Collection Development, Cultural Heritage, and Digital Humanities

BOOK CONSERVATION AND DIGITIZATION

THE CHALLENGES OF DIALOGUE AND COLLABORATION

by

ALBERTO CAMPAGNOLO

ARC HUMANITIES PRESS

BOOK CONSERVATION AND DIGITIZATION

ARC HUMANITIES PRESS

COLLECTION DEVELOPMENT, CULTURAL HERITAGE, AND DIGITAL HUMANITIES

This exciting series publishes both monographs and edited thematic collections in the broad areas of cultural heritage, digital humanities, collecting and collections, public history and allied areas of applied humanities. The aim is to illustrate the impact of humanities research and in particular reflect the exciting new networks developing between researchers and the cultural sector, including archives, libraries and museums, media and the arts, cultural memory and heritage institutions, festivals and tourism, and public history.

FOR PRIVATE AND
NON-COMMERCIAL
USE ONLY

BOOK CONSERVATION AND DIGITIZATION

THE CHALLENGES OF DIALOGUE AND COLLABORATION

by

ALBERTO CAMPAGNOLO AND CONTRIBUTORS

ARC HUMANITIES PRESS

ARC HUMANITIES PRESS

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

© 2020, Arc Humanities Press, Leeds

The author asserts their moral right to be identified as the author of this work.

Permission to use brief excerpts from this work in scholarly and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is an exception or limitation covered by Article 5 of the European Union's Copyright Directive (2001/29/EC) or would be determined to be "fair use" under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L.94–553) does not require the Publisher's permission.

ISBN (print): 9781641890533
eISBN (PDF): 9781641890540

FOR PRIVATE AND NON-COMMERCIAL USE ONLY

www.arc-humanities.org
Printed and bound in the UK (by Lightning Source), USA (by Bookmasters), and elsewhere using print-on-demand technology.

CONTENTS

FOR PRIVATE AND NON-COMMERCIAL USE ONLY

ILLUSTRATIONS

Figures

ARC HUMANITIES PRESS

FOR PRIVATE AND NON-COMMERCIAL

ARC HUMANITIES PRESS

Lists

FOR PRIVATE AND NON-COMMERCIAL USE ONLY

Tables

ARC HUMANITIES PRESS

FOR PRIVATE AND
NON-COMMERCIAL
USE ONLY

INTRODUCTION

> *Like men, books have a soul and a body. With the soul, or the literary portion, we have nothing to do at present; the body, which is the outer frame or covering, and without which the inner would be unusable, is the special work of the binder.*
>
> *(William Blades,* The Enemies of Books*, 1902, p. 96)*

THE SUCCESSFUL TRANSMEDIATION of books and documents through digitization requires the synergetic partnership of many professional figures that have what may sometimes appear as conflicting goals at heart. On one side, there are those who look after the physical objects and strive to preserve them for the future generations—conservators and curators—and on the other those involved in the digitization of the objects, the information that they contain, and the management of the digital data—digitization professionals, and then digital humanists. These complementary activities are generally considered as separate, and when the current literature addresses both groups, it does so strictly within technical reports and guidelines, concentrating on procedures and optimal workflow, standards, and technical metadata. In particular, more often than not, conservation is presented as ancillary to digitization, with the role of the conservator restricted to the preparation of items for scanning, with no input into the digital product, and this leads to misunderstanding and clashes of interests.

Digitization projects have become increasingly crucial for memory institutions and cultural heritage in general and have been at the core of the field of digital humanities from its beginnings, inevitably with a profound influence on conservation practice since the early 2000s.[1] Here, as hinted at in the title, we will be concerned with the digitization of books and documents. All the fields involved in this diverse practice landscape, besides the specifics included in technical manuals in regard with best-practice procedures, have largely distinct authorship and readerships that do not overlap. This book tries to fill this gap. It strives to do so by showcasing, on the one hand, the need to understand the informational content of books as objects and the role that conservators could have in the creation of broader digital products and tools, and, on the other, the transformative and transcendent value that digital surrogates can and should bring to the table. The inclusion of a series of real-life case studies and expert contributions aims at highlighting both the conservator's and the digital humanist's point of view, keeping in mind the significance of both aspects of the conversation: the importance of the original object, and the merit and enhancing nature of digital surrogates.

The quote by Blades that opens this introductory chapter speaks of a soul and a body of books, the former being its (literary) content, and the latter being its physical form.

---

1 Kirschenbaum and Werner, "Digital Scholarship and Digital Studies," 417; Bülow and Ahmon, *Preparing Collections*, 18–20.

