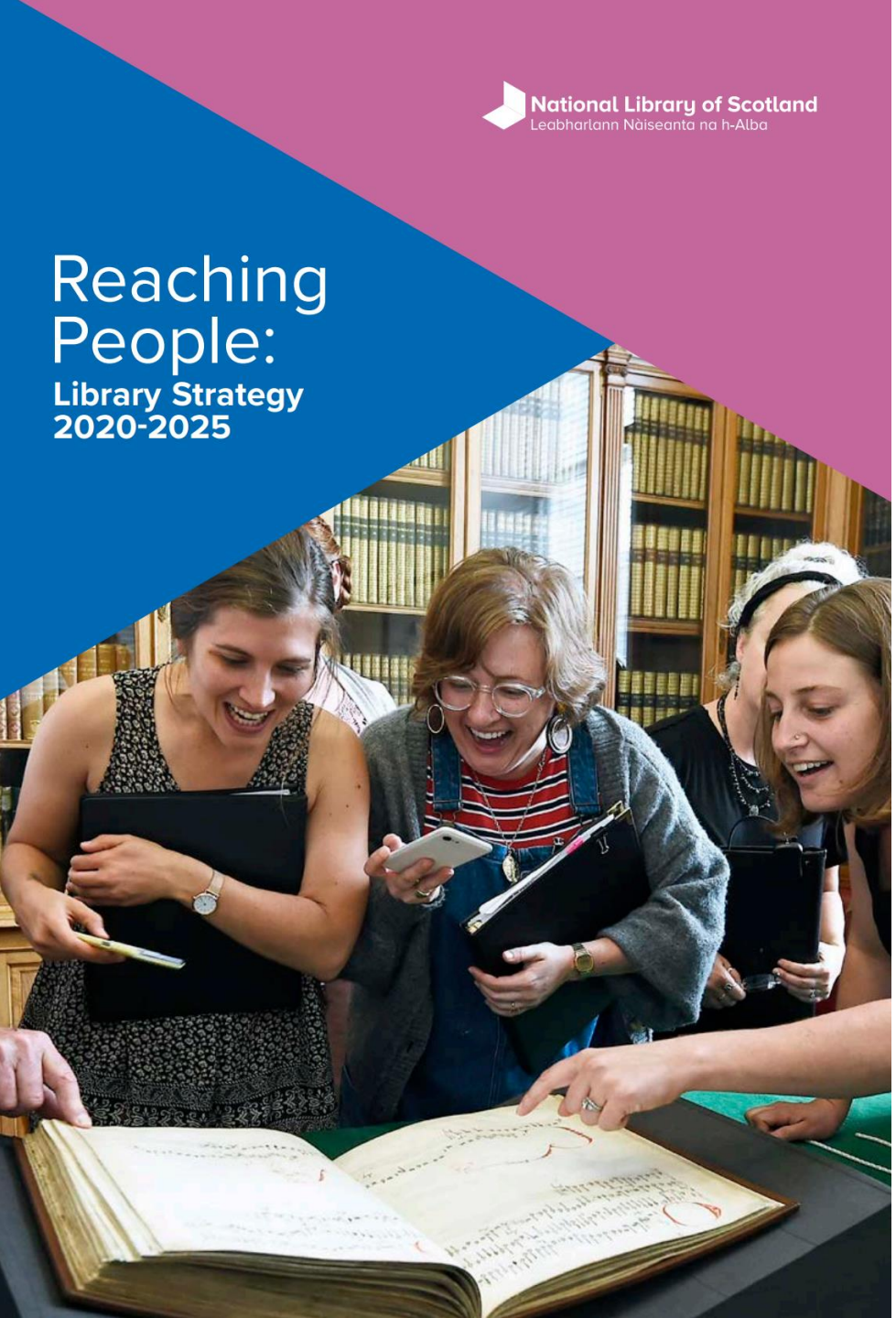
National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Collections as data at the National Library of Scotland: access, engagement, outcomes

Sarah Ames
Digital Scholarship Librarian
National Library of Scotland
@semames1 | #NLSdata | sarah.ames@nls.uk



National Library of Scotland
Leabharlann Nàiseanta na h-Alba

The National Library of Scotland

Over 31 million items in collection (excluding web archive, and owned-in-perpetuity digital collections)

'One Third Digital' by 2025

In-house mass digitisation programme

Strategy 2020-2025, 'Reaching People': strong focus on engaging with new audiences through the collections

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Introducing the Digital Scholarship Service

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Digital Scholarship Service

ENCOURAGE, ENABLE & SUPPORT USE OF COMPUTATIONAL RESEARCH METHODS WITH THE COLLECTIONS

ENSURE THAT THE COLLECTIONS ARE USED TO THEIR FULL POTENTIAL

ESTABLISH A LIBRARY CULTURE WHICH UNDERSTANDS DIGITAL SCHOLARSHIP

PRACTISE AND PROMOTE TRANSPARENCY IN OUR DATA CREATION PROCESSES

ANTICIPATE THE FUTURE OF RESEARCH

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

1. Making data available

TEXT

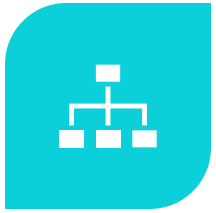
IMAGE

METADATA

AUDIOVISUAL

MAPS

WEB ARCHIVE



ORGANISATIONAL

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

2. External engagement

Collaboration

Projects

PhDs, residencies, fellowships

Beyond research community

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

3. Internal engagement

- Awareness
- Training
- Champions
- Culture

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Collections as data and the Data Foundry

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

No-nonsense data

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Identifying user needs

Pro	Intermediate	Beginner
All the tech skills! Will find a way to get the data no matter how presented But – has expectations of existing standards (where they exist) & consistency	Limited tech skills Understands value of different formats and approaches for research questions: theoretical rather than practical understanding Wants to get hold of the data easily to check what's there Likely to employ an RA to do the work	No tech skills Wants to use online tools to explore datasets Just wants the text

Plus broader audience includes other libraries (standards, presentation of data etc)

