GILLIAN BOAL

Conservation for Digitization at the Wellcome Library

INTRODUCTION

The Wellcome Library is one of the world's major resources for the study of medical history with holdings of four million books, two hundred and fifty thousand works of art on paper, and thirteen hundred paintings. The Library has begun the process of digitizing its entire collection. This paper outlines the considerations needed for the co-operation between all interested parties to enable the completion of this major digitization project in a safe and timely manner. It lays out the expected level of care and protection of the physical objects during the digitization of all of the Wellcome Library's holdings.

Plans to digitize the Wellcome Library collection has involved a collaboration with our development partners, Digirati, for the purpose of designing and building a viewing platform for the digitized material, known as 'The Player'. At present the Wellcome Player is able to render digitized books, archives, multiple-volume works, images, audio and video. Key features include deep zoom, the ability to download high- and low-resolution images and PDFs, and some 'pro' features e.g., the ability to embed an object such as a video clip.

All Wellcome Library staff are involved in the digitization process and have a shared responsibility for the care and handling of the collection during digitization. At first the role of Conservation in Digital Preservation was not obvious, but our contribution has become clearer as the overall project has unfurled, and conservation for digitization has now become one of our primary foci along with exhibitions. The learning curve over the last six years has been steep, and will continue to be so, for all involved. Some strict conservation standards have been challenged in our attempt to establish best practice and treatment fit for purpose.

What we have all come to understand is that Digital Preservation is about preserving digital and born-digital images while creating wider access to the collections. Digital Preservation should not be understood as preservation of the analogue, although it does broaden access to the analogue from which the images are derived. It is now widely recognized that digitization is no guarantee of the survival of content; indeed, digital files are more vulnerable than analogue originals.

It should be kept in mind that the digital images are not just virtual but also material, since they reside in the physical world. Even the Cloud is on servers somewhere. At the Wellcome the bits are stored in a Safety Deposit Box (SDB) v4.4, which is the digital object repository for all of the Wellcome Library's digital content, both digitized and born-digital. We are scheduled to migrate to Preservica Enterprise Edition late in 2014. The SDB is hosted within the Wellcome Trust and all digital content is stored locally on Wellcome Trust servers. With this in mind, in collaboration with the Digital Curator, we have composed a preservation policy encompassing all of our physical collections as well as digital and born-digital images, in a document entitled *Wellcome Library Preservation Policy for Materials held in Collections,* by Dave Thompson and Gillian Boal (2014).

In the digital era everything has changed but, no, nothing has changed. From a conservation perspective digital capture is fundamentally no different from other previous reformatting technologies, such as photocopying and microfilming. Conservation concerns have always focused on the safe handling of collections during these procedures and on the need to establish ways of minimizing the damage inherent in *all* forms of use.

The context of use—in this case, digitization—does make a difference. Researchers in the reading room and photographers in the studio during digital capture have different intentions, different requirements and make use of human faculties. Eyes can look around corners and cope with undulations, but capture equipment requires the item to be flat. On the other hand, photographic technologies are able to record information outside that of normal 'naked eye' vision. By this is meant the ability to range beyond the visible spectrum—ultraviolet, infrared and X-ray. If you see shadow in

Presented at the Book and Paper Group Session, AIC's 42nd Annual Meeting, May 27–31, 2014, San Francisco, California.

Fig. 1. Evidence of shadow.

an image, as in figure 1, then no crease has been made to the foldout during digital capture.

The digital photographer is intent on getting the best image. The current technology is much better than it was, but it is still limited. Digital capture should not be at the expense of the physical items being digitized. Conservators have a vested interest since it is they who will be doing the resulting repair work. The role of Conservation and Collection Care is to ensure that respect for the item is maintained across contexts of use: working closely with digitizers, creating condition surveys of collections that allow for good preparation, learning the physical needs of the equipment, writing guidelines and giving handling training can all help instill a respectful approach to digitization.

HISTORY OF DIGITIZATION PROJECTS

The Wellcome began with the digitization of the Arabic manuscripts in 2009. This was a collection with unique characteristics but we were able to learn about the overall needs of the digitization project.