The physical form comprises the book's cover and bindings—as stated by Blades—but also the form and materials of its pages, inks, decorations, letter shapes, usage accretions, stains, and so on. While digitization tends to concentrate on the remediation of the content of books, "with the soul, or the literary portion, we have nothing to do at present," and we will instead argue for an increased interest in the transmediation of the body too, to the extent that is possible. When considering the digitization of the bodies of books, one has to keep in mind their specific physicality and materiality, and that, as Hayles illustrates, these are two related but distinct concepts.[2] While physicality is a permanent quality, an infinite set of physical attributes that make up the reality of the object, materiality, instead, is an emergent property that depends on the attention of some observer who isolates as meaningful some particular attributes, setting them aside from the continuum of physicality. Materiality results from human mediation and interpretation in an act aimed at the identification and discernment of information. Materiality—and its information—is extrapolated from the physicality and each aspect of it can be represented and manipulated, also digitally. This digital representation and manipulation of an object's materiality—or, indeed, materialities—is achieved through different means. Among these, metadata designation—certainly not limited to the digital realm—is one of the most established processes, one that transcends the object/information juxtaposition, one that can flatten hierarchies and bring objects and the information infused in their materiality onto the same level.[3]

The digitization of books is generally understood as the capture of the page contents through photography and imaging. Not all features of books can be digitally acquired in this manner—we will refer to these as *untransferable* characteristics— and models and descriptive metadata are necessary steps to computerize important information about the structure and materiality of documents. For this reason, we assign a broader meaning to the digitization (or computerization) of books. We include not only imaging activities—that is the scanning of books, page by page, to attain digital surrogate images—but also any action directed at the computerization and transmediation of books and their features, materiality included, into digital media, and the use of such data.

Digitization can do much more than reproducing books as texts to be read, and books are much more than flat sequences of pages: there is much information in books that has been largely ignored—so far. Other than showing sequences of pages of texts, digitization can make readable invisible texts or features, flatten irremediably distorted or rolled documents, rearrange pages beyond their physicality, and evidence the history and use of documents. Digital surrogates have the potency to be more than mere replacements of the original objects, and so only of use for a limited number of activities, as the term *surrogate* would suggest. Instead, when the transformative nature of the digitization process is more fully harnessed, they can become digital cultural objects: digital objects that transcend the originals, work in synergy with them, and

FOR PRIVATE AND NON-COMMERCIAL USE ONLY

2 Hayles, *How We Think*, 91–92; Shep, "Digital Materiality," 326.

3 Geismar, "Defining the Digital"; Geismar, *Museum Object Lessons*, 61.

make them something more. However, these advanced uses of digitization are only possible through synergetic collaboration among conservators, digital humanists, and all others involved, as this encourages connections and resource sharing to achieve genuinely innovative outcomes.

This book develops in three parts. The first part, "Books as Objects and Their Digitization," sets the scene by presenting the issues related to the understanding (or lack thereof) of the artifactual value of books as objects, in turn leading to the understanding of the delicate balance between the aims of book conservation and the needs of digitization endeavours. The first chapter, "Understanding the Artifactual Value of Books," illustrates how, despite the increasing interest in the book as an object and its artifactual value, probably linked to the distance that new technologies have allowed establishing between us and the object, books are generally black-boxed, without deep consideration for their inner working. The concept of the artifactual or intrinsic value of books is complex: it is an emergent quality linked to material features that depend on research interests and points of view. The idea of intrinsic value of documents in need of preservation came about when archives and libraries began to look into reformatting as a preservation technique. It became apparent that specific characteristics of the documents would not be easily (or at all) transferable into a new medium, leading to loss of information. These untransferable qualities of books are very much still relevant, even if digitization is not regarded as a preservation methodology any more, as, for the same reason, digital surrogates fail to represent most of the materiality of digitized documents. As a way to bring together the diverse landscape of research interest in our written heritage, we advance a framework for the study of the book as an object that develops along four research axes. The first focus is on the *materiality*, looking into raw materials, structure, and appearance. Then comes the *context*, focusing on the relationships of the object and with the object, and its *history*, with its functions and uses. Finally, we look into the *significance* of books, covering their psychological and academic values. Through these research patterns, and the data thus accrued, one can draw a series of conclusions that aid a more inclusive interpretation of the object. The model encourages a cohesive investigation into books as material objects, bringing to the foreground what observable data can be gathered from their physicality, and putting them in relationships with other objects and information sources, fostering the production of a network of comparative and supplementary data. In turn, these views can be put in relation with the book contents and applied to the wide-ranging understanding of books that is necessary for their conservation (balancing their value as objects, and their use as content delivery tools), and to produce more inclusive digital surrogates.

