Digitizing and Archiving of All Information Taken From Rare Blockbooks

Peter Meinlschmidt; Wilhelm-Klauditz-Institut, Fraunhofer-Institute for Wood Research; Braunschweig, Germany
Markus Brantl; Munich DigitiZation Center (MDZ)/Digital Library, Bavarian State Library (BSB); Munich, Germany
Irmhild Schäfer; Institute of Book and Manuscript Conservation (IBR), Bavarian State Library (BSB); Munich, Germany
Bettina Wagner; Department of Manuscripts and Early Printed Books, Bavarian State Library (BSB); Munich, Germany
Rahel Bacher; Department of Manuscripts and Early Printed Books, Bavarian State Library (BSB); Munich, Germany
Volker Märgner; Institute for Communications Technology, Technische Universitaet Braunschweig; Braunschweig, Germany

Abstract

Xylographic books of the 15th century, the so-called blockbooks are objects of extreme rarity and considerable value. Due to their age and as a result of paper corrosion caused by the copper content of coloring pigments, some books are in rather poor condition.

Digitization opens up new perspectives for research in such rare and fragile materials, as many of the surviving blockbooks have never been reproduced in print. Furthermore, the technology of digital infrared photography for the first time allows a comprehensive documentation of the watermarks which occur in the paper on which blockbooks were printed.

Introduction

Xylographic books produced in the 15th century, the so-called blockbooks, are exceedingly rare specimens of our cultural heritage. Only about 600 such books have survived worldwide. In the early 20th century, scholars distinguished 100 different editions containing 33 different works – ranging from the 'Bible of the Poor', biblical texts and religious tracts in Latin and German to works for everyday use, e.g. calendars, grammars (the 'Donatuses'), travel instructions (the 'Mirabilia Romae') and even a chiromantic treatise.

A common feature of blockbooks is not only the technology of printing from woodcuts with which they were produced, but also the often close combination of text and images: most blockbooks are illustrated or, rather, consist of illustrations to which explanatory texts have been added as part of the same woodcut. Thus, blockbooks can be seen as an intermediary between illustrated manuscripts and incunabula. Like manuscripts, they allow texts and images to be combined on a page in a very flexible manner, and like incunabula, blockbooks can be produced in large numbers and almost identical copies. While the type forms from which incunabula were produced were as a rule taken apart immediately after printing, the woodcuts used for blockbooks could be stored for many years, and new copies could be produced with little effort – an early form of 'publishing on demand'. For this reason, most blockbooks do not contain any information about the date of printing, and copies of the same edition can be printed on different paper. Often, the woodcuts deteriorated over long periods of storage or use, and sometimes deliberate changes were made for later impressions, which allow scholars to establish a relative chronology by comparing the surviving copies.

For these reasons, more precise conclusions regarding the year of production can often only be drawn from the dating of the watermarks of the papers used.

Figure 1. German blockbook of about 1470: Page 7v of the "Bible of the Poor" with a scene from the New Testament (Christ's baptism) surrounded by scenes from the Old Testament prefiguring it (urn:nbn:de:bvb:12-bsb00038196-0).

For a long time, it was assumed that blockbooks were a precursor of books printed with movable type, being produced with a more 'archaic' technology. This assumption has been proved wrong, and we now know that the production of xylographic books extended over the entire second half of the 15th century, thus coinciding with the incunable period. Accordingly,