A further pilot project was undertaken with several different components, for example, the Medical Officer of Health Reports which were prepared in the conservation studio for digitization off-site: thanks to this procedure we were able to understand the issues related to borrowing from other institutions, and the process of packing and dispatching the MOH Reports to Holland for digitization. The Genetics Books project involved materials digitized in-house by Wellcome photographers, and the Early European Books project was contracted to Proquest, and then subcontracted to the Numen Company for on-site digitizing at the Wellcome Library.

Currently, in Phase 2, there are several projects under way in-house simultaneously; among them, Medieval Manuscripts and 19th Century Books, which are being digitized by our own Photography Department and by the Internet Archive.

GENERAL REQUIREMENTS FOR THE DIGITIZATION PROCESS

With good planning, sufficient resources, a pragmatic approach to the condition of the object, and by withholding and putting out of scope where necessary, most of the collection can be safely digitized.

Conservation and Collection Care is one component in a complex workflow. The greatest risk to any physical object is in the handling, and the nature of the digitization process means that objects are often handled in new ways in new contexts, and with greater intensity. It is the risk of damage or loss from handling that all who are involved aim to mitigate. Communication among all parties in the digitization process is key.

Each digitization project will have specific values for each of the following variables: the scope of the collection to be digitized; its current physical condition; the demands of the equipment to be used; the experience of the trained personnel involved; and other resources available to the project.

Early participation of Conservation in a digitization project is necessary. A member of the Conservation team should attend preliminary meetings at the inception of the project. Their continued level of involvement rests on a number of factors, listed below.

GUIDING PRINCIPLES

- To prevent loss
- To ensure safe preparation of materials for imaging
- To provide handling training to all concerned in order to minimize damage from use, to be extended to digital preparators and imaging technicians
- To be prepared to mend after digitization

IDENTIFIED RISKS

Examples of risk factors from digitization at the Wellcome Library include:

- *Risk of loss*: particularly fragile items have been found in many collections selected for digitization and risk of loss has often been mitigated by ensuring that these items are imaged in-house where trained conservation support is immediately available. Fragile items may also be rehoused before image capture (e.g. using Mylar sleeves) to facilitate handling.
- *Risk of damage*: early printed books selected for digitization have at times had fragile bindings with restricted openings, where forcing them to the 110° required by a book cradle would have resulted in irreversible damage to the spine. As such, books that could not open comfortably to 110° were flagged for inclusion in a supplementary project workflow using an image capture device capable of photographing with lower opening angles.

Fig. 2. Gutter inaccessible.

- *Risk from repeated handling to capture a suitable image*: a variety of options are needed to address this issue, ranging from preparation by Digitization staff and Photographers, to technical repairs performed by trained Conservators. In the case of the Medical Officer of Health Reports, there was a need to disbind the reports to facilitate image capture. In some cases e.g. where the binding is too tight, a lower quality image may be a necessary compromise (e.g., allowing text to disappear into the gutter of bound material) (fig. 2).

SELECTION OF COLLECTIONS FOR DIGITIZATION

The selection of material for digitization is the first step in the preparation phase of a digitization project. The general physical condition of the collections to be digitized and the kind of image capture equipment being considered should be discussed early on in the selection of each project.

PROJECT LEAD TIME

Successful care of collections throughout a digitization project requires adequate lead time for planning, preparation and execution. This is necessary to achieve the appropriate balance between necessary collection care activity and project expediency during digitization. The time needed for a conservation assessment or survey turned out to average two to four weeks but extra time was allocated to accommodate unplanned contingencies. The initial scoping and assessments should inform the project lead time. The amount of time needed for pre-digital stabilization will vary depending on the outcome of these assessments.

CONSERVATION CONDITION ASSESSMENT

While the emphasis has been on facilitating the digitization process where possible, some books may nevertheless have to be deemed out of scope. This may be through lack

Fig. 3. Uncut pamphlet.

