Digitisation Strategy

Vision for 2017

By 2017, we will have digitised a significant amount of items from collections held at the University of Leeds, to help create, advance and disseminate knowledge. Digitised content will be delivered via the Library's digital collection repository "Leeds Digital Library" to ensure the unique and outstanding quality of the University's collections is reflected online.

This supports the Library's strategy, *Library Futures*, which states that 'increased use of digital technology to support and engage users' is one of the Library's priorities for the next five years.

To help deliver this priority we will develop an infrastructure to manage digital collections in a sustainable manner, supported by a robust IT infrastructure created for sharing and preserving content. We will create an online presence where students, researchers and the public come together to explore and interact with content in collections related to their interests and needs.

In addition to collecting, storing and preserving content, we aim to be a trustworthy creator of digital surrogates from original materials through our digitisation programme. Undertaking this programme of work will enhance our standing as an international university and support the University's strategic objectives to deliver an exceptional student experience centred on inspirational learning and teaching and to recruit and support high quality students from all backgrounds.

Digitisation objectives

These objectives will support the delivery of the digitisation strategy vision and ensure the needs of different users are met:

1. **Enhance the student experience with increased digitised content**
   a) Provide good quality, cost effective scans for OCR
   b) Provide digital images and OCR text to support undergraduate and postgraduate study
   c) Digitise the most frequently requested PhD theses
2. **Increase research carried out using Library collections**
   a) Identify and digitise key collections based on a survey of Faculty Team Librarians, Special Collections staff and academics (online questionnaire / focus groups)
   b) Liaise on project bids and provide guidelines for digitisation
   c) Undertake digitisation to support successful funding bids
3. **Increase digital access to Special Collections and Gallery collections**
   a) Provide images for customers on demand
   b) Create sample collections for Special Collections and Art collections in Leeds Digital Library[1]
   c) Provide images for the Library website, Leeds Digital Library and Turning The Pages[2] kiosks to support events and exhibitions
   d) Create digital surrogates for frequently requested items
   e) Provide images to support community engagement activities
4. **Support long term conservation and preservation of Library collections**
   a) Provide digital surrogates for fragile / deteriorating print materials
   b) Convert deteriorating or obsolete audio-visual formats to digital
   c) Provide images for condition reports of loan items
   d) Establish best practice guidelines for handling materials
   e) Develop procedures for long term preservation of digital files (as outlined in the DP strategy to follow)
   f) Use results of conservation audit to establish priorities for digitisation
5. **Build partnerships and strengthen existing collaboration through digitisation activities**
   a) Undertake external orders for partner organisations
   b) Liaise with local and national organisations to share best practice
   c) Develop agreements with partners for outsourcing
   d) Unite physical collections shared between the University of Leeds and other institutions

---

[1] Digitised Library and University collections available at http:\\digital.library.leeds.ac.uk
[2] Turning the Pages (TTP) kiosks display digitised images of books and manuscripts as an "ebook". Pages can be turned by swiping the kiosk's touchscreen. The Library has two kiosks, one located in the Brotherton Library entrance hall and the other in the Art Gallery. TTP software can also be used to display books via a website.

Digitisation Services

The work of the Library's Digitisation Studio covers a number of services:

- Online Course Readings
- Digitisation requests
- Etheses
- Digitisation programme

These reflect the different areas of work to be carried out in order to support the vision and the objectives. Table 1 (page 7) gives detail of the work that will be undertaken in order to achieve the Library's vision and objectives for digitisation services. The table also gives indicative throughput for each service and for the studio overall.

Online Course Readings (OCR)

This existing service provides book chapters and articles digitised under the CLA licence for use by students. Prior to the establishment of the Library's Digitisation Studio, the OCR team carried out some scanning in house and paid for other scans from the British Library. Since the beginning of this academic year, studio staff have supplied OCR content instead – producing scans of a higher quality and more cost effectively than those received from the British Library. They have also processed requests, created metadata and uploaded scans. This helps provide a more streamlined service for the delivery of readings to academics.