blockbooks were an alternative to books produced with movable type: the technology was better suited for particular types of illustrated texts which were in high and continuous demand as well as limited in size (most blockbooks consist of a maximum of 40–60 pages), and blockbooks may well have been intended for particular audiences: for clerics at the propaedeutic levels of religious instruction or lay readers with limited needs for pragmatic books rather than for theologians or university students and teachers. These questions, however, can only be examined once a comprehensive survey of the surviving copies has been established. Yet even then their provenances may not yield more general conclusions, as most blockbooks we know today have been preserved in monastic libraries while a large proportion of the books owned privately by lay people or parish priests have perished. Thus, until today, blockbooks raise a wide range of research questions, with regard to the origin, date, the organization of the production and distribution, and their readership. In order to address these, both access to full reproductions of the surviving copies and detailed descriptions are required. Yet only a few of the major collections of blockbooks have been cataloged so far. Among them, the collections of the Bibliothèque de France in Paris (54 items) and the British Library London (40 items) stand out. Their holdings of blockbooks have been described in the course of long-term enterprises to catalogue incunabula. This is also the case for some Bavarian collections, for which a series of incunabula catalogues have been produced since the 1960s. The catalogue of the largest collection, held at the Bayerische Staatsbibliothek (Bavarian State Library) in Munich and comprising more than 20,000 books printed in the 15th century, was only completed in the year 2000, with two index volumes published after that date. However, it did not contain descriptions of the 48 blockbooks preserved in the library, as this was beyond the scope of the project. Therefore, a separate project to catalogue the blockbooks was initiated after completion of the incunable catalogue, and the availability of state-of-the-art digitization technology at the library opened up perspectives to answer diverse needs: to make the blockbooks accessible to researchers worldwide, and to gain new bibliographic insights on the basis of high-quality reproductions, also of the watermarks. In order to broaden the range of the material subject to this detailed analysis, all institutions in Bavaria in which blockbooks are preserved were invited to join the project.

As a result, over 90 blockbooks from 14 different institutions, extending from university and state libraries to church and private collections as well as museums, are now accessible online.

The paper is organized as follows. In the next section an insight into the scanning project is given. A short introduction into the infrared technique for watermark detection follows. Then the scanning workflow is presented in detail. Finally, the paper concludes with a discussion of the results.

The scanning Project

Since 2009, the Bavarian State Library (BSB) in Munich [1, 2] has been carrying out a project funded by the Deutsche Forschungsgemeinschaft (DFG) which focuses on digitizing and archiving of xylographic books of the 15th century, the so-called blockbooks.

Up to now, these rare and valuable specimens of early printing were accessible only to a small and selected group of scholars. Digitization opens up new perspectives for research in such rare and fragile materials. Many of the surviving blockbooks have never been reproduced in print and thus could not be analyzed and compared with each other in detail. Furthermore, the new technology of digital infrared imaging allows the first time a comprehensive documentation of textual content together with watermarks which occur in the paper on which blockbooks were reproduced.

After two years of work on the project, the online presentation of the BSB comprises high-quality images of more than 90 blockbooks. All parts of the blockbooks are reproduced and archived, including the bindings (covers, pastedowns, flyleaves) and empty verso pages – a feature typical for blockbooks. On an internet site [3], the digital images are displayed in a standard viewer developed with support of the DFG which allows easy navigation and offers important additional functions, like an electronic table of contents, access to catalogue records and a download of images as a PDF document. Infrared images of the watermarks have been made available on the same site for more than 50 copies so far. The digital photographs were taken with equipment invented and developed by the Fraunhofer-Institute for Wood Research (WKI) and the Technische Universitaet Braunschweig, Institute for Communications Technology (IFN) in Braunschweig and tailored to the specific needs of blockbooks, by using a book cradle to ensure that the books are opened at an angle hardly wider than 90°, and by controlling the exposure to infrared radiation during the imaging process [4-6]. The system consists of a metal plate heated to 30-50° Celsius and a digital infrared camera which transmits images directly to a computer; the leaves are placed between those two devices. For every leaf, a series of images are taken within milliseconds, and successively, the image with the best contrast is selected for analysis, manual enhancement and storage. Even for leaves which have been glued together with a thickness up to 0.35 mm, the process has generated impressive results. For the first time, the watermarks of the blockbooks in Bavarian collections have been documented comprehensively, which greatly facilitates comparison and identification and forms an important preliminary step for cataloguing.