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

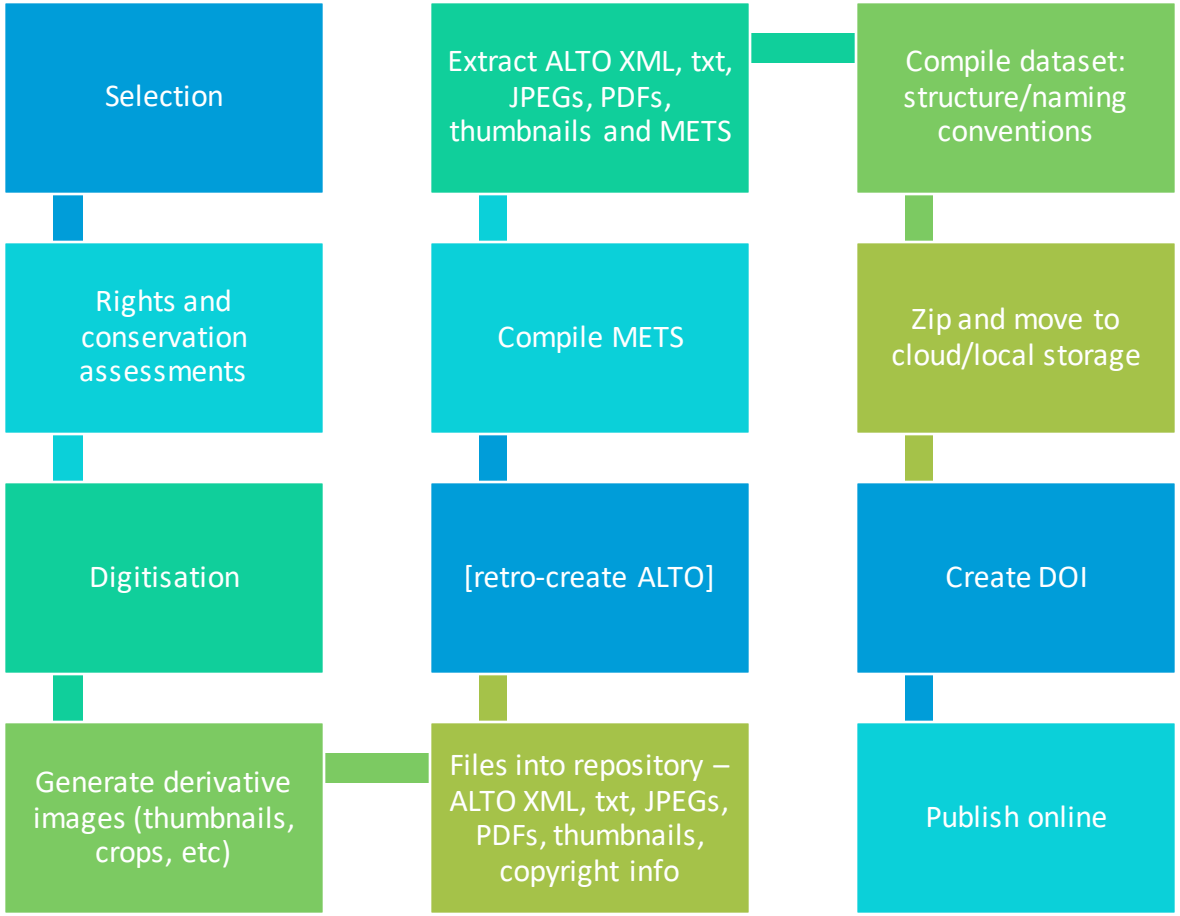
Identifying user needs

Pro	Intermediate	Beginner
All the tech skills! Will find a way to get the data no matter how presented But – has expectations of existing standards (where they exist) & consistency	Limited tech skills Understands value of different formats and approaches for research questions: theoretical rather than practical understanding Wants to get hold of the data easily to check what's there Likely to employ an RA to do the work	No tech skills Wants to use online tools to explore datasets Just wants the text

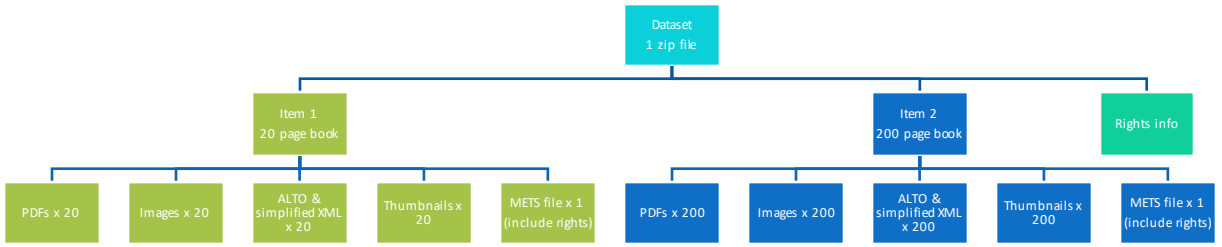
Plus broader audience includes other libraries (standards, presentation of data etc)

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Changing processes: digitisation to data