Demand for the OCR service continues to grow year on year. It is difficult to predict how this will continue, but it is likely there will be an increase in scan requests for the digitisation studio.

Digitisation requests

Requests include digitisation of Library collections for Special Collections customers and Library staff, as well as digitisation of content owned by individuals or other institutions.

Special Collections has provided digital copies of the collections for a number of years. A combination of SC staff with existing photographic knowledge, and the University's Print and Copy Bureau have fulfilled the requests. From January 2012, all requests for image capture have been fulfilled by the digitisation studio. The administration of orders continues to be handled by SC staff. Orders are passed to the studio and delivered within two weeks of payment, or sooner if paid for.

Fees for digitisation reflect the quality and resolution of the image the customer receives. The pricing structure (Appendix 1) is based on analysis of other comparable services at UK universities and cultural institutions. Images are delivered to customers on CD, DVD or via email. Any images which are suitable will also be made available online via the Digital Library and/or Turning The Pages.

Library staff are also able to request images on demand. Images are used to support content on the Library website, small scale projects and Library and Gallery events. These requests

are only for a small number of images. Any larger requests would need to be assessed as part of the Digitisation Programme below.

The studio has undertaken a number of orders for the digitisation of external content. Successful work has been done for Leeds Library and University of York, and is in the planning stages with the University of Sheffield. These orders are not only a potential source of income, but also a valuable way of building partnerships with appropriate organisations. However, this work does take the studio away from supporting the strategy's key aim of making Library collections available online. It is necessary to balance these conflicting demands and take on external orders where it is felt the relationship is important to the Library / University.

Etheses

A programme of PhD thesis digitisation started in April 2012. The most heavily requested theses have been identified. These are being scanned and made available online via the open access repository, White Rose eTheses Online (WREO). Specific titles are also digitised for customers via the request service.

Digitisation programme

In order to support the aim of making Library collections available digitally, it is necessary to develop a long term sustainable programme of work for the digitisation studio. This work would need to be carried out in addition to the on-request services listed above. Currently, a small amount of "programme" work is undertaken to support Library and Art Gallery events and projects such as Literary Archives. This level of work does not allow for larger scale projects, or engagement with academics to support teaching and research. A programme of digitisation will enable the studio to meet the needs of a wider range of users / customers. Regular production of new content will also ensure the Digital Library and TTP are dynamic and can showcase Library collections.

The development of a programme of digitisation will involve an assessment of collections and user needs. A Digitisation Prioritisation Group (Appendix 2 – suggested membership) will be created to oversee this process. The group will use a prioritisation model based on the Make It Digital scorecard developed by Digital New Zealand (Appendix 3). The following streams of work will be considered by the Prioritisation Group:

- Conservation / preservation (including audit)
- Special Collections (frequently requested items, collection highlights, treasures etc.)
- Library / Gallery events and exhibitions
- Marketing
- Academic requests
- Digitisation audit (including survey of academic and Library staff)

The Prioritisation Group will establish an annual workplan for the studio. This will be reviewed on a six monthly basis.

An Academic Advisory Group will be created to advise the Digitsation Prioritisation Group and make recommendations for the annual workplan.

Table 1   Digitisation services overview – showing staffing and studio capacity

Digitisation activity	Digitisation strategy objective	Number of pages / images per year	FTE required
Provide OCR scans	1a	25,000 (0.66min/page)	0.2
Carry out OCR admin	1a		0.4
Digitise and upload etheses to WREO	1c	50,000 (0.92min/page)	0.5
Produce condition report images	3b	100 (10min/page)	0.01
Create on-request images (customers)	4a	9,000 (10min/page)	0.9
Create on-request images (Library)	4c	500 (10min/page)	0.1
Digitise non-UoL content	5a	1,000 (10min/page)	0.1
All on-demand or 'essential' activity		**85,600**	**2.21 FTE**