The second chapter, "Conservation and Digitization: A Difficult Balance?," showcases how conservation and digitization efforts relate to one another, and how they can work in synergy for the production of better digital representations of books. The balance between meaning and use often tips towards the latter, losing crucial historical evidence. The same is also mostly true when considering traditional routine digitization programs, whereby the sole scope of the reproduction seems to be the transfer of textual information into the new medium. Books, when regarded as archaeological objects, contain a lot of evidence, evidence that should at least be recorded (and transferred to

the digital products). Modern book conservation strives to safeguard evidence—while preserving use—and this is now transposed also into the digitization process, which can be seen as an intensive, and potentially damaging, use of the book. To curb this potential damage, conservators are more and more involved in digitization projects from the onset, guide in the selection of the equipment and its use, and help to preserve the documents by selecting materials that can undergo digitization, treating them, offering handling training, and maintaining an appropriate and stable environment. Digitization is, however, also an opportunity for the conservator, a chance to implement long-term conservation and preservation programs, for example, or a way to collect reliable and high-definition visual evidence to document the state of conservation of documents at a specific date; digitization can also help preserve the objects by limiting subsequent handling.

Most importantly, digitization also helps to achieve results that are not possible with the original document or with standard conservation, transcending, in this manner, the material object. In turn, conservators can offer a means to address the untransferability of material objects, by helping to describe, through pertinent metadata, untransferable features, and by bringing into the fold their expertise, knowledge, and understanding. In this way, in the cohesive view of the book as an object fostered by the application of the framework that we have advanced, conservation skills and knowledge can be leveraged to record and transfer untransferable features, adding precious data to the digital representation of books.

The second part, "Conservation and Digitization in Practice," portrays a series of real-life case studies aimed at showcasing the role of the conservator in digitization projects of diverse nature, from a varied set of institutions, encompassing libraries, archives, and universities. Each chapter introduces a particular digitization project with its challenges, goals, and achievements. The case studies have been selected to cover as many different situations as possible, providing examples of a full breadth of practices. We have examples of large-scale digitization projects in bigger institutions, both from the library and the archive world, with their specific problems. On the one hand the reader will find the narration of the challenge of digitizing the multitude of manuscripts in the collections of the Vatican Library, and on the other the piecemeal approach to conservation for digitization at The National Archives in London. Other case studies concentrate on collection-specific digitization in larger and smaller institutions, such as the Qatar Foundation Project at the British Library or the medieval manuscripts at the Wellcome Trust Library—both in London, half a mile apart, yet entirely different in their approaches and the solutions adopted for equally successful projects. Particularly important and difficult objects may need bespoke digitization efforts, such as the Great Parchment Book project at the London Metropolitan Archives, in collaboration with the UCL Centre for Digital Humanities (UCLDH). Similarly, research libraries need well-established good-practice standards, as exemplified by the account of the digitization endeavours at the Herzog August Bibliothek Wolfenbüttel, Germany. Finally, increasingly, specialized digitization techniques are been put into use in memory institutions worldwide for difficult-to-digitize features, as is the case of the methodology for the description of bookbinding structures and materials fostered by the Ligatus Research

FOR PRIVATE AND NON-COMMERCIAL

Centre of the University of the Arts London, and the multispectral imaging techniques showcased by Fenella France and Mike Toth.

The majority of the institutions and individuals involved are European and drawn in from personal connections because of the familiarity of this author with their work. This selection, however, was also strategic inasmuch as the main object of focus here is the book in codex format; in addition, the vast number of volumes in European collections, and the relative lack of funding, compared to North America for example, places these institutions in a situation where the balance between conservation and digitization is more challenging. Notwithstanding this, the experiences and lessons showcased here are transferable to other contexts, and Quandt's expert contribution balances the European focus, just as Terras's brings multispectral imaging within a European context.

The contributing authors were asked to think about projects they had been involved with, considering the issues and challenges that they faced in setting up the project. They were requested to consider the kind of amendments they had to perform to accommodate for the problems they encountered and the compromises that had to be established to ensure the beginning of a fruitful project.

In particular, the authors were to report on the kind of relationship that was established between the team of object experts (conservators, object curators, etc.) and that of the digital data experts (photographers, data curators, database engineers, online content managers, etc.), also thinking about how this changed during the project. Alongside this, the authors were to think about the relationships with other institutions, the practical solutions that had to be devised, and any outreach avenue that was established. Often, conservators reach out to colleagues in other institutions to seek advice, exchange information, and develop shared practices, especially looking for practical solutions to routine problems. It is also increasingly becoming customary to set up outreach activities, such as blog posts, to showcase the work of the conservator that would otherwise remain unseen and unnoticed. These allow the sharing of ideas and experience, and also offer an informal but public way to establish a unified narrative between all people involved in the digitization project, from the conservator to the digital data experts. An example of this is the Polonsky Foundation Digitization Project.[4]

Ángela Núñez Gaitán, head of conservation at the Vatican Apostolic Library, describes how the colossal endeavour to digitize their almost 82,000 manuscripts is linked to the Library's century-long tradition of preserving but also making freely available their textual heritage, in a balance of meaning and use. She puts the current digitization project into the perspective of a service that the Library has always pursued, having used photographic technologies since the nineteenth century, also as preservation techniques for the textual information of the manuscripts. In the past, conservators had been required to adapt old books to the modern user's needs—by disbinding and sewing on tapes, for example—but this kind of invasive treatment, while certainly making reproduction easier, as it allows books to be opened at 180 degrees, inevitably results in the loss of historical evidence. Still, the needs of such a large-scale digitization project meant that