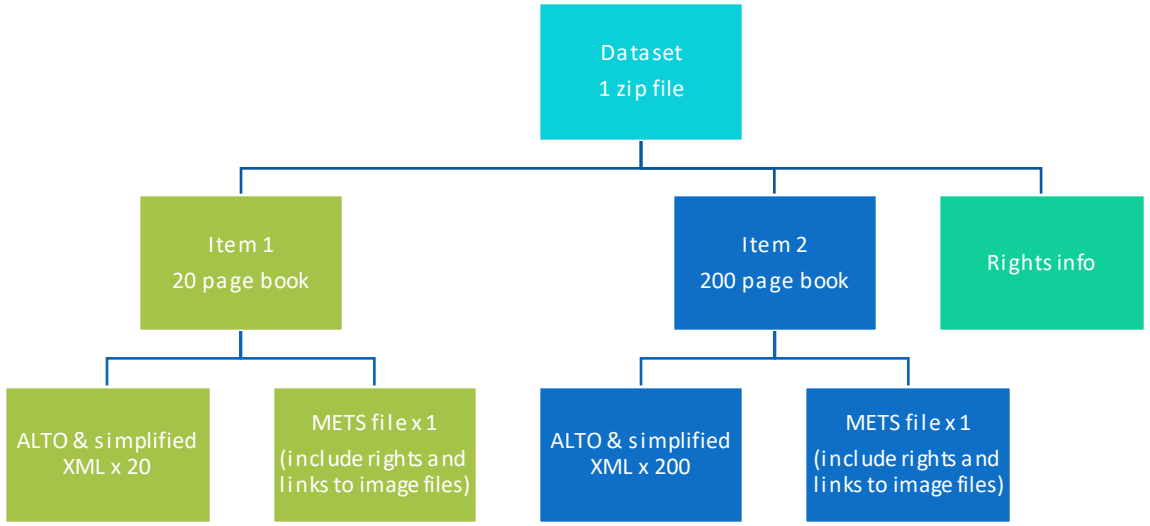


A

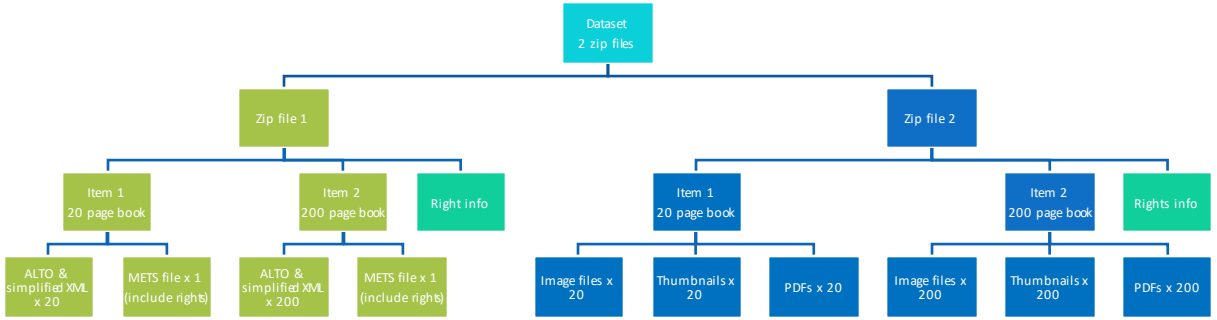


B

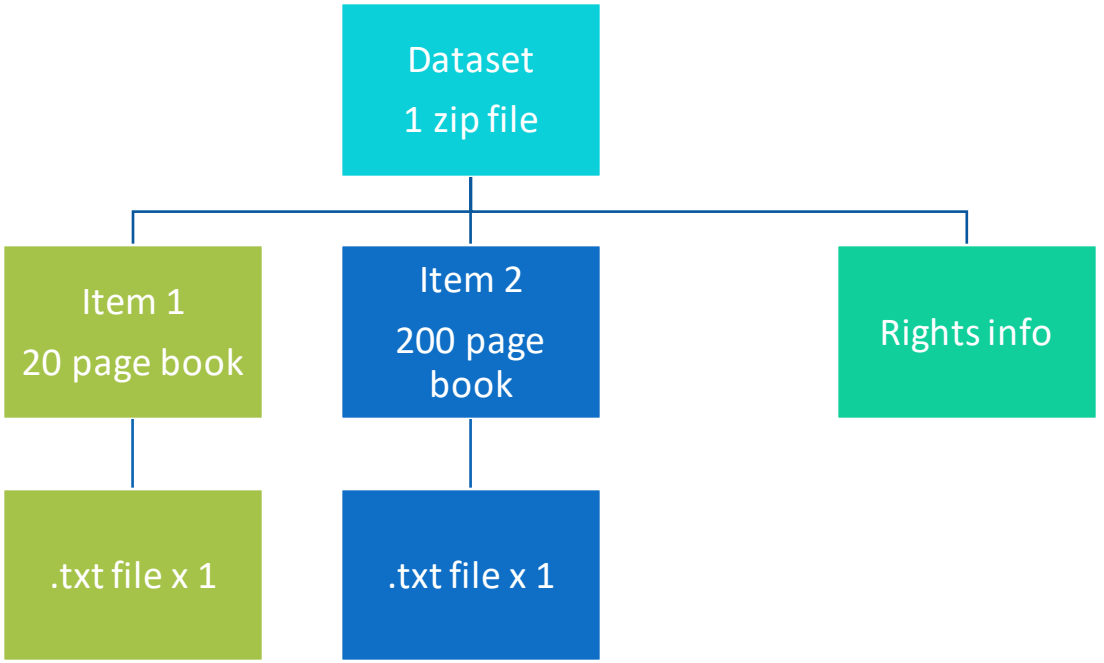
National Library of Scotland
Leabharlann Nàiseanta na h-Alba



C



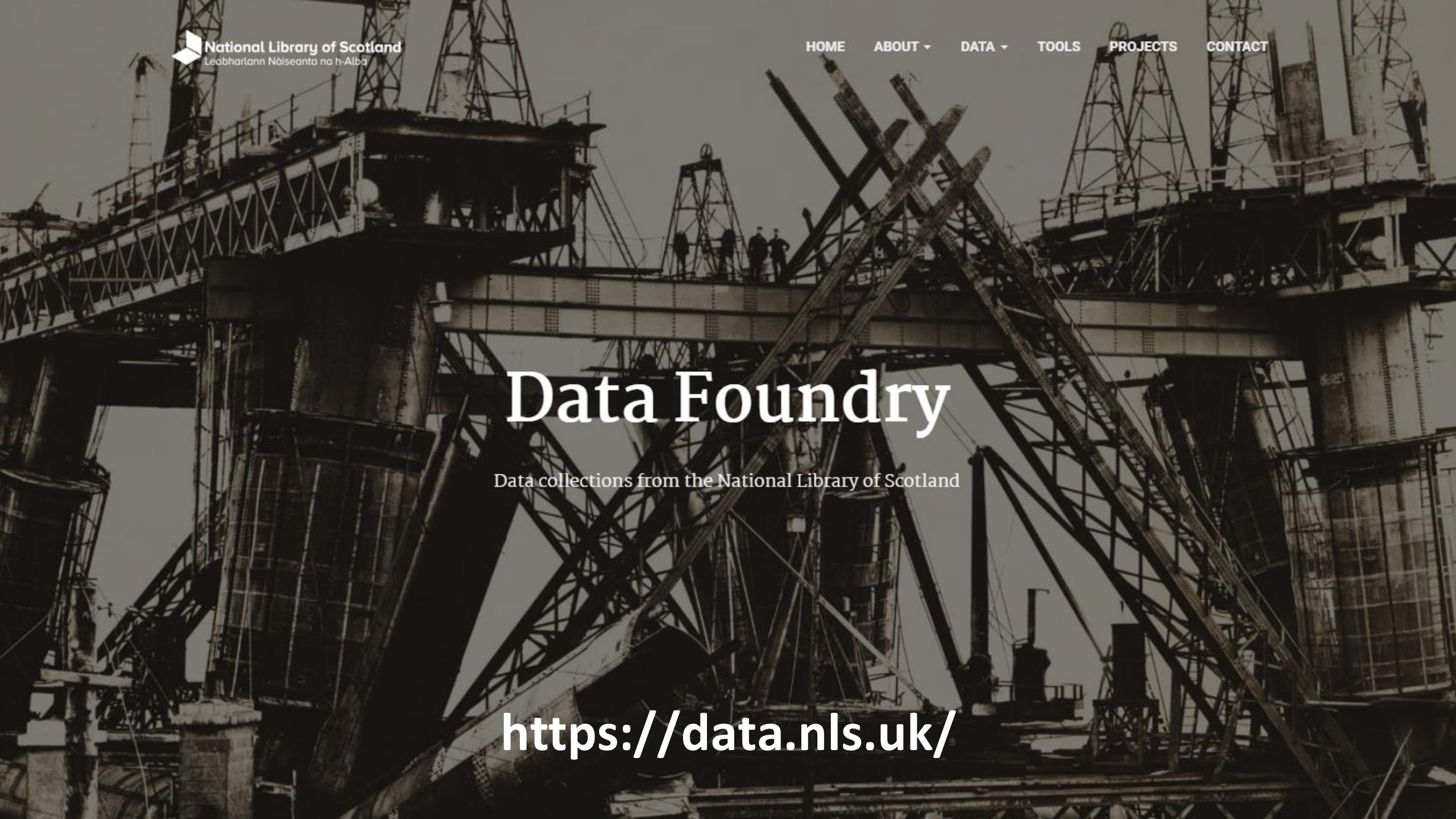
D



National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Making decisions

- Standards:
  - METS/ALTO and Plain text
  - MARC/Dublin Core
  - Tiered downloads
- Image sizes/quality to include
- Storage (local/cloud)
- What metadata to include and what is available
- How to be transparent: gathering and presenting dataset context
- Now moving on to more metadata collections, web archive data, spatial data


National Library of Scotland
Leabharlann Nàiseanta na h-Alba
HOME
ABOUT
DATA
TOOLS
PROJECTS
CONTACT
Data Foundry
Data collections from the National Library of Scotland
https://data.nls.uk/

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Our principles

National Library of Scotland data

Open

The National Library of Scotland publishes data openly and in re-useable formats.