Infrared technique for watermark detection

At first glance, the visualizing of watermarks in paper appears to be a very simple task. If the paper is placed between a light source and the observer, the watermark can be seen instantly in this transmitted light. For digitization, the observer's eye simply needs to be replaced by with a digital camera.

This technique works satisfactorily as long as the paper is sufficiently transparent and the variations in density are rather small to allow recognition of the watermarks contours. But the limitations of the technique are revealed once the surface of the paper is covered with different types of inks or drawings or the paper reaches a certain thickness.

In order to overcome the problem of handwriting or drawings on the paper surface interfering with the visibility of the watermark in transmitted light, a solution is to use a certain part of the spectrum in which the writings and drawings are invisible. Tests in the near-infrared (NIR - 700nm to 3µm) and in the mid-infrared spectrum (MWIR - 3µm to 7µm) have shown much better watermark appearance in MWIR than in NIR.

The new infrared technique for visualizing watermarks was invented in close collaboration between the Fraunhofer Institute for Wood Research (WKI) and the Institute for Communications Technology (IfN) of the Technical University of Braunschweig [7-8].

This method is based on the one hand on the fact that most of the different types of ink are invisible in the infrared wavelengths and on the other hand on the certainty that the transmitted infrared radiation is scattered or absorbed differently by the paper and the watermark.

The process is as follows. Infrared images are taken using a stirling cooled sensor with 384 x 288 pixels and a temperature resolution of better than 20 mK. To obtain a homogeneous infrared radiation from the rear, a copper plate 3 mm thick and painted black is heated from behind by means of a rubber heating mat glued to the back of the metal sheet (Figure 2, left). A piece of passe-partout about 10 mm thick is placed on the metal (Figure 2, center) to secure a specific distance between the paper under investigation (Figure 2, right) and the heat source.

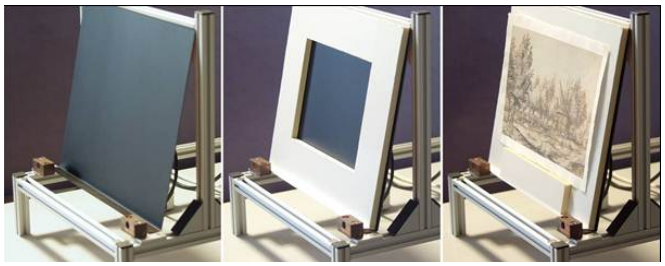

Figure 2. *A copper plate painted black and at a temperature of about 40°C serves as an infrared emitter (left), while a piece of passe-partout (center) secures a certain distance from the valuable drawing (right).*

With the camera in front of the system and focused on the paper, high qualitative watermark images can be obtained within seconds as shown in Figure 3.

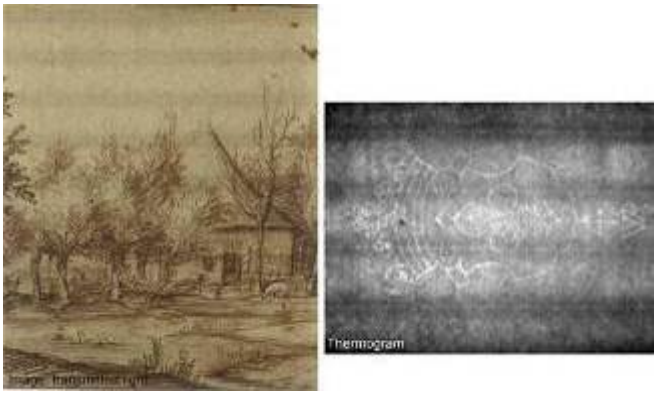

Figure 3. *The photograph on the left shows the drawing 'Farm and trees by the water', pen and brown ink, made around 1665 by Jan Lievens, Inv. No. Z 103 (copyright: HAUM, Claus Cordes). On the right the watermark emblem can be seen in the infrared image without interference from the drawing.*

Figure 3 shows a photograph of a drawing in brownish ink made around 1665 by Jan Lievens [9]. At the time the watermark was taken it was not quite clear whether the drawing really belonged to the original school of Rembrandt or was a later reproduction of the 18th century [10]. Recognition of the emblem proved that the paper and therefore the drawing was an original.