Digitisation programme (3 options below)	3b		
Minimal production (ensures Library events are supported, DL and TTP remain dynamic)		5,000 (10min/page)	0.5
Substantial production (in addition allows for significant proactive digitisation to support research and teaching)		15,000 (10min/page)	1.5
Extensive production (would enable UoL to move to more innovative digital content within 3 years – does require extra equipment with space implications)		30,000 (10min/page)	3
Programme activity		**5,000-30,000**	**0.5-3.0 FTE**

Total resource required for digitisation activity (range depends on size of digitisation programme)		**90,600-115,600**	**2.71-5.21 FTE**

These figures do not include the FTE required to support line management, work co-ordination, advocacy and any other related activities. Figures do include minimal metadata creation and ingest of materials in Leeds Digital Library.

NB TTP content creation may be included as part of 'digitisation programme' – but this does not include the additional processing required to make images suitable for TTP software.

Current capacity of studio (fully staffed)		**156,850**	**5 FTE**

This is an estimate of what it would be possible to digitise on current studio equipment. The figure includes full time production on high volume, lower quality scanners, so it is higher than the figures for the Digitisation Programme proposed above. This gives an indication of what capacity there is for us to undertake additional digitisation – in addition to the Library's core digitisation and digitisation programme activity. The figure does not include the FTE required to support line management, work co-ordination, metadata creation, ingest of materials, advocacy and any other related activities.

Digitisation standards

The following standards will be set and adhered to for all work undertaken as part of the Library's digitisation services. Table 2 (10) gives an overview of theses standards for each digitisation service.

Resolution

Resolution for image capture reflects the number of pixels or dots per inch (ppi/dpi) captured. The higher the number of pixels captured, the higher the quality of the image will be. However, it is not always appropriate to capture images at the highest possible resolution. An increase in resolution increases the size of the file created which can be problematic for users trying to access items. Therefore it is important to capture images at different resolutions, depending on their intended uses.

File format

As with resolution, file formats should be selected on the basis of the intended uses of the files. Ideally, an open non-proprietary format should be used. This enables the files to be used across a range of software rather than tied to a particular one. Different formats also allow for the appropriate size and quality of images to be selected.

- TIFF – archival quality image file
- JPEG – image file for printing, online and personal research purposes
- PNG – image file required for TTP software
- PDF – provide searchable text capability for etheses and distribution copies of bound volumes
- WAV – archival quality audio file
- MP3 – distribution copy for audio files

Further detail is given in Table2, but our guiding principle for file formats will be that TIFF (or similar) will be used for preserving an archival copy of an image and JPEG, PNG or PDF will be used for presenting an image or other digital content to users.

Method of delivery

We have a number of systems which provide users with access to the outputs of the digitisation studio.

- VLE – items to support learning and teaching
- WREO – open access etheses
- Digital Library – a range of digital content available online, with some open access items
- K-Emu – collections management system for Special Collections and the Gallery
- Turning The Pages – items selected to showcase Special Collection and Gallery content. These will be used to enhance exhibitions, particularly displaying whole copies of bound volumes where possible.

All content delivered via WREO and the Digital Library will also be searchable via the Library catalogue.

Metadata

- Good quality metadata is vital to the effective management and online discovery of our digital objects.
- Different recognised standards will be used depending on the requirements of the items and methods of delivery.
- We will capture both descriptive and technical metadata where appropriate. This will be carried out within the digitisation workflow as it enables a more streamlined process.
- Many of the collections would benefit from additional information. To support this, it is intended to participate in community engagement activities and crowd-source some metadata.

Storage

- The digital output of the studio requires a significant amount of file storage. Archival quality images can be as large as 600MB each, so it has been necessary to purchase additional server storage from ISS. This will be reviewed as the output of the studio continues to increase.

Licensing

- Licences applied to digital content from the studio will reflect the aim of increasing access to Library collections.
- Content is created for OCR under the terms of the CLA licence and so must only be made available under those terms.
- For on-request and programme digitisation, where the copyright is owned by the University or where the item is out of copyright, content will be made available under a Creative Commons licence for re-use.
- The Creative Commons licence (Attribution Non-Commercial Share Alike http://creativecommons.org/licenses/by-nc-sa/3.0/) selected allows for sharing and re-use for non-commercial purposes. This would support educational use. Applying this licence allows users to know how they can use the content they find in the DL.
- A reproduction fee will be charged if people want to use the content for commercial purposes or are not willing to release their work using the same licence.