Transparent

We take the provenance of our data seriously, and are open about how and why it has been produced.

Practical

We present datasets in a variety of file formats to ensure that they are as accessible as possible.

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

A whole-Library effort

Developers

Curators

Metadata

Rights

Access

Digitisation

…and the National Librarian! (‘Data Foundry’)

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Engagement activity and outcomes

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

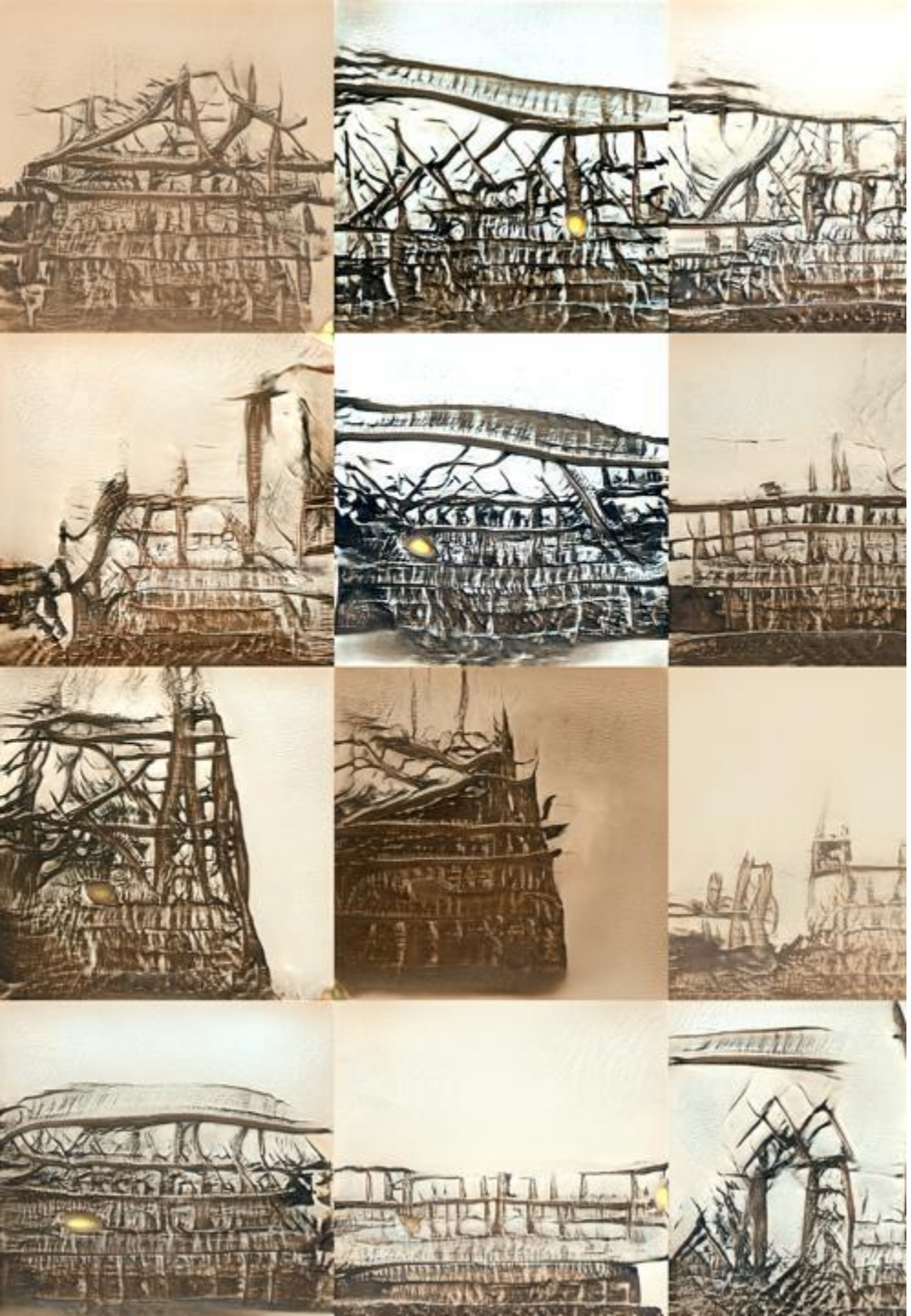
Teaching and learning

National Library of Scotland
Leabharlann Naiseanta na h-Alba
1870 - 1900
Birth of poster
& Art Nouveau
1900 - 1914
1914-1919
PROPAGANDA
1919 - 1950
1950 -
1972
GLOBAL
POSTER
COLLECTION

