>
Medium and long term requested items are held at Boston Spa in Yorkshire in a low oxygen warehouse, using robot to retrieve items. Boston Spa also has a reading room too where you can request items. In total, the library has nearly 1000 km of shelving, growing by 12 km every year.
<CLICK>
The British Library manages the UK public lending right, that is a living author’s right to payment/royalty each time their books are borrowed from public libraries.
Roly Keating our CEO, launched his vision for the Library in 2015. Whilst most people understand the Library’s role is as a custodian of knowledge it is also one of the largest research library’s on the planet and we carry out our own research. Our vision states our purpose is to make our intellectual heritage accessible to everyone for research, inspiration and enjoyment for everyone and be the most open, creative and innovative institution of its kind by it’s 50th anniversary.

To achieve this, we also help Businesses grow through a national network of Business and Intellectual Property centres (BIPC) offering IP advice, access to huge amounts of resources such as business intelligence, patents as well as events and training, we also engage in significant commercial activity through our commercial services.

We are also a cultural heritage organisation and have an ambitious programme of cultural activities to include exhibitions such as Harry Potter: History of Magic and host spectacular events during the day and when the library is closed concerts, performances, and soon even an Algorave, look it up!

We have a range of activities to support learning such as onsite and online courses for school children to adults and other activities around the local community in London as well as UK regionally.

Finally, we work with partners internationally on a range of projects to advance knowledge and mutual understanding.
The key to working with us on collaborative projects is to understand the way we see the world. In fact, we try to ensure that our requests for collaboration are prioritised and closely aligned to our vision. I would encourage you to delve deeper into our vision, especially if you would like to collaborate with us, it’s actually really fascinating read and it has helped me really understand the complex place where I work.

Slide 5

The picture you can see is inside the main building in London, it’s the King’s Library – King George the Third’s personal library, Mad King George! Sometimes known as the ‘stack’, I walk past this everyday and it gives me a sense of awe and reminds that the collections the British Library have are truly staggering and almost impossible to comprehend.

We currently estimate them to exceed 180 million items, representing every age of written civilisation and every known language. Our archives now contain the earliest surviving printed book in the world, the Diamond Sutra, written in Chinese and dating from 868 AD....only around 8% of our collections are books and as you can see we have so much more, please note the numbers are only really guesses as to exactly what we do have. If you saw 5 items a day it would take you over 80,000 years to see the whole collection.
Have you got X?

Looking for Physical Content in the British Library

For me, this is what it is like trying to find a physical item at the Library. It feels like a huge hypermarket, or perhaps even a factory or warehouse. It stocks a random assortment of things and if you ask the assistants they can tell you about things that are simply not visible on the shelves in huge storage facilities.
Moving on to our digital collections which is where my work largely sits. What percentage/proportion of our physical collections are digitised?

What surprises many people is that only an estimated 3% of our physical holding are digitised. This is because digitisation costs time and money and we have to achieve this through partnerships with commercial and other philanthropic organisations.

Through one of the first BL Labs project, ‘Sample Generator’ we discovered that our digital collections are not truly representative of our physical collections. There will be all sorts of reasons why certain items get digitised and others do not. In reality, all our collections be them digital or physical have selection biases. Our collections are hundreds of years of decisions made by people as to which items to buy, keep and which ones to discard.

In terms of licensing and using/reusing digital collections, a Lab like ours has further challenges. Out of our over 720 digital collections, only around 15% have an open license. The remainder are only available onsite at the moment. This is in part because many legacy digitisation projects didn’t always consider licensing when items were digitised. Trying to retrospectively establish rights and licensing on previously digitised collections costs time and money.
As a National Library, we have been collecting born digital items for over many decades. We are the home of the UK Web Archive that periodically captures billions of UK websites to keep for posterity and research and we are the home of the Alan Turing institute centre for AI and data science where we are an active research partner. For example, Living with Machines is a five-year £9.2 million research project that will take a fresh look at the well-known history of the Industrial Revolution using data-driven approaches.