ARC HUMANITIES PRESS

4 Polonsky Foundation, "Polonsky Foundation Digitization Project."

the Library had to select high-production scanners (capable of capturing two pages at once) that force the books to open at 180 degrees, or 130 degrees with the insertion of supports. Even if some items have been set aside, waiting for more sympathetic technologies to become available, as is often the case, these scanning issues created a conflict between the conservator and photographer, as the preservation concerns of the former clashed with the need to produce the highest image quality possible of the latter. The conflict was resolved through constant dialogue and collaboration among conservators, digitizers, and IT specialists, helped by an integrated management system, specific handling sessions, and the production of a "technical phrasebook" to create a common language between the different teams. Items are routinely assessed before and after digitization and prepared for reproduction with minimal conservation treatments. The conservators then realized that the digitization process could be considered as one extensive use of the manuscript. Also, the high quality of the digitized images could mean that the object would then be handled less since the majority of the scholars should be satisfied with the surrogate, as textual features are preserved and enhanced. Even items with compromised legibility, in fact, can be read thanks to spectral imaging. Indeed, no digitization is a complete representation of the original, since the historical and cultural aspects and the materiality of books are inevitably left aside. These are better represented by metadata, but due to the high volume of manuscripts, it was not thought feasible, in this project, to record exhaustive metadata sets for each manuscript. Even the records in the conservation database are reduced to the bare minimum.

Catt Thompson-Baum, formerly senior conservation manager of the digitization services at The National Archives in London (now an independent accredited conservation manager and collections consultant in the UK), recounts the history of digital surrogate creation at The National Archives (TNA) in London. She describes the beginning of the large-scale digitization programs and the fundamental role of the conservator as fully integrated within the workflow. TNA have a long history of reformatting, on microfilm first, and then through digitization. Given the nature of the collections, family historians are an important category of users, and with them come the commercial companies who manage family history resources. As a result, many mass digitization projects are funded and run by external companies. Making all stakeholders understand the value of conservation was a challenge at the beginning, sometimes only implemented as an afterthought. A first attempt at integrating conservation and digitization required in-house conservators also to manage commercial digitization needs. The situation was stretching the Archives' resources, and a new approach was devised whereby the new role of the Digitization Support Conservator was established within the Preservation team, to manage large-scale projects, build relationships with all digitization stakeholders, and carry out assessment surveys and conservation for smaller collections. Conservation was now always included in any project from the beginning. A second radical change was the move of the digitization conservation team away from collection care and into the commercial services department, with all other commercial digitization stakeholders. This inherently embedded conservation into the digitization team, in a separate conservation studio, next to the digitizers, and made "conservation for digitization" non-negotiable and funded by the commercial partners. The new setting allowed stronger relationships to

FOR PRIVATE AND NON-COMMERCIAL

be built so that everybody involved understood the needs and work of the others. For example, conservators are required to sit with the digitizers and even try their procedures to get a deeper understanding of what it means to handle documents for digitization, thus informing their conservation work. A set of conservation guidelines were put together to help conservators and digitizers understand what kind of work was deemed necessary and sufficient. After condition assessment, trusting the handling training provided to all digitizers, conservation treatments only aim at stabilizing documents and guaranteeing maximum legibility, easier handling, and capture, with minimal documentation. The experience accrued and the success rate has meant that the Archives have become a leader in conservation for digitization, confirming the success of their model.

Flavio Marzo, formerly conservation manager for the British Library/Qatar Foundation Partnership (now managing conservator at the Cambridge Colleges' Conservation Consortium, Cambridge, UK), reports on the Partnership's digitization of materials relating to the Gulf history and Arabic science. He focuses on the role that the modern conservator has to take, shifting away from the bench and becoming an integral part of the digitization workflow in all areas of interest. Marzo describes the development of the digitization workflow, highlighting the importance of the involvement of Conservation from the planning stage and the communication and sharing of experiences with other conservators who had previously been involved in large-scale digitization projects. There are many stakeholders involved in these projects, all coming with different expertise, knowledge, and jargon. It is, therefore, essential to foster a shared understanding of each other's role, duties, and language. In this case, a koine was produced with the writing of a "Guidelines for Conservation" document, the painstaking process of devising the menus of the joint project management software, and handling training sessions. The integration process was also facilitated mainly by the physical working proximity of all parties involved. A secondary effect of the "Guidelines" was the fact that, by standardizing conservation treatments, these allowed the keeping of minimal and efficient condition assessment and conservation records (understandable by everyone, within the project). Given this close shared understanding of the objects' and project's needs, exceptional cases could be treated successfully on a flag-up basis, and handling and bound-volume opening concerns could be communicated and addressed without problems or damage. Even if the conservation records were necessarily minimal, a particular field in the database allowed the provision of technical descriptions of specific binding and structural features that could then be used for catalogue entries by some of the curators. This project makes available an enormous amount of records that were previously difficult to access, but the digitized documents are not the sole digital output. Through the website, blogs, and microblogging accounts, all stakeholders, with their set of specialized knowledge—including conservation with video tutorials and articles about the history of bookbindings, its craft, and related topics—have engaged in outreach activities, thus enriching the online surrogates with content pieces to help online users to better understand and enjoy the material.