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

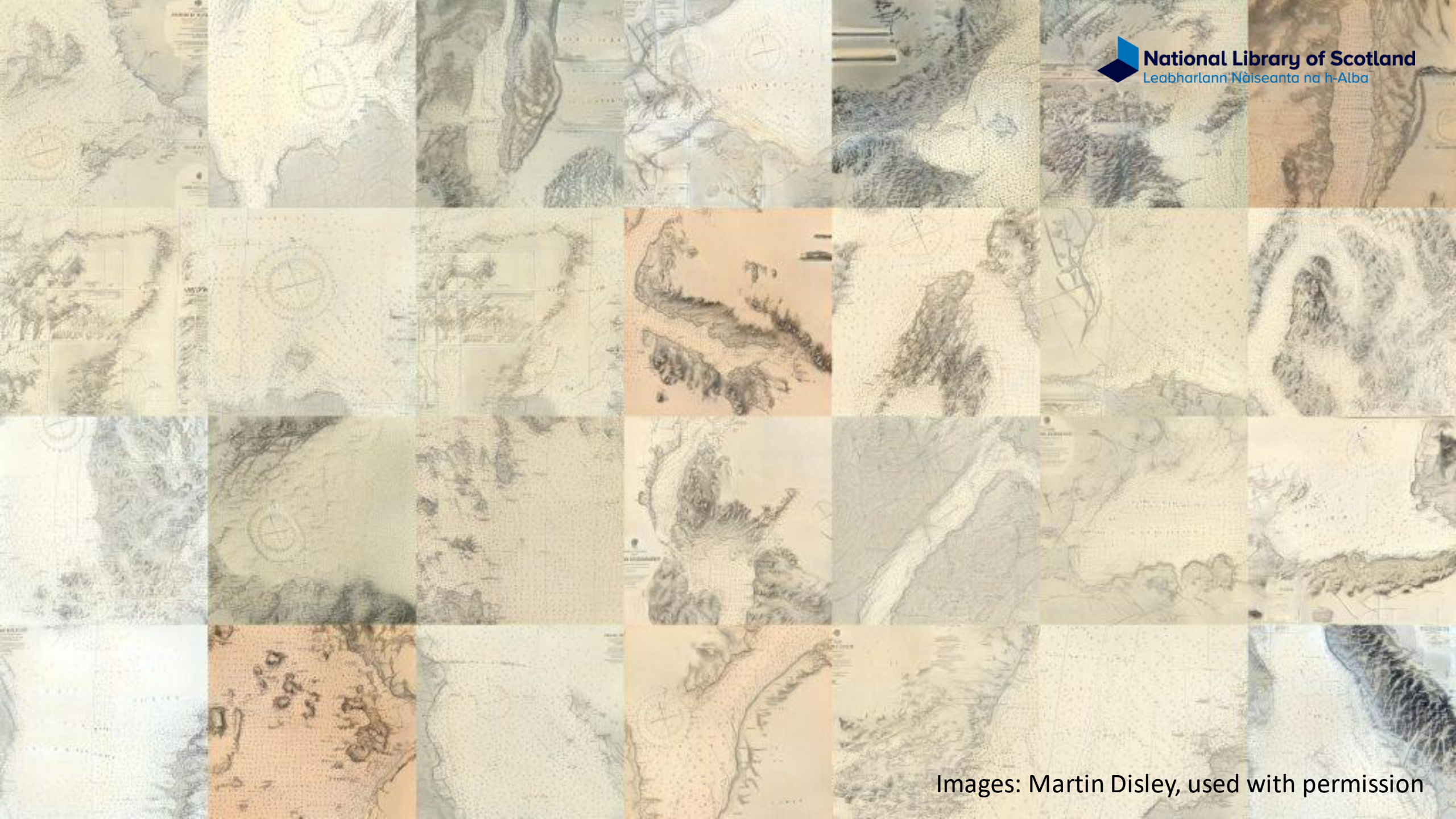
Creative partnerships



National Library of Scotland
Leabharlann Nàiseanta na h-Alba

the dataset is not the map is not the territory

- Martin Disley @martin_disley
- Funded by Creative Informatics (University of Edinburgh)
- Using GAN techniques with the collections
- https://data.nls.uk/projects/artist-in-residence/



National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Images: Martin Disley, used with permission

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

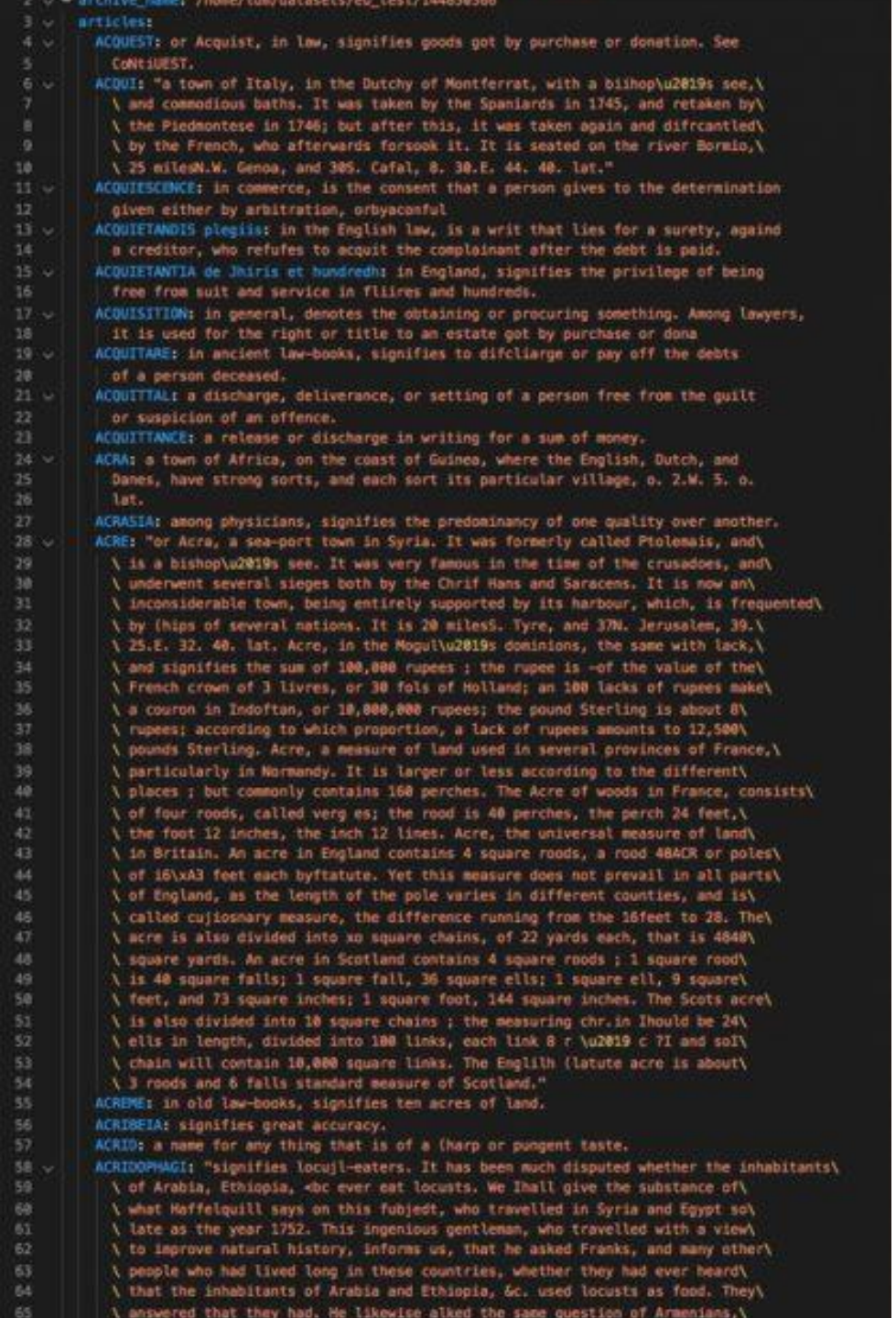
Is it true? The post-truth archive factory

- Marion Carré @Ma_rionC
- Collaboration with Goethe-Institut Glasgow, Alliance Française Glasgow and Institut Français d'Ecosse
- Challenge: how does AI open up new ways of interacting with library and archival collections and what are the challenges and dangers of using this technology in archival research?
- Will you archive 'fact' or fake news?
- Take part: https://data.nls.uk/projects/artist-in-residence-marion-carre/