CLICK

A new digitisation programme Heritage Made Digital is trying to learn from past digitisation projects, especially on digitising collections based on research demand.

Since 2012, we have developed a more systematic approach for agreeing terms of access to digital collections and data. What’s important to understand that the process is sometimes subject to competing concerns such as reuse, open research, licensing, ethics, revenue generation.
What’s important to understand is that if you really want to work with our digital collections, it sometimes pays to learn the ‘back story’ of how the collection came about, this was a really early and important lesson I learned. Knowing it, can have a significant impact on what you might want to do with it. On the screen you can see many factors. I simply haven’t got time to go into them all, but perhaps the most important one is the last one. Is there a human being around in the organisation who can tell you about the collection, as communicating with them may be the quickest way to learn more about the digital collection you want to work with. Often, they will have access to important information that isn’t written down.
As a Labs manager, I faced a significant challenge of how we would enable those who want to access to our 85% of digital collections that are only available onsite at the moment. If we look at this, we can see onsite at the Library may mean that the digital materials are only available in the reading room on a specific PC, or that the materials are still on their original storage media and may need obsolete equipment to access them or they still are as yet to be transferred onto a more modern system.

Some digital materials are only available through payment

Only a small fraction of digital materials are in the shiny happy carefree open web (more about how to access these later)

What we developed to tackle this situation was to develop a ‘Residency Model’ initially through an annual competition that we ran and now this has evolved in application process where researcher’s can apply to carry out digital research onsite at the British Library. These researchers in residence have special access to digital collections that our regular readers do not have. Access is strictly controlled depending what they would like to access and what they want to do with the materials.
Have you got X?

Looking for Physical Content in the British Library

So remember this? What is it like looking for digital collections at the British Library?

Have you got X digitised / in digital form?

Looking for Digitised / Digital Content in the BL

It can often feel like this...It’s much smaller, we have some free stuff, some can only be consumed on site, some you need to buy. If you speak to shop keeper, they may be able to get you to see what’s under the counter, because they couldn’t display it. You might be able to get special permission to get a look in the warehouse at the back of the shop which has even more goodies there. If you are looking for vegetables you have come to the wrong shop!
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

How do you find our open cultural heritage collections? One way is to use our collection guides which offer subject pathways into our collections, each guide will have a section non what’s available digitally if at all.
I haven’t got time to go into all our digital collections. Here is a small snapshot, which you are welcome to explore at your leisure at a later date, there are lots of links on this slide. The important thing to note is that it’s not just digitised books we have. It’s also playbills, magazines, newspapers, images, manuscripts, maps, usage data, catalogue data, broadcast news on TV and radio, sounds, music, sheet music.
What engagement does the BL have with people wanting use our digital content?

- 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…

What actual engagement do we have when we speak to people wanting to use our digital content? Well first of all people think you are a walking catalogue? And secondly, remember we only have around 3% of physical stuff digitised?

‘You are in luck’, we have what you are looking for! Or ‘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!**

We tend to attract projects with ‘fuzzier’ research boundaries.

We have a tendency to work on more flexible, interdisciplinary and collaborative research proposals.

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

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
I haven’t got time to go into detail about this is slide, but these are the kinds of things we were hoping researchers would do with our digital collections, especially things that would be very difficult to do manually!

Slide 19

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
Why are we doing this? (2)

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
- Artists
- Librarians
- Curators
- Software Developers
- Archivists
- Entrepreneurs
- Educators

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
Phases of interaction at BL Labs

1. **Phase 1: Exploration**
   - Rethink

2. **Phase 2: Query-Focussed**
   - Iterate

3. **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!
- 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

---

https://goo.gl/tVwqHF
Slide 30

Hard work, no magic formula!

https://goo.gl/tVwqHF

Slide 31

A dozen BL Labs Lessons!