Stefania Signorello, a senior book conservator, covers the digitization of the 335 medieval Western manuscripts at the Wellcome Collection. Illustrating the development of the reformatting workflow, she touches on the significant research potential that is

embedded in the collection, and on how only through collaborations and partnerships is it possible to properly take advantage of it, advancing the knowledge about these artifacts. In the beginning, the project was run in batches of manuscripts but, to facilitate the work of the photographer, it became clear this approach had to be changed, thus allowing the digitizers to organize their work through selecting different kinds of materials. The books are digitized by the Library's photographers, already known to all parties involved for their trustworthiness when it comes to the preservation needs of the materials. Therefore, there was no need to establish new relationships and a common language; rather, because of this, the conservators, in their pre-digitization surveys, were able to flag up for conservation presence during scanning and for pre-digitization treatments strictly when unavoidable. When necessary, conservation treatments were minimal so as not to cancel or hide any evidence. To avoid obscuring any information, it was preferred to digitize before treatment. The digital images would then act as a reference to control deterioration in the future. The survey, besides immediate conservation needs, covered foliation, binding issues, opening angle for bound volumes, after-digitization restrictions, and re-housing needs. Despite the full range of supports and digitization infrastructures (which included a conservation book cradle), some manuscripts had to be excluded, for the time being, from reproduction, because of their fragility and difficult opening characteristics. Due to the small nature of the collection, the pre-scanning survey was a particularly rich experience, as it allowed the establishment of a long-term preservation and conservation program, but also the noting of distinctive or unusual features as areas of potential research (such as candidates for multispectral imaging). The digitized manuscripts have become part of a vibrant digital information ecosystem published on the website, whereby users can be guided in the use of the collections through blog series or can experience materials enhanced by videos and other curated interfaces.

Almuth Corbach, head of collection care and conservation at the Herzog August Bibliothek (HAB) Wolfenbüttel (Lower Saxony, Germany), describes in detail the kind of damage that can occur to books during digitization, and the tools and practices to avoid them. The HAB has a long history of reproduction, and in the past, books had been forced open at 180 degrees to be photographed and microfilmed. Today, all books are assessed by a senior conservator before digitization. About 10,000 books are assessed each year, and the number is steadily increasing. The guiding principle of the assessment is that no evidence can be lost during reproduction and that no loss of information is permitted in the digital copy (such as incomplete capture due to very narrow gutters). About 30 percent of assessed items are flagged up as not to be digitized for the time being. In order to permit digitization, limiting damage to the minimum, a series of new tools have been developed and acquired to adapt as much as possible to the various materials and their needs. A "Checklist on Assessment for Digitization" has been produced to standardize the selection criteria and as a communication tool. The opening characteristics of books depend on many aspects of the object, such as the presence of decorative elements and titling, spine movement, and joint stability. These are assessed pre-digitization, and specific digitization tools are thus selected to address and respect them. New tools include: the Wolfenbüttel Book Reflector (for capturing at angles as

FOR PRIVATE AND NON-COMMERCIAL USE ONLY

narrow as 45 degrees); the Cobra Book Scanner (which allows two-page scanning); the Grazer Conservation Cradle (to capture without contact with glass for illuminated manuscripts and other fragile materials); the Traveller's Conservation Copy Stand (for small formats and ease of transportation) with its companion, the Austrian Watermark Imaging System (an infrared photography system for watermark capture in transmitted light); and a versatile divided table for large formats. Some items still require special handling, and ad hoc supports with the presence of a conservator. All information acquired during the assessment is recorded in a database. Finally, Corbach highlights how digitization can be seen as an occasion to focus on preventing and mitigating damage, degradation, and loss in the collections as part of preventive conservation activities. Digitization can also protect from constant handling, and therefore be regarded as a preservation measure, but it is the reformatting procedures that need to adapt to the demands of the object and not the other way around.