Images: Marion Carré, used with permission

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Research collaborations



National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Annual fellowship

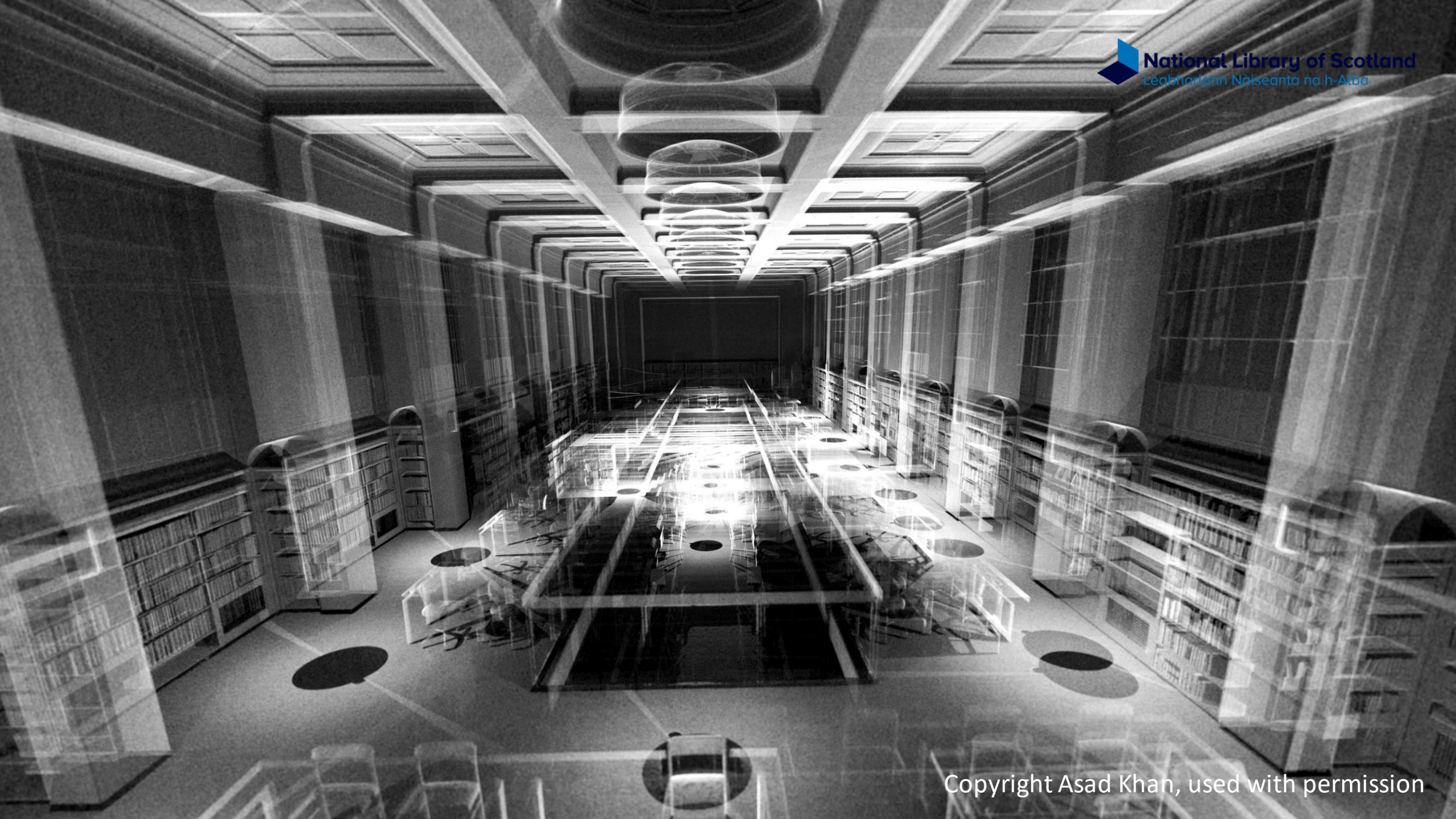
- Dr Giles Bergel (University of Oxford)
  - VGG computer vision tools
  - Chapbooks dataset
  - Exploring what can be learnt from their illustrations about chapbooks' origins; about relationships between chapbook printers, publishers and distributors; and about the type and range of imagery available to their readers.
- Dr Rosa Filgueira (St Andrews University)
  - AI toolkit for the collections: bringing AI tools to those who can't code
  - Initially with the Encyclopaedia Britannica dataset
- Dr Gustavo Candela (University of Alicante)
  - Linked Open Data – Moving Image Archive metadata

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

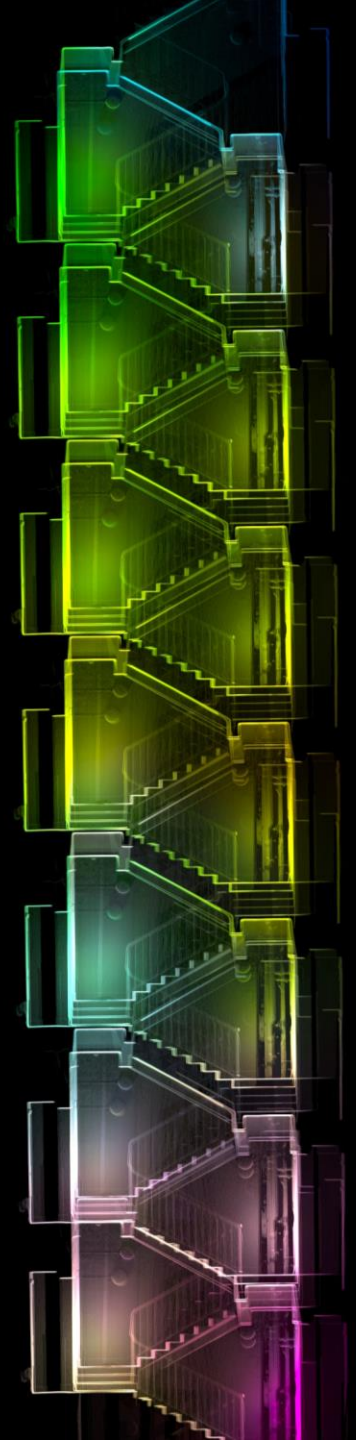
PhD studentships

Joe Nockels
'Adopting Transkribus in the National Library of Scotland: Understanding how handwritten text recognition will change management and use of digitised manuscripts'.
Supervisors: Professor Melissa Terras (University of Edinburgh), Dr Paul Gooding (University of Glasgow), Dr Sarah Ames and Stephen Rigden (National Library of Scotland).
Funder: Scottish Graduate School for Arts and Humanities, AHRC Collaborative Doctoral Award

Ash Charlton
'Slavery and Race in the Encyclopaedia Britannica (1768-1860): A Text Mining Approach'.
Supervisors: Professor Melissa Terras and Professor Diana Paton (University of Edinburgh), Dr Sarah Ames and Robert Betteridge (National Library of Scotland)
Funder: Scottish Graduate School for Arts and Humanities, AHRC Collaborative Doctoral Award