https://www.flickr.com/photos/26765085@N08/3365589331

https://goo.gl/tVwqHF
Early ‘BL Labs’ lessons 1

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

Early ‘BL Labs’ lessons 6

Policies and processes for digital re-use are critical
Slide 38

Early ‘BL Labs’ lessons 7

Be brave! Fail fast!

https://goo.gl/tVwqHF

Slide 39

Early ‘BL Labs’ lessons 8

Reject perfectionism, enemy of rapid progress!
Good enough is sometimes…good enough!
(This can be difficult message for Libraries…metadata will never be perfect!)

https://goo.gl/tVwqHF
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.
Early ‘BL Labs’ lessons 11

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.

https://goo.gl/tVwqHF

Slide 43

Early ‘BL Labs’ lessons 12

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!

https://goo.gl/tVwqHF
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

Not the British Library!

Finding things in ‘messy’
Optical Character Recognised (OCR) text

Not the British Library!
**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](https://twitter.com/memotv) on twitter) for the inspiration!
Victorian Meme Machine (2014)

@BL_Labs @mahendra_mahey @uvic @uviclib @britishlibrary labs@bl.uk
https://goo.gl/tVwqHF

Victorian Meme Machine (2014)

@BL_Labs @mahendra_mahey @uvic @uviclib @britishlibrary labs@bl.uk
https://goo.gl/tVwqHF
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 'Bishop...
Black Abolitionist Performances & their Presence in Britain (2016) – Hannah-Rose Murray

http://frederickdouglassinbritain.com/

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

https://www.bl.uk/collection-guides/digitised-printed-books

This is where true magic happens!

30 August 2012

This is where true magic happens!

30 August 2012
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
Slide 59

Tagging a million images

Iterative Crowdsourcing

- James Heald
- Mario Klingemann
- Chico 45

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

- Average for ‘crowd’ is 1 tag per person
- Mobile games for ‘Ships’, ‘Covers’ and ‘Portraits’
- Interface for tagging
- Cardiff University’s Lost Visions Project

- Metadata Games
- Human Tagger
- Computational methods
Adam Crymble: Crowdsourced Arcade

Using Arcade Games to help Tag images

Game Jam

30 mins talk
Labs Symposium (2015)

Results of a Game Jam

https://youtu.be/xoCgHo2rwN4?t=82
Slide 62

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

Geotagging maps
Labs Symposium (2015)

Slide 63

SherlockNet: Karen Wang, Luda Zhao and Brian Do

bit.ly/sherlocknet

>15.5 million tags added
>100,000 captions

Tags

Captions

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

Pooled surrounding OCR text on page from similar images

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

Using Convolutional Neural Networks to Automatically Tag and Caption
the British Library Flickr Commons 1 million Image Collection
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/Hb6l4A

David Normal - Artist

'It was beyond my wildest dreams'
We then brought this to the British Library!

Slide 71
Mario Klingemann – Code Artist
Our first Artistic Award winner!

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

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

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

Hey there young sailor

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

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

Virtual Reality with Unity 3D
Exhibition: 4 April to 14 July 2019

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
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
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

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)
Alumni
Ben O’Steen – Tech Lead
Hana Lewis (0.5) BL Labs Project Officer
Digital Scholarship team at the BL

We would love you to experiment with our digital collections.
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
<table>
<thead>
<tr>
<th>Prompt Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>I didn’t understand….</td>
</tr>
<tr>
<td>Why did you…</td>
</tr>
<tr>
<td>What if…</td>
</tr>
<tr>
<td>What’s the best thing about…</td>
</tr>
<tr>
<td>If you could have your time again, …</td>
</tr>
<tr>
<td>I am not sure I agree about…</td>
</tr>
<tr>
<td>What was the most successful thing about…</td>
</tr>
</tbody>
</table>

| Can you tell me more about…            |
| I am not sure about…                   |
| Why didn’t you…                        |
| What was the worst thing…              |
| How did you…                           |
| What was the biggest challenge…        |
| Who did…                               |