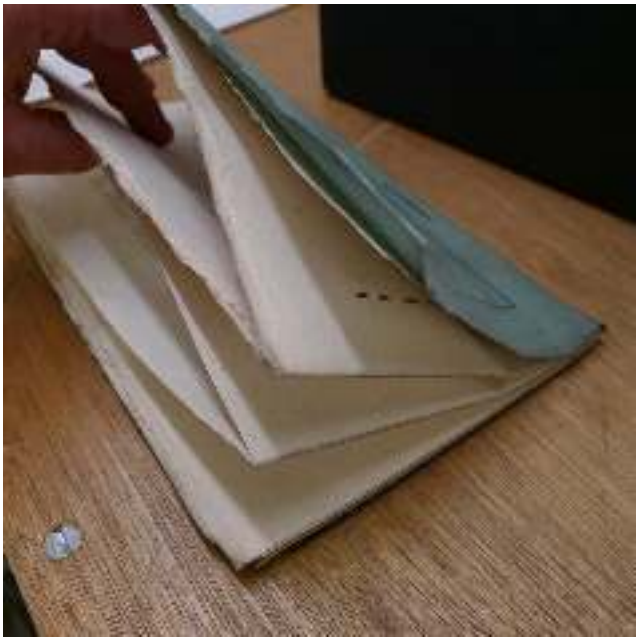

Fig. 4. Uncut pamphlet.

of resources but may also be for reasons relating to the historical value of the book as artifact, where treatment or other preparation for digitization would compromise the original attributes of the individual object and thus its research value. The uncut pamphlet in (figs. 3–4) is an example of an item that should not be cut to facilitate digitization. Such uncut items are invaluable for the study of publishing and printing history. This pamphlet can still be studied in the reading room as eyes can look around corners.

From a conservation perspective there are three methods of assessing the condition of bound and unbound materials: (1) a written condition assessment, (2) a random condition survey, and (3) for unbound collections typical in archives, a random visual survey. The approach taken will depend on the quantity of material surveyed. The time taken will vary according to the extent and nature of the collection in question. They can also determine the required resources.

1. A written condition assessment using an Excel spread sheet should be undertaken by a trained conservator at the initial scoping stage of the project. This is appropriate for a small project and produces a general scoping document — examining each item for its condition and opening angle and is limited to a project with fewer than a thousand items. This might typically take three to five days over a two to three week period.
2. A random condition survey, including a written report, should be performed by a trained conservator who determines the level of stabilization needed. Random surveys are appropriate for of a large collection, numbering a thousand items or more. It identifies specified conditions and the levels of stabilization required to ensure successful image capture. This can take three to five days over a two to three week period.
3. A random visual inspection is appropriate for archives which include both bound and unbound material housed together: paper; photographs; books; documents and letters. As the media are identified and scoped, and as the digitization program progresses, random visual inspection compiles information to highlight issues concerning the levels of pre-digitization stabilizing needed.

CONSERVATION AND COLLECTIONS CARE TEAM AND THE DIGITAL SERVICE OFFICER'S ROLE AND RESPONSIBILITIES

Conservation decisions and treatments are based on three axioms: minimal intervention (in the context of a research library); re-treatability; and the 'fit for purpose' principle. All treatments in preparation for digitization are undertaken with the aim of stabilizing objects for digital capture.

STABILIZATION LEVELS

Stabilization occurs before and during each project. The levels are based on the Library of Congress preservation guidelines (*see* bibliography) and have been established by the Conservation Department at the Wellcome Library. There are two levels of preparation for digitization with a view to stabilizing the object for image capture:

- *Level 1*: Basic preparation of both bound and unbound items for digitization is performed by trained digital preparation staff. For example, removing staples or pins where necessary, or opening folds or creases that interfere with image capture, dry cleaning materials, inserting brittle material into Mylar sleeves, or relaxing creased paper with a Teflon folder, if necessary with a tacking iron. These should be carried out close to the area of digital capture to allow ongoing discovery of preparation needed, as the items are digitized (figs. 5–9).
- *Level 2*: Technical conservation treatment of both bound and unbound items, usually performed by trained Conservators or trained Digitization Preparation staff. This treatment needs to be carried out in the conservation studio: for example, mending paper tears, disbinding bound collections, or humidification and flattening.

Responsibility for Stabilization Level 1—namely, pre-digitization preparation and preparation discovered during digital capture—will be extended to Digitization Preparation staff and Photographers, who have been trained by Conservation staff. Stabilization Level 2 will be undertaken by Conservators or trained Digitization Preparation staff in the conservation studio, except for items that are declared out of scope. Any digital preparation that involves moisture, e.g. disbinding or simple paste and Japanese tissue mends is done in the conservation studio and may be performed by trained staff and/or conservation staff.