Table 2 Digitisation standards overview

	Resolution / quality	File format	Method of delivery	Metadata	Storage	Licensing
OCR	300dpi greyscale (and colour where appropriate)	PDF	VLE	DEDOCR	N drive	CLA licence
Etheses	300dpi greyscale (and colour where appropriate)	TIFF and PDF	WREO	DC	SAN	Open Access / Creative Commons
On-demand /Programme (image)	600-1200dpi colour	TIFF and derivatives (PDF with searchable text / JPEG / PNG)	DL K-Emu	DC	SAN	Creative Commons where possible / copyrighted where necessary
On-demand /Programme (audio)	MP3 – 256kbps	WAV / MP3	DL K-Emu	DC	SAN	Creative Commons where possible / copyrighted where necessary
On-demand /Programme (video)	To be decided	To be decided	DL K-Emu	DC	SAN	Creative Commons where possible / copyrighted where necessary

Advocacy and communication

- Communication for the Digitisation Service will raise awareness of the service and its capabilities, and showcase images of Library collections.
- Advocacy for the service will be carried out by Faculty team librarians, Digital Content and Special Collections staff.
- Leeds Digital Library will be used as the primary dissemination tool for digitised images. This will be supported by a range of different communication media (email, Twitter, blog etc), as outlined in the communication plan (Appendix 4).

Risk management

- Digitisation does pose a number of risks to Library collections, including potential damage to fragile materials.
- There are risks associated with establishing a strategy and not being able to achieve the aims and objectives.
- A clear governance model and procedures outlined in the strategy itself will help mitigate these risks.
- It is also important to consider the risk of not digitising. These are outlined along with the digitisation risks in Appendix 5.

Success factors

- This strategy establishes a number of key performance indicators (see Appendix 6) in order to better support the work of the Library.
- The Digitisation Service will produce an annual report for LT outlining the work of the studio mapped against the KPIs.
- Statistics from Leeds Digital Library will also be gathered and used to assess the impact of digitised content.
- Evaluation will include both quantitative and qualitative measures to ensure user needs are being met.

Evaluation and review

- This strategy and accompanying documents will be reviewed annually by the Digitisation Steering Group and modified as needed to reflect current priorities and infrastructure.
- Feedback will be gathered using a survey/focus groups with users (staff/students) to further help evaluate and shape the service.
- An annual report will be circulated widely outlining activities and achievement in the past year and plans for the next year.

Action plan

The following outlines the actions which will be undertaken in order to implement the digitisation strategy, and ensure the digitisation programme is realised in a systematic way with clear governance.

	Objective	Completion date	Staff responsible
Establish handling guidelines	4d	January 2013	Sharon Connell / Beccy Shipman
Digital preservation strategy	4e	April 2013	Jodie Double
Review prices	3a	July (annually)	Studio / DPG
Establish Digitisation Prioritisation Group (DPG)	-	February 2013	TBC
Assess / survey collections			
Survey academics (focus groups?)	2a	August 2013	TBC
Meetings with FTLs and SC staff to discuss collections	2a	June 2013	TBC
Use prioritisation model on current wishlist of Special Collections suggestions for digitisation	3c/d	March 2013	TBC
Identify frequently requested items	3d	June 2013	TBC
Identify collections to support upcoming events and exhibitions	3c	Set annual date	TBC
Publicise digitisation studio and support for research projects via FTLs and website	2b	Set annual date	DPG
Survey funding opportunities and UoL applications/awards	2b	July 2013	TBC
Develop template for submitting items to DPG	-	August (annually)	Beccy Shipman / DPG

Appendix 1 – Digitisation & publication pricing structure

Standard image files (JPEG 72 dpi files)