Caroline De Stefani, conservation studio manager, and Philippa Smith, head of collections, from the London Metropolitan Archives in London, illustrate the work done, in partnership with UCLDH, on the Great Parchment Book (GPB), highlighting the fundamental role of the conservation prior to photographic capture to reveal obscured parts of the document and on the powerful response of the project's outreach avenues. The GPB is an important document containing the 1639 survey of all the estates in Derry~Londonderry managed by the City of London. The volume was caught in the 1786 Guildhall fire, with devastating results due to the heat and the subsequent water damage: 165 sheets survived, but in a very poor state of conservation, with general concave distortions, and very deep text-obscuring creases and folds. The sheets are now loose, and current knowledge does not allow establishing its former codex structure. The document was not accessible to scholars for over 200 years. Traditional conservation could not flatten the membranes without further damage, and traditional digitization approaches could not cope with the deep creases and three-dimensional distortions. To tackle the problem, the project invested in a partnership between different disciplines (archivists, paleographers, conservators, computer scientists, digital humanists). It was decided to attempt only minimal conservation treatments to ease deep creases and reveal any hidden text, while a novel digitization method, a multi-view stereo technique, was adapted and devised to cope with the distorted topography of the membranes. The process entailed taking between fifty and sixty images per side and processing them to obtain texture maps and virtual reconstruction. Subsequently, a virtual flattening algorithm computed a map to flatten the three-dimensional surface onto a 2D plane with as little distortion as possible. The results were incredibly useful, to the point that the document could be withdrawn from access (the virtual surrogate being much more informative than the original, for what concerns the textual content). The project's success was only possible thanks to the constant communication and relationships between all stakeholders. The project had also a very successful outreach program, with blog posts and videos, and all data (before and after virtual flattening images, and TEI transcriptions) was published on a dedicated website. The GPB project is a clear example of how digital surrogates can transcend and make the original usable, not merely replace them as secondary copies.

ARC HUMANITIES PRESS

The last three contributions of this part cover a series of projects that can be accounted under the remits of digitization, as understood in the broader sense of the computerization of books outlined above, applied to the cultural heritage sector. The contribution by Athanasios Velios and Nicholas Pickwoad, directors of the Ligatus Research Centre of the University of the Arts London, showcases the potential of semantic web technologies for the preservation and digitization of information regarding books as artifacts. One of the reasons that have curbed proper bookbinding descriptions within catalogues is the lack of a consistent and recognized terminology. This makes the production of relevant literature and other descriptions, such as within conservation documentation, challenging, but it is even more problematic for the creation of database schemas. While traditional free-text documentation can be suitable for single-object records, schemas, databases, and controlled vocabularies are a better solution to document entire collections, as free text limits information retrieval and multi-language access. In turn, the integration of different databases requires a more flexible approach. One solution is mapping the database schema to an abstract model (to which all databases of the same domain can relate) and referring to a universally agreed (domain-based) system of concepts (often published in the form of thesauri). The Language of Bindings (LoB) thesaurus, developed by Ligatus, attempts to foster this kind of solution, by advancing a description system based on semantic web technologies (RDF and SKOS) to implement database integration through the CIDOC-CRM (an ontology explicitly developed for the cultural heritage domain). Binding descriptions should arrange data according to strict book-structural principles, following the sequence of operations necessary to bind it, in a hierarchical approach, from top concepts (endleaves, structures, spines, boards, endbands, covers, decoration, fastenings and furniture, enclosures) to narrower and more precise ones. The LoB is organized precisely in this manner, through SKOS, in hierarchies, and following the CIDOC-CRM. With the help of domain experts, the core terms have been selected from the concepts currently in use in bookbinding documentation. Each concept is assigned a preferred label (in a language), a scope note, and a stable and permanent URI, ready for semantic web applications, also through a SPARQL endpoint. The LoB has been included in the thesauri to be used in MARC bibliographic records. It is therefore hoped that with this new technology and reference model, the discipline may at last reach a mature stage, and more compatible descriptions and schemas may be implemented and queried.

The last two contributions by Fenella France, chief of the Preservation Research and Testing Division of the Library of Congress (Washington, DC, USA), and Michael B. Toth, president and chief technology officer at R. B. Toth Associates (Virginia, USA), draw attention to the implementation of multispectral imaging technologies. France reports on the potential that these cutting-edge digital technologies offer to conservators. Traditionally, spectral imaging has been applied to palimpsests or underdrawings for the recovery of hidden texts and graphics. However, through close collaboration between preservation professionals, conservators, and curators, this imaging technique can be used for much more. Spectral imaging works by measuring the reflectance of materials to assess their chemical composition. Through these measurements, spectral imaging can be used to track change over time, assess the efficacy of conservation treatments,