The new device for watermark detection

For the reproduction and digitization of watermarks within the blockbook project, a special setup was developed consisting mainly of a book cradle, an infrared camera and a warm blackened copper plate (Figure 4). The book holder or cradle can be adapted to almost every size and angle of books while the heated copper plate has to be brought in a save position during turning of the book pages.

Figure 4. *Set-up for IR-watermark detection with a book cradle that can be adapted to sensitive books and with the heated copper plate in its highest position for choosing the page under investigation. The camera (orange) is focused on the plane were finally the paper under investigation will be.*

The infrared camera is focused on the plane were finally the paper under investigation will be. Special attention has to be taken that good images of watermarks can be captured even if the opening angle of the book is sometimes only 60°. For security reasons, the black phyroelectric detector shown in Figure 4 in front of the camera controls the maximum paper temperature.

Figure 5: *Detail of the setup showing that every single leaf of the book will be held flat in a certain position by a vacuum bar.*

If the book is in position and a leaf under investigation is chosen, a vacuum will be applied to a perforated flat rod to hold the paper in a straight position in relation to the camera. Infrared

images or videos can now be taken always at the same distance and therefore with the same scale.

Workflow for the scanning of blockbooks

The scanning of the blockbooks was done at the BSB in the Scanning Center of the Munich DigitiZation Center (MDZ). The scanning process is a sub-process within the entire digitization workflow.

The digitization workflow consists of the following major process steps:

- Preparation of the books and setups
- Scanning, i.e. image capture with different scanning devices; in the blockbooks project there was an additional process with the thermographic scanning device for the image capture of the watermarks
- Indexing or production of metadata (enrichment), including administrative, bibliographical, technical and structural (incl. full-text) metadata
- Storage and long-term preservation
- Access, to support the material-specific presentation on the internet; nowadays ready for multi-devices, like smartphones, tablets and desktops.

The individual processes always depend on the structure and composition of the original object and the intended presentation mode on the internet (e.g., browsing, search, 3D animation) [11]. The empirical formula is: the older an object is or the more metadata it needs - for example, information for search and retrieval - the more complex and costly the entire production process will be. To standardize the production processes and ensure a consistent production of digital data, the Central Digital Asset Management System (ZEND), a software tool with different document and workflow management components was developed by the MDZ in 2003. ZEND controls the entire production process from the preparation up to the automated transfer of the digital master files to archival storage.

The Scanning Center of the MDZ is today equipped with 22 different scanning devices, including book scanners and special equipment scanners (including four automatic book scanners, a thermography scanner for watermarks, and a 3D scanner for book covers), which are able to scan format sizes up to DIN A0 standard (e.g., historical maps) with high optical resolution and different book cradles in order to avoid any damages by mechanical handling of the books.

Images of the blockbooks were produced with the following parameters:

- Optical resolution of 400 ppi, relative to the original size of the object (1:1; due to the different sizes of the book scanners)
- Digital master storage format for all images, even watermarks: Tagged Image File Format (TIFF) (version 6.0) uncompressed. This results in color scanning file sizes of 20 to 800 megabytes per image, depending on the format, the size of the book, and the image resolution
- Color depth of 24-bit or 8-bit for watermarks
- Color management and media neutral image production: the entire production is based on the use of a color management system (in RGB and LAB as transformations color space), i.e. scanners and monitors are color calibrated.
- Targets for visual control of color, grayscale, depth of sharpness and scale are scanned once per object and stored with the digital master of the object. For watermarks there was built up a special copper target.