	Price (excl.VAT)	
	External customers	University of Leeds
1-50	£2.00 each	£1.60 each
51 to 100	£1.75 each	£1.40 each
101 to 200	£1.50 each	£1.20 each
201 upwards	£1.25 each	£1 each

High quality image files

Format	Price (excl. VAT)	
	External customers	University of Leeds
JPEG 300dpi (1-50)	£10 each	£8 each
JPEG 300dpi (51 upwards)	£6 each	£5 each
JPEG 600 dpi (1-50)	£20 each	£16 each
JPEG 600dpi (51 upwards)	£12 each	£10 each
Master TIFF (1-50)	£25 each	£20 each
Master TIFF (51 upwards)	£15 each	£12 each

Subject to conservation requirements, images will be supplied within two weeks of receiving payment. For an express service, images can be supplied within two working days at double the prices listed above.

Delivery charges

	Delivery via CD			URL via email
	UK	Rest of Europe	Rest of the world	
Price (excl. VAT)	£5	£7	£10	Free

Price for CD collection = 25p

External Users for Academic Publication = Charge for images & 25% of the Publication Fee

Leeds Academic Publications = No Publication fee. Charge just for images

Image publication charges – print run of up to 10,000

	One country (excl VAT)	World, all languages (excl VAT)
Book, magazine, ebook – editorial	£50	£100
Book, magazine, ebook – cover	£160	£280
CD / DVD / Record – cover	£200	£400
Newspapers – ½ page	£40 (regional) £60 (national)	
Newspapers – full page	£50 (regional) £75 (national)	
Video / CD-ROM / DVD	£80	£160
Website		£75 (1yr) £200 (in perpetuity)
Film / TV	£100	£200
Merchandise	To be negotiated	

Appendix 2 – Group membership

Digitisation Prioritisation Group

Brian Clifford – Deputy University Librarian (chair)
Member of LSAB to be decided

Role	Name
Conservation	Sharon Connell
Digital Content / Repositories	Jodie Double
Digital Content / Copyright	Beccy Shipman
Faculty Team Librarian (Medicine / Health / Sciences)	Sally Dalton
Heritage Collections / Faculty Team Librarian (Arts)	Maureen Pinder
Online Course Readings	Paul Cave
Marketing	Katy Sidwell
Metadata	Ian Jennings
Special Collections	Joanne Fitton

Academic Advisory Group

TBC

Appendix 3 – Prioritisation model

Prioritisation Model / "Digital Scorecard"*

Each of the principles used for the scorecard has four criteria that need to be weighted before undertaking the final scoring. A threshold for meeting or exceeding criteria will need to be determined to prioritise collection or item digitisation.

SCORECARD:
WEIGHT SCORING:

0	1	2
Not important	May be important	Important

FIT SCORING:

0	1	2
Does not meet	Partially meets	Meets

Collection or item name:					
Question	**Weight score*** add in score from chart above	**Fit score**	**TOTAL SCORE** F + W	**Meets or exceeds criteria (✓)**	**Fails (✓)**
1. Purpose of Copying					
The proposal demonstrates that digitisation will:					
a. digitisation will **improve availability** of the content, such as by increasing the opportunities for access, the potential audience, or the number of uses					
b. digitisation will produce **a faithful representation** of the material sufficient to be accessed as an alternative to the original					
c. digitisation will create **a complete replacement** of the material sufficient to be use permanently as the access copy instead of the original					
d. digitisation will create **a copy that transcends the original** enabling new ways of use,					

