



Digitisation Strategy for Cultural Collections at the University of Sheffield 2024-2032

PRINCIPLES

- Digitisation should be driven by the needs of the University and its communities, in support of the University’s vision “to deliver life-enhancing research, innovation and education that not only transforms the lives of our graduates but shapes the world we live in”.
- The strategy should align with the [Library Comprehensive Content Strategy | Library | The University of Sheffield](#) and the [Library’s Five Year Plan](#).
- All cultural collections held by the University are potentially within scope for digitisation.
- Content should, as far as possible, be open, freely discoverable, and re-usable. (FAIR principles)
- Digitisation should be carried out according to recognised professional and industry standards and reflect best practice in the field.
- All projects should include impact evaluation, ensuring that decisions are based upon evidence.
- Digitisation should deliver efficiency and good value for money, taking advantage of philanthropic and research funding where available to ensure sustainability and future development.

AIMS

Digitisation is undertaken in order to:

- Widen access to resources for researchers and other audiences.
- Facilitate outreach and public engagement.
- Preserve material that cannot otherwise be easily consulted.
- Make material available for use in publications, exhibitions, and teaching.
- Enable new research techniques, e.g. computational research including AI, to be used on collections and metadata.
- Lower the carbon footprint of the University, reducing the need for remote researchers to travel to consult materials on site.
- Increase the accessibility of collections to audiences who might otherwise have been excluded from consulting them due to socio-economic, health, or other circumstances.

PRIORITIES

When deciding which material to prioritise for digitisation the following criteria should be taken into consideration:

- Uniqueness. If material is not available elsewhere then digitisation becomes a higher priority.
- High demand. If the item is in high demand for research or teaching, digitisation may facilitate the item's wider use. (If multiple items within a collection are frequently requested for reproduction, then it may be more efficient to digitise the whole collection.)
- University need, e.g., if the item is required for engagement, outreach and/or impact, exhibition, hand-outs, or online learning resources.
- Potential for creation of searchable text or datasets. Material which facilitates the application of new research techniques, including AI.
- Conservation needs. If the item cannot be consulted because of its fragility and the risk of damage, then the creation of a digital surrogate can preserve the original from harm (noting that conservation work may need to be funded and scheduled into the process in order to prevent the process of digitisation from damaging the item further).
- Obsolete and fragile formats. If the item cannot be consulted because of its format, or the format is going to deteriorate, then digital preservation is required. (Digital images created for preservation purposes will not necessarily be made available publicly in all cases.)

A balance will need to be struck between fragile formats urgently needing digitisation for preservation reasons, and unique material in high demand for research. The strategy for 'at risk' material needs to accord with the University's Digital Preservation Policy.

APPROACH

Standards

The aim should be to create high quality files which meet current best practice. Ideally both an access copy and a preservation copy should be created. For preservation images, 600 dpi, TIFF format is currently the preferred minimum, although standards may change over time. Colour-matching should be used where appropriate. In all cases, non-proprietary formats should be used. PDF-A rather than PDF should be used for preservation purposes. Standards should be assessed at the start of each project. IIIF should be used for images to facilitate user access. Ideally TEI should be supported. Standards should be reviewed regularly to ensure that they are still fit for

purpose. Consideration should be given whether to destroy historic digitised files which fall significantly below current standards.

Legal rights

Collections prioritised for digitisation will be assessed for any Data Protection issues and the ownership of rights. If digitisation would infringe legal rights, or copyright clearance would be obtainable only at prohibitive cost or with excessive effort then digitisation will not proceed. Where ownership of rights is unknown, a 'best efforts' approach will be taken to identify rights holders. A risk assessment should be carried out and procedures put in place to manage any inadvertent rights infringement effectively. Infrastructure should be able to accommodate rapid take-down of images if necessary.

Workflows

Discovery and delivery systems need to be fully integrated to allow for a logical workflow for loading both objects and metadata. Infrastructure needs to support both discovery and delivery of images, and their re-use in multiple ways by end-users.

Metadata creation

While metadata regarding the attributes of the digital image can only be created at the point of digitisation, wherever possible metadata regarding the content of the item should be in place prior to digitisation. Automated processes should be employed to extract such metadata from existing systems (and indeed to generate metadata regarding the digital object), minimising staff time requirements. Opportunities to explore the use of AI should also be taken. This is still going to be the most time-consuming aspect of digitisation.

Licencing, aggregators, and commercial partners

The Library's aim is to make material and its associated metadata freely available for scholarly and private use. Current policy on licences for commercial use, where TUoS holds the copyright, will not change. Approaches from commercial partners to digitise content will be considered on a case-by-case basis, with full assessment of legal, contractual, and financial issues. Accepting a time-limited paywall may be considered only if this is a viable route to making key collections that would not otherwise be digitised available on open access within a reasonable timeframe.