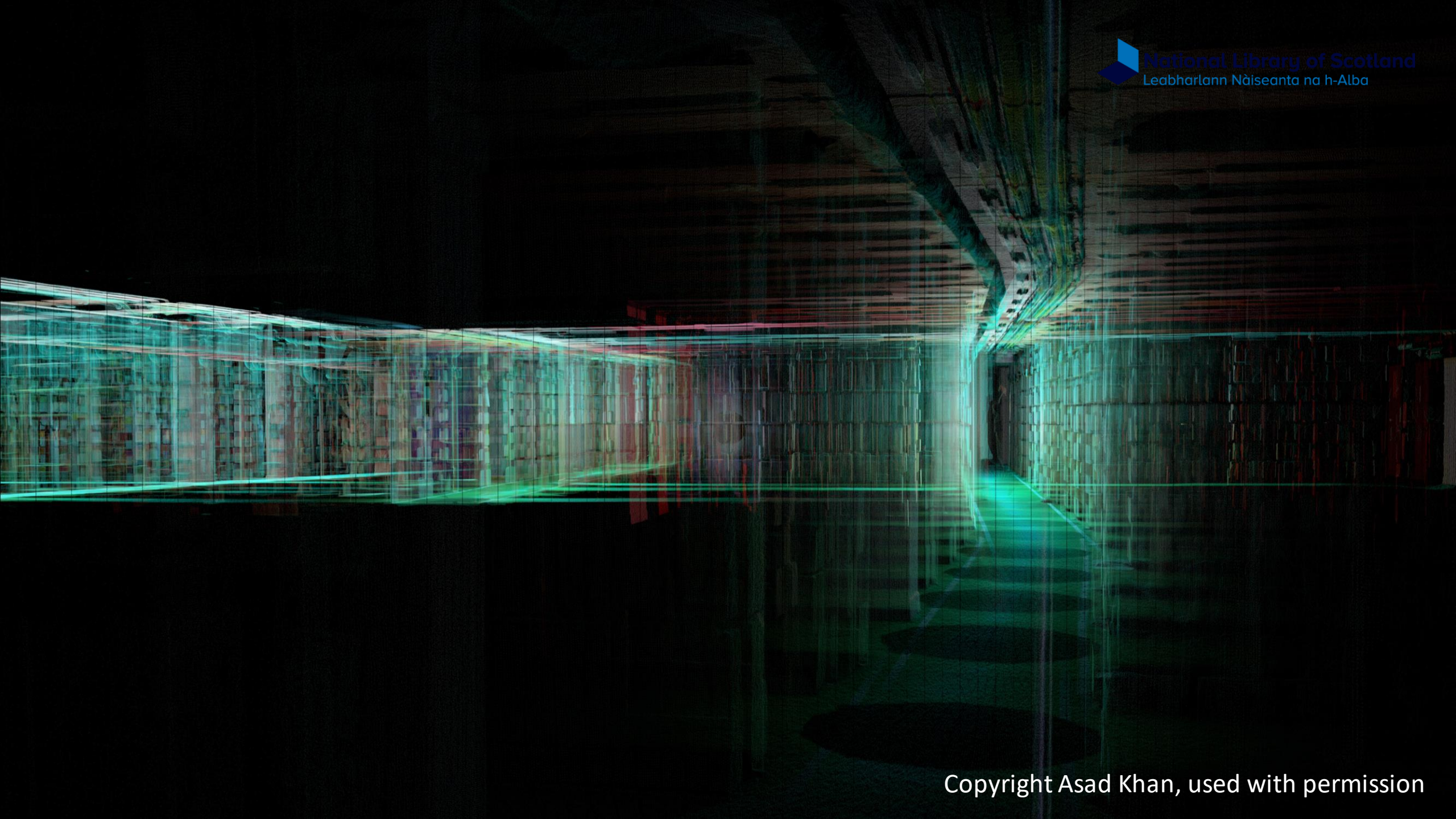


Copyright Asad Khan, used with permission

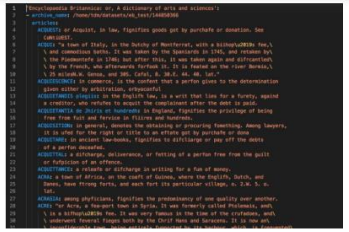


National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Copyright Asad Khan, used with permission

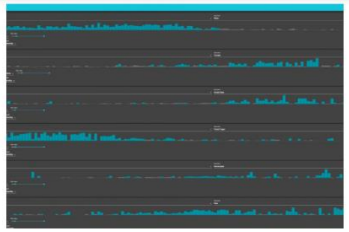


Copyright Asad Khan, used with permission

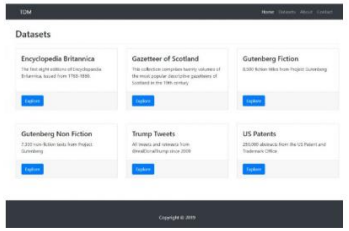


High Performance Computing meets Encyclopaedia Britannica

Data visualisation student projects

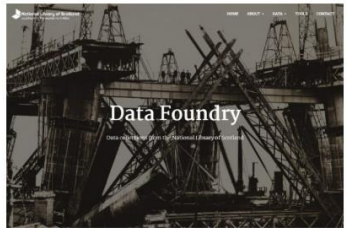


3,000 Scottish chapbooks...as music!

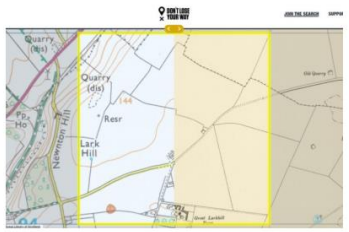


Text and data mining platform

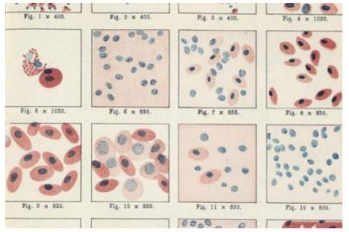
Artist in Residence



The National Librarian's Research Fellowship in Digital Scholarship



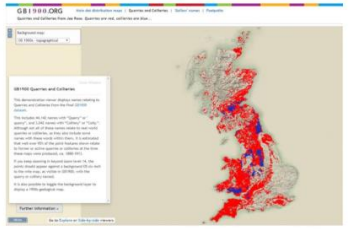
Finding lost footpaths using GB1900



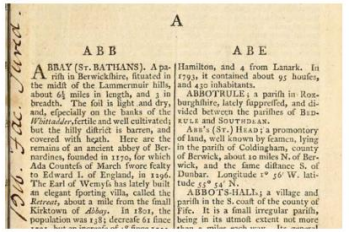
Teaching digital humanities with A Medical History of British India

GOETHE INSTITUT

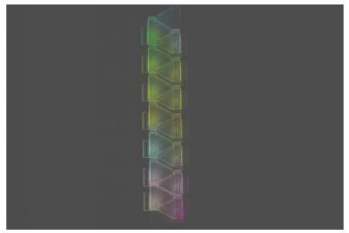
AI in Residence



Mapping quarries and collieries



Geoparsing the Gazetteers of Scotland



The Library as LiDAR (coming soon)