Other ongoing and post-project conservation collection needs may be identified by Conservators, Digitization Preparation staff and Photographers during the project, for example, re-attaching loose or detached spine components or cover boards, or other more extensive repairs.

Materials requiring special attention are most likely to impact the project timeline. In some cases these are objects requiring Level 2 stabilization. They will be identified during the conservation condition assessment as to whether they will be included in the project or taken out of scope.

In most cases these objects *may* be prepared for digitization but they will need extra time and resource planning for both preparation and image capture. In all cases where a high level of intervention would be required for successful digitization a curatorial decision is required. The decision may also be to remove such items from the project's scope. Such special conditions in objects that may require immediate

Fig. 5. Staple removal: X-ray attached to paper.

Fig. 6. Bundles of Ticehurst certificates.

Fig. 7. Ticehurst certificates prepared for digitizing.

Fig. 8. Contents of each bundle placed in a folder for digitizing.

Fig. 9. Envelopes returned to archival box.

attention are mold and pests, vulnerable media and extreme fragility of substrates (figs. 10–11).

Surveys are needed to predict some of the preparation time and resources required for digitization projects. When whole books are uncut they will need to be cut by a digital preparer or conservation staff. This procedure is often carried out when a reader requests it in the reading room, but it is necessary to release these pages so that the book can be digitized. The conservation survey of five thousand books, in the Wellcome Library, dated 1850–1920 has suggested that there will be two hundred books that are uncut. It is estimated that this will take two to three hours per book, which will take upwards of ten weeks of work for the digital preparer over the two year project. It has been established that digital imagers are permitted to release up to five uncut pages with training in order that digitization can proceed without serious interruption (fig. 12).

In Phase 2, one of the projects, entitled "The Asylum and Beyond", will focus on mental hospital records in the 19th and 20th centuries. These are from the Ticehurst House Hospital collection founded in the 18th century and having

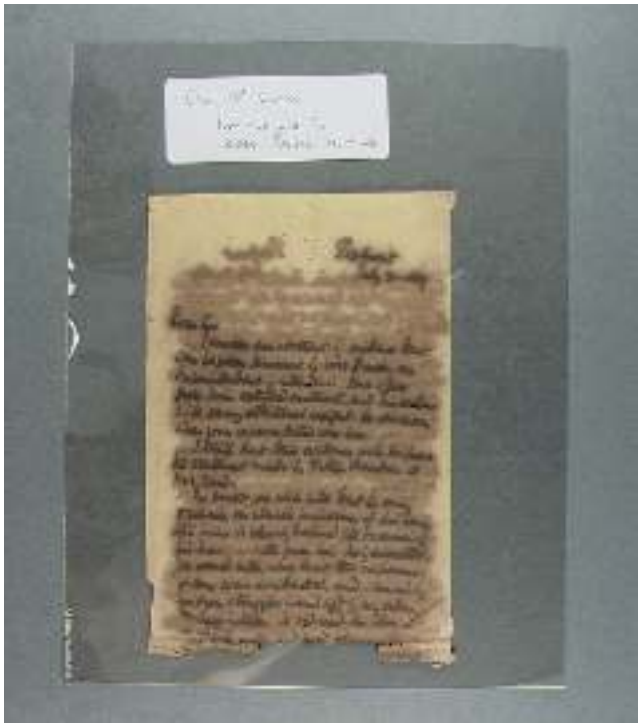

Fig. 10. Brittle material in polyester sleeve: out of scope.

Fig. 11. Brittle, torn pages: out of scope.

Fig. 12. Whole book with uncut pages.

Fig. 13. Ticehurst records purse removed.

Fig. 14. Ticehurst purse.

detailed patient records going back to this time. Digitization of the Ticehurst papers is estimated to take fifteen months.

Projects such as Medieval Manuscripts and 19th Century Books will require specialized preparation, identified as a result of the surveys. Before they are selected and delivered for digitization, they will need preparation. Archives are proving to be the most challenging material, as they typically need more preparation than a book. Archivists already prepare materials for handling in the reading room. They remove rusty pins and staples and replace them with brass paperclips; thus it has been established that such items can be prepared for digitization by trained staff (figs. 13–14).

