The Journey to AI Ready Collections at the University of Leeds

Jodie Double
Digital Content and Copyright Manager
University of Leeds

JISC Making Your Collections AI Ready Webinar 22 November 2023
Digital Content
Digitisation at the University of Leeds

- **2009**
  - Pre 2009 – commercial digitisation and small scale in-house

- **2009–2011**
  - LIFE-SHARE project = Creation of a Studio Team

- **2012–2019**
  - Internal Projects, commercial partnerships, UKMHL – Internet Archive

- **2015**
  - Risk Management Framework

- **2020–2022 +**
  - PANDEMIC

- **2022–2023**
  - Mass Digitisation Scoping Project

- **2024 onward**
  - Strategic Digitisation
RISK MANAGEMENT WORKFLOW FOR DIGITISATION OF IN-COPYRIGHT MATERIAL

- Identify item or collection to digitise
- Material covered under the Data Protection Act 1998
- Copyright holder known?
  - YES
    - Record status in EMu
    - DO NOT DIGITIZE
  - NO
    - Contact details?
      - YES
        - YES
          - HIGH RISK
          - Low Risk
          - MEDIUM RISK
          - HIGH RISK
        - NO
          - LOW RISK
          - MEDIUM RISK
          - HIGH RISK
      - NO
        - YES
          - LOW RISK
          - MEDIUM RISK
          - HIGH RISK
        - NO
          - LOW RISK
          - MEDIUM RISK
          - HIGH RISK

- Determine level of risk using Risk Calculator
  - Record risk number from calculator in EMu and level of risk (Low/Med/High)

- Undertake Level of Diligence
  - Copyright holder identified?
    - YES
      - Register work on EU Orphan Works register
    - NO
      - YES
        - Register work on EU Orphan Works register
      - NO
        - YES
          - Register work on EU Orphan Works register
          - DIGITIZE**
        - NO
          - YES
            - DIGITIZE***
            - NO
Working Towards AI and Research Ready Collections

- OCR
- HTR
- Georeferencing
- Transcription
- Catalogues as Data
- Full collection digitisation
- Learning from academics
- Etc...
Mass Digitisation Scoping Project (MDSP): Jan 2022-July 2023
18-month project surveying collections across the Library (Main and Special Collections) to understand the potential for mass digitisation with resulting feasibility study and Final Report.

Supporting new models of online delivery and collaboration to improve the user experience and enable access to material that cannot be obtained online from other sites or services.

Greater use of the library’s collections internally and externally.

Reduction in staff time for retrieval and processing for off-site collections due to digital access.

Digitised collections can be used at scale for collaborative research projects.

Ability to apply digital scholarship and/or digital humanities tools and research methods on collections.
So where are we on our journey to AI Ready Collections?
Partnerships, Collaboration and Sharing is key

- Foster and grow your networks
- Be an advocate for collections
- Make collections OPEN when possible and ready to mediate when necessary
- Work in partnership
- Continuous learning is the norm
- Be ready for change to happen at unexpected pace (slow and fast)
Where are we on the journey to AI readiness?

- Digital Library Infrastructure Project (DLIP) - 4 year project started Autumn 2023
- Strategic - Sustainable - Scalable Digitisation (moving away from the word MASS)
- Digital Creativity and Cultures Hub (DCCH)
- Developing closer working relationships with Library Teams
The Digital Creativity and Cultures Hub

Fostering digital research in the Arts, Humanities, Cultures and beyond

FIND OUT MORE →
Tools for Digital Research

A selection of tools for digital research: a mixture of free (opensource) and proprietary tools, arranged by broad area of practice. Existing UoL subscription is indicated where available. By no means comprehensive: if you have a favourite tool that isn’t listed here, please let us know!

- **Working with Texts**
  A selection of applications for textual analysis and editing.

- **Tools for Data Analysis and Visualisation**
  A selection of applications for visualisation, exploration and analysis of data.

- **Tools for Geospatial Analysis**
  A selection of tools for working with Geospatial data.

- **Tools for Archiving, Databasing and Exhibiting**
  A selection of tools for managing research and versioning, creating databases, archives and exhibits.
Managing risk

- RISK MANAGEMENT APPROACH
- KEEP GOOD RECORDS
- LICENSING STRATEGY
- TAKEDOWN POLICY
- GUIDANCE, SUPPORT AND TRAINING
Takeaways

- Remember that it is a journey and we are all learning along the way.
- Work in partnership inside and outside of the organization.
- Be brave and experimental.
- Keep good records at decision points.
- Develop or adapt a risk management strategy.
Thank you

Jodie Double
Digital Content and Copyright Manager
University of Leeds

j.l.double@leeds.ac.uk