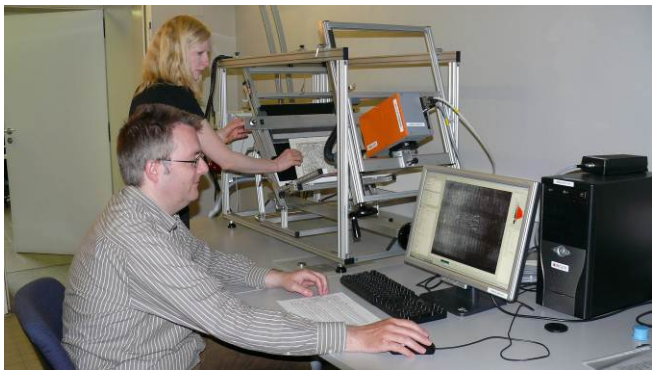

Figure 6: *Photo which was taken during the scanning campaign at the BSB.*

The process step of the scanning set-up was doubled:

1. For the color production of the pages and
2. For the thermographic watermark reproduction (see Figure 6).

Depending on the value, size, and condition of a document, the set-up on the selected scanner can take from a few minutes up to an hour. In the case of the blockbooks it was a time-consuming process, which was accompanied by a conservator. It consists of the following steps:

- Position the book on the book cradle of the scanner
- Scan targets
- Constantly re-adjust the depth of focus, due to the varying thickness of the book as the pages are turned. This is inevitable when working without a glass plate. Only the constant readjustment of the book cradle height will compensate for the changing focus.
- Separately scan recto and verso pages that are linked afterwards by the software (this is a critical, error-prone process due to the fact that the books often lack a pagination), if required due to the scanning device and opening angle of the book.
- Scan the book covers (out- and inside)
- Perform “what you see is what you get” (WYSIWYG)-quality control throughout the scan using the monitor
- Remove the book from the cradle after finishing the work.

During the digitization process it is essential to save the generated digital master files and their derivatives, like images, metadata and if available full text, within an appropriate storage solution as the primary storage capability. After the completion of the digitization process, the files will be organized and managed in two copies and in two separate archive systems, with the secondary and tertiary data storage serving as long-term preservation. This was covered by ZEND workflow.

An automated process is used to register the files, and a semi-automated process to generate the metadata for the digital objects:

1. The digital master files (e.g. watermarks) are separated accordingly to their ZEND-ID and all images of an object are stored in their own directory.
2. The master TIFF files are then converted for standard Web presentation into a JPEG format with two different resolution versions.
3. Next the simple structural metadata of the object is generated in TEI-XML to describe the logical and physical structure of the object. With this description, the book can be browsed later and edited online with the ZEND Table-of-Contents Editor.

As a last step in digital preservation, the digital master files for images and other metadata are transferred in their long-term stable file formats (TIFF and XML) to long-term archival systems in two setups: to a Network Attached Storage System (NAS), which currently has 220 terabytes of net capacity and to the Tape Library of the Leibniz Supercomputing Center in Munich-Garching, which is running under IBM Tivoli Storage Manager software. The files are also stored in another copy in another location.

As soon as the final quality assurance in the ZEND-system has been carried out, the object will be published on the internet immediately. An automatic message is generated saying that the object is ready on the Web and available under its persistent internet address, which is sent to the local library system online public access catalog (OPAC) as well as to other reference systems (union catalog, European, etc.). As a result, the digital object on the internet can be accessed in various and multiple ways, for example via the Digital Collections of BSB/MDZ, the Bayerische Landesbibliothek Online (regional portal for Bavarian history), catalog systems, search engines, etc. In addition, a free and full pdf download is possible for every object, which enables the data transfer to a user-PC and facilitates offline work. Furthermore, the bibliographical metadata can be exported for re-use in other information environments, for example via an OAI interface to the European or the World Digital Library. An integrated and structured presentation of the digitized blockbooks together with their watermarks is available under http://www.bayerische-landesbibliothek-online.de/