Projects

https://data.nls.uk/projects/

Examples of collaborative and individual projects over the past 6 months

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Challenges

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Change! Workflows, processes, culture…

Resourcing! No developers

OCR quality

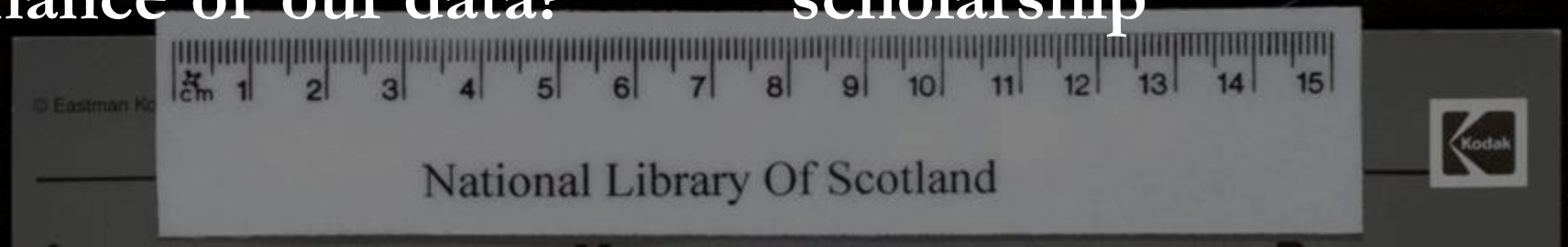
What does collections data look like?

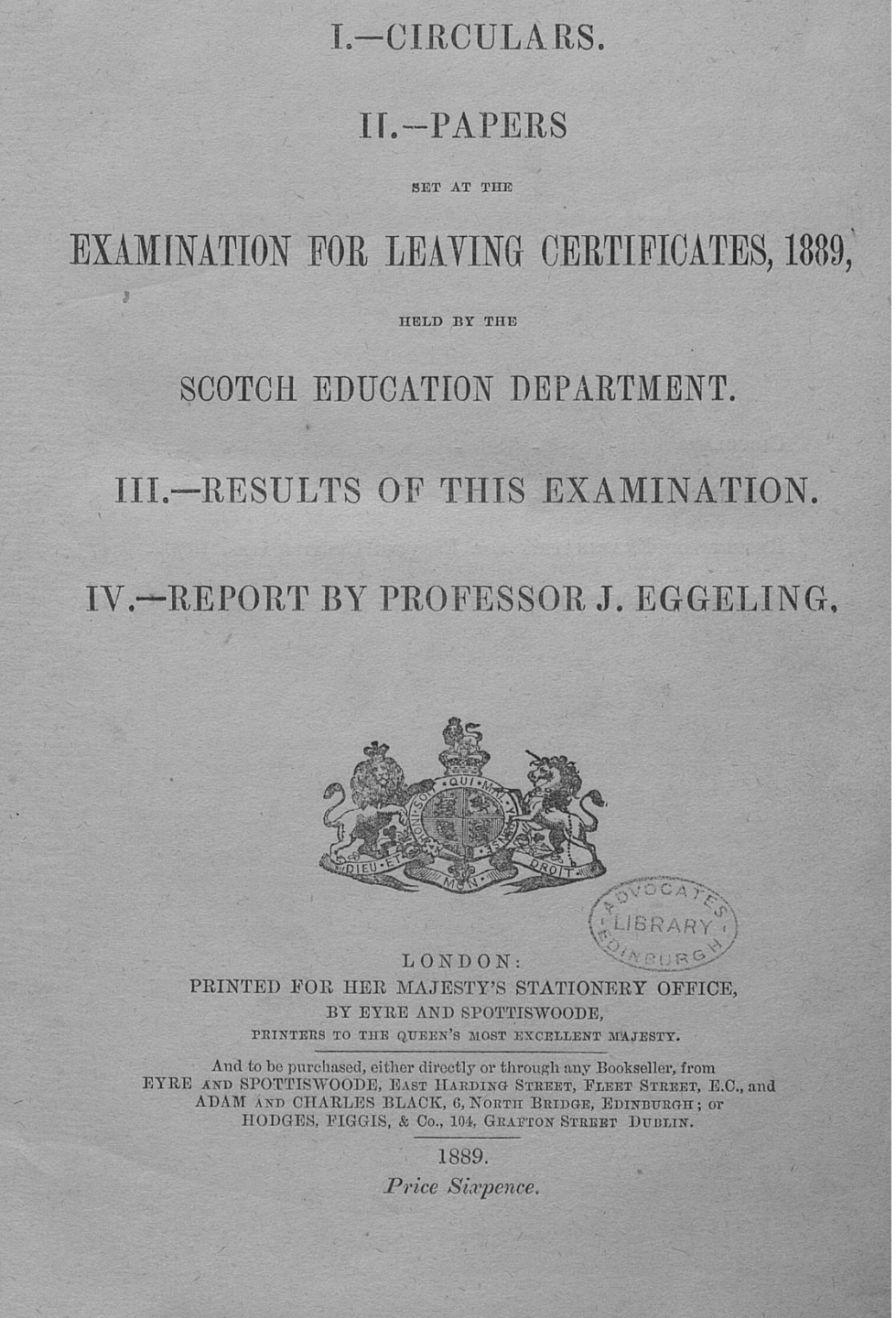
Copyright

Versioning

How can we be clear about the provenance of our data?

Scale and scope of digital scholarship





I.—CIRCULARS.

II.—PAPERS

SET AT THE

EXAMINATION FOR LEAVING CERTIFICATES, 1889,

HELD BY THE

SCOTCH EDUCATION DEPARTMENT.

III.—RESULTS OF THIS EXAMINATION.

IV.—REPORT BY PROFESSOR J. EGGELING.

ADVOCATES LIBRARY EDINBURGH

LONDON:
PRINTED FOR HER MAJESTY'S STATIONERY OFFICE,
BY EYRE AND SPOTTISWOODE,
PRINTERS TO THE QUEEN'S MOST EXCELLENT MAJESTY.

And to be purchased, either directly or through any Bookseller, from
EYRE AND SPOTTISWOODE, EAST HARDING STREET, FLEET STREET, E.C., and
ADAM AND CHARLES BLACK, 6, NORTH BRIDGE, EDINBURGH; or
HODGES, FIGGIS, & Co., 104, GRAFTON STREET DUBLIN.

1889.

Price Sixpence.

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

What have we learned so far?

1. It's worth putting effort and resource into presenting and promoting data collections
2. There is more to digitisation than online image galleries
3. DS has implications across the Library
4. Historic practices affect everything: it's difficult to break the cycle
5. Our users have exciting, refreshing, innovative approaches to using collections, which we can learn from

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

Thank you!

sarah.ames@nls.uk
@semames1 | #NLSdata