FOR PRIVATE AND NON-COMMERCIAL

characterize and identify inks and pigments, study data stratigraphy in early printing techniques, and for the recovery of hidden or faded text. At the Library of Congress, spectral imaging is implemented as a first investigative tool. Objects are imaged with a standard LED illumination at 45 degrees to map the spectral response of inks and colourants, characterize stains or discolorations, and look at layers. Raking light illumination at 15 degrees helps to capture the morphology of a document, and transmitted light can be used to capture watermarks, text within laminated layers, and detect treatment areas. In addition, the detection of fluorescence, a secondary response to light whereby energy is re-emitted at a lower frequency, can help in identify deterioration areas, and assess organic colourants. Capturing is only one part of spectral imaging. Data thus acquired can be processed (generally through PCA) to further reveal features and to map and visualize spectral responses through false-colour images that, for example, differentiate between features that may look the same to the naked eye, such as two different inks, or forgeries. Mapping can also highlight areas in need of further study with other analytical techniques (XRF, FTIR, FORS, etc.). Spectral imaging is also an excellent technique to archive data on the materiality of objects (at a specific date) and to enhance conservation documentation. The proper file formats need to be used (TIFF) to ensure archival quality data. For the same reason, comprehensive metadata need to be saved and embedded in the images. Also, because of the large nature of the files, any spectral imaging endeavour requires clear and firm relationships with the IT department of an institution to warrant long-term preservation and access to machines with enough processing power.

Toth addresses the issues and challenges related to the use of multispectral imaging on special collections and high-profile items. The pioneering Archimedes Palimpsest Project involved a close relationship with conservation from its onset, setting the de facto standard for imaging following preservation and conservation guidelines, enforcing safe handling, adequate environmental conditions at all stages—scanning included—and minimum light exposure. This also included the very specialized imaging at the Stanford Synchrotron Radiation Lightsource, Stanford Linear Accelerator Center (SLAC), to penetrate the paint layers on forged leaves to reveal the undertext. The Archimedes Palimpsest Project saw the continuous development of the prototype multispectral imaging, explicitly developed for the manuscript since previous applications of the technology had not involved books and documents, but satellite imaging and astronomical phenomena. A second system, utilizing LEDs, was subsequently developed and the manuscript reimaged. The new system, now in its fourth generation, displays an important conservation characteristic: it does not generate heat, and it exposes documents to narrowband light sources, thus limiting light damage, despite the use of UV light sources. It was calculated that an imaging session would be equivalent to a couple of days in exhibition conditions. The St. Catherine's Palimpsests Project did not include collaboration with conservators but was able to capitalize on the years-long conservation survey run by Ligatus on the same collection. The processed data has now been released, but only with the undertext in mind, and no images highlighting conservation issues or areas of conservation interest have been produced and released. The imaging of the Syriac Galen Palimpsest, fragmentary and dispersed in the collections of five different institutions, on the other hand, proved successful thanks to the close relationship between all the

stakeholders involved: scientists, curators, managers, and conservators. Toth also illustrates how image processing can be accomplished with open-source software, and a tool specifically developed to render the operation easily implemented by curators and conservators. As a result, multispectral images are beginning to be considered as information sources by some conservation scientists and conservators to support long-term preservation efforts.

The last part, "Conservators and Digitization Experts in Dialogue," showcases two contributions from two prominent figures in the fields of book conservation and digital humanities. One portrays the point of view of a conservator who has had considerable hands-on experience in state-of-the-art digitization projects of tremendously important (and valuable) manuscripts, the other that of a digital humanist who has been involved in cutting-edge digitization of cultural heritage artifacts for a significant part of her career.

The authors describe the role that they have played in several digitization projects and how they perceive the relationship between the two fields. They focus on what would not have been possible without the active collaboration and dialogue between the two fields, considering whether current training allows practitioners of both fields to understand each other, and how they see things changing in the future.

Abigail Quandt, head of book and paper conservation at the Walters Art Museum (Baltimore, USA) and lead conservator of the Archimedes and the Syriac Galen Palimpsest projects, presents the book conservation point of view by reporting on over a decade of digitization projects at the Walters. On top of being a world-leading conservator, Quandt is also very well acquainted with digitization and multispectral imaging efforts. The Walters Art Museum, beginning in the late 1990s under the leadership of Will Noel, has pioneered the digitization of manuscripts and their digital access. Thanks to a series of National Endowment for the Humanities (NEH) grants, the Walters has acquired specialized digitization equipment and run a series of successful manuscript digitization projects that involved close collaboration between the conservation department, the curators, and the digitization specialists. Special care was observed from the outset in the adaptation of the equipment, guaranteeing adequate environmental conditions, and in working closely with the digitizers, offering support and handling training. Having bought all the necessary equipment with the first grant, the following funding was able to cover a part-time project conservator to survey and prepare the manuscripts for digitization. Eventually, the Walters agreed to absorb into the institutional budget the two digitizers, the maintenance of the equipment, and the processing and archiving of the data, while one conservator position was partially dedicated to supporting the ongoing digitization programs. An ad hoc database was prepared to survey the manuscripts and to record materials and binding information, current condition, past treatment history, current conservation needs, and minimal documentation. Given the particular nature of the material, aside from the usual repairs limited to tears, losses, and marginal binding stabilization, an essential aspect of the pre-digitization conservation was the painstaking illumination treatment to secure flaking pigments and inks. At the Walters, all aspects that could be captured photographically are generally included in the digital surrogates, including stubs, endleaves, bookmarks and other inserted material, dealer's tickets, and covers—both current, and earlier detached bindings, if available. For those

FOR PRIVATE AND NON-COMMERCIAL USE

items that cannot be safely imaged, the choice is to digitize only what is possible, from just the binding (if they do not open safely) to partial textblocks. Quandt also speaks of the difficulties in funding project conservators and the consequences of this limitation and criticizes that fact that, despite their in-depth knowledge of the items and their materials and structures, conservators are very rarely involved in the documentation and cataloguing of books being digitized. It is hoped that the situation may change in the future considering how conservators are now routinely trained in the needs of digitization and should, therefore, understand the value of metadata as a means to add their insights to the digital sphere.