COMMUNICATION

The Digital Service Officer (DSO), the digital operators and the conservation staff all communicate to establish best practice. There should be an understanding as to how and where damage is likely to occur with the items that are being digitized. The staff should all maintain an open channel of communication throughout, with a member of the Conservation and Collection Care Team assigned to each project.

With good training and communication, conservation concerns can be raised and communicated back to the conservator, before and during digitization. The conservation team working with the DSO maintains a continuous dialogue with our digital contractors to clarify our expectations on book handling, and to learn from them about their equipment and workflow. Both our digital preparers and our photographic staff understand careful handling of items during the digitization process. We have also included training in the use of basic conservation tools and straightforward tasks that save everyone's time, such as unfolding the corners of pages overlooked by the digital preparator. This means that digitizing proceeds without undue interruption. Digital Preparation is mostly a matter of stabilization to facilitate the digitizing process.

The relationship between the physical object and the photographic equipment being used should be well understood by Conservators, Digitization Preparation staff and Photographers. Instructions to the digital capture staff concentrate on the use of equipment in such a way that no harm comes to the items being digitized—for example, how to hold the cradle off the open pages being photographed so that there is no pressure on the book during the period when the imager needs to leave their set-up for a protracted break, such as to lunch or a meeting.

DIGITAL EQUIPMENT

There exists a wide variety of photographic equipment and associated methods of image capture (figs. 15–18).

When image capture is being undertaken, whether on- or off-site, specific criteria for the equipment used, in relation to the materials to be digitized, should be met:

- Image capture equipment should suit the items to be digitized with regards to size. Scanning beds and book rests must be as large as, or larger than, the item being imaged so that it can be supported and digitized safely. This particularly applies to large items such as fold-outs, maps and charts.
- Vulnerable items such as those with restricted openings (less than 90 degrees), or having fragile components, should be digitized using appropriate equipment such as the Conservation Cradle.
- Photographers should make efforts to minimize light and heat exposure from photographic equipment with respect to the items being digitized, and to minimize the amount and duration of the pressure a book is placed under during digitization.

It is important that Conservators review image capture procedures with photographers before a digital project begins. Even familiar equipment may perform differently with different items. Such a review will both enhance the safety of collections and increase the chances of successful image

Fig. 15. Internet Archive book cradle: 110 degree opening.

Fig. 16. A regular copy stand with 90–110 degree opening.

Fig. 17. Conservation Copy Stand 6545: less than 90 degree opening.

Fig. 18. Guardian Copy Stand System: 180 degree opening.

capture for high-risk items. Vulnerable bindings require specialized equipment such as the Conservation Copy Stand 6545 ('Conservation Cradle') (see fig. 17) or the copy stand setup (see fig. 16) albeit operating at a reduced rate of image capture. The Guardian copy stand is a more sensitive piece of equipment and is the recommended digitizing equipment for flat items or books that open 180 degrees. *(*see fig. 18). During digital capture the book is raised up to the glass by a foot pedal, but it can also be done mechanically, when a gentler movement is required. With other equipment by contrast the glass is brought down onto the object and flattens the item, rather than just touching it, as the Guardian does, with this equipment there is a faster rate of image capture.

ACKNOWLEDGEMENTS

I would like to acknowledge our Wellcome photographers for supplying images of their equipment and helping me understand the capabilities of various pieces of equipment. I would also like to acknowledge the Wellcome Conservation Team and the Digital Service Officer for their commitment to providing ongoing support for the digitization projects and for their contribution of certain images. We have grown in knowledge and understanding of all facets of digitization together.

BIBLIOGRAPHY

Bülow, A. and Ahmon, J. (2011), *Preparing Collections for Digitisation*, London: Facet Publishing.

Library of Congress, *Preservation Guidelines for Digitizing Library Materials*, Library of Congress Collections Care. Available online: http://www.loc.gov/preservation/care/scan.html

JISC, *Selection Procedures for Digitisation*, JISC Digital Media. Available online: http://www.jiscdigitalmedia.ac.uk/guide/selection-procedures-for-digitisation/

GILLIAN BOAL
Head of Conservation and Collection Care
Wellcome Library
London, United Kingdom
g.boal@wellcome.ac.uk