Collection or item name:					
Question	**Weight score*** add in score from chart above	**Fit score**	**TOTAL SCORE F + W**	**Meets or exceeds criteria (✓)**	**Fails (✓)**
access or representation					
e. digitisation will support learning and teaching at the University of Leeds					
f. digitisation will support research at the University of Leeds					
g. digitisation will support events, exhibitions and marketing for the University					
2. Access to the original					
a. the original is **rare, unique, part of a Heritage collection or has few viewable or usable copies** (regardless of form) that limits access					
b. there is **significant existing demand** to access the original					
c. the original is being put **at unacceptable risk of damage or loss** due to the current level of access					
d. the original is **difficult or costly to access** due to factors such as its location, conditions placed on access, or preparation required for access					
e. the original is deteriorating and valuable material will be lost if it is not digitised					
3. Digitisation technique					
a. the material can be **readily prepared** for copying					
b. the digitisation technology is **specifically designed** for the purpose					
c. the digitised copy can be **accurately described** to make it usable					
d. the digital format and carrier chosen can be **managed over time**					
4. Value of the copy					
a. the digitised copy will have **evidential value** that enables it to be referenced with confidence					
b. the digitised copy will have **value as a resource**, able to be utilised to create new					

Collection or item name:					
Question	**Weight score*** add in score from chart above	**Fit score**	**TOTAL SCORE F + W**	**Meets or exceeds criteria (✓)**	**Fails (✓)**
works or opportunities					
c. the digitised copy will be able to **convey the intrinsic meaning or qualities** of the original					
d. the digitised copy will **have integrity** through mechanisms such as quality controls, administrative metadata, authenticated access, or continued ownership					
e. the digitised copy can be made available online without any copyright restrictions					
Overall tally					

* **Adapted from:** makeit.digitalnz.org

Appendix 4 - Advocacy and Communication plan

To follow

Appendix 5 - Risk register

Likelihood has been estimated based on the digitisation strategy being approved and permanent staffing being allocated to the studio/strategy implementation (it will need changing if staffing is not allocated)

	Severity	Likelihood	Mitigation
Risk of not digitising			
Continued / increased handling and damage to fragile materials	High	Moderate	Assessment by SpColl staff, possible restrictions to public access
Loss of unstable formats	High	High	Conservation work to stabilise formats
Reputation of the Library / University next to comparable institutions	Medium	Moderate	Communicate other areas of strength prioritised over digitisation
Failure to meet needs of University staff and students	Medium	Moderate	Clear communication to staff and students
Failure to meet expectations of external users	High	Moderate	Clear information on Library website about lack of digitisation service
Indirect costs of not using equipment purchased	High	Low	Sell equipment
Risk of digitising			
Damage whilst digitising	High	Low	Assessment by SpColl staff, SC, Digi staff Conservation protocols and training
Increased requests to see / handle original	Moderate	Medium	Regular condition assessments for heavily used items Offer digital surrogates
Digitise copyright material (including orphan works)	High	Low	Collect copyright information in new Special Collections system Digitisation staff to receive training in copyright law / CLA licence Takedown policy on Digital Library New government register of orphan works
Increased footfall to Special Collections	Moderate	Medium	Good thing – need to plan for impact on staffing requirements?
Increase in digitisation requests	Moderate	Medium	Good thing – need to plan for impact on staffing requirements?
Increased digital storage needs	High	High	Planning with ISS will mitigate this, will involve increased cost for

			storage
Loss of digitised data	High	Medium	Create and implement digital preservation strategy as soon as possible
Risk to digitising / fulfilling the strategy			
Lack of internal capacity	High	Moderate	Long term plan for staffing needs
Loss of skilled digitisation staff	High	Moderate	Long term plan for staffing needs – extend staff contracts
Inadequate staffing in the Systems Team/ISS to support infrastructure (creation, discovery and preservation)	Medium	Medium	Build digitisation support into Systems Team operational plan, agree SLAs with ISS
Equipment failure	High	Low	Maintain service contracts
Lack of engagement with service	Moderate	Moderate	Implement communication plan
Lack of digital preservation strategy – could lead to backlog of digital content which will need preservation metadata.	Medium	Low	Create and implement digital preservation strategy as soon as possible

Appendix 6 – Key performance indicators

- Numbers of digitised items
  - Etheses
  - OCR requests
  - Special Collections materials
- Percentage of scans processed for OCR team within agreed timeframe
- Percentage of images produced for digitisation requests within two week limit
- Number of times digitised items accessed online / downloaded
  - WREO
  - Digital Library