Melissa Terras, Professor of Digital Cultural Heritage at the University of Edinburgh's College of Arts, Humanities, and Social Sciences, and former director of the UCL Centre for Digital Humanities, represents the goals and expectations of a digital humanist, without forgetting the importance of the original artifacts. Terras is a leading researcher in digital humanities of worldwide fame. She teaches and researches in the digitization of cultural heritage materials in the library, archive, and museum sector, and has been involved in several high-profile projects that required active collaboration with conservation practitioners, which she describes in her contribution. While most memory institutions now aim to create digital representations of documents, artifacts, and objects to improve access and foster understanding of their material—with plans to digitize 52 percent of collections across Europe—there are opportunities for advanced imaging to analyze primary sources, beyond routine reformatting projects. Indeed, as routine digitization requires conservation, and can thus be an occasion to implement preservation measures on the collections, conservators can also highlight which documents may benefit from advanced imaging to go beyond the limits of physical conservation, and help understanding materiality issues, accessing the material, drawing attention to how new methods and practices can be best applied. Advanced imaging projects depend on the synergy between imaging scientists, conservators, historians, and digital humanists, to develop suitable digital imaging approaches. In this manner, scientific imaging technologies can answer research questions about our cultural heritage and develop best-practice methods and recipes for specific tasks, so that others could then successfully apply these techniques. Inter- and cross-disciplinary research and endeavours of this kind suffer from frequent issues, such as lack of a common language and understanding (or appreciation) of disciplinary-specific knowledge, skill, methodologies, and tools, with distinct difficulties in engaging—institutionally—within liminal spaces. These can be overcome through dialogue and networking, leading to the clear benefits of developing ways to optimize methods and technologies and a solid awareness of other disciplines. For the future, the next generation of researchers and practitioners should gain the appropriate skills that are needed to work across a broader range of digital (and non-digital) applications.

ARC HUMANITIES PRESS

Throughout the text, both in the introductory chapters and in the case studies and the expert contributions, a series of main recurring themes are advanced repeatedly. We find good-practice procedures and desiderata, as well as issues related to the delicate balance of meaning and use as specifically applied to the digital and digitization. There are also reflections on the positive and the problematic aspects of the reformatting of

books, including the transcendence of the digital in relation to the readability of text and the tackling of conservation problems not safely solvable through traditional approaches. The importance of documentation and metadata is often underlined, in addition to the intrinsic and untransferable value of books, and the problems related to the digitization of the materiality of these objects. Finally, evident is the need for inter- and cross-disciplinary approaches and implementations for successful digitization projects. Such a liminal practice, however, brings a particular set of issues and challenges that can be overcome only through dialogue and synergy among all parties.

When these differences are overcome, and effective collaboration takes place, as noted in the Coda, a new kind of digital product emerges. The original item becomes the common ground on which all stakeholders base their investigations and reformatting processes, leading to the production of digital cultural heritage objects that are not mere reproductions of the artifacts. Considering the emergent quality of materiality, these are never conceived as complete or final. Instead, they are capable of integrating more information as it becomes available, cumulatively, and, in making evident elusive features, create a new balance between the physical and the digital, whereby one informs the other in a recursive manner. It is this new balance that all stakeholders involved in the digitization of books should aim to produce. It is this new balance that this volume strives to foster.

Acknowledgements

The initial seed for this book came at the suggestion of Dymphna Evans from Arc Humanities Press to publish a monograph based on the work I presented at the inaugural session of the UCLDH seminar series in 2014, "Scalpels and Magic Wands: From Physical World of Book Conservation to the Magic of Digital Humanities," by invitation of Professor Simon Mahony, then program director for UCL's MA/MSc in Digital Humanities, now director of the UCL Centre for Digital Humanities. My deepest gratitude goes to both, for their support and encouragement.

I would like to thank the authors of the case studies and the expert contributions for their input to this volume. It is their hands-on experience that makes this volume valuable, and that has informed and (re-)shaped my own understanding of the equilibrium between book conservation and digitization. My sincere appreciation also goes to the reviewers of the book proposal and the manuscript for their careful and focused indications on how to improve the text. Any interpretative shortcoming remains mine.

FOR PRIVATE AND
NON-COMMERCIAL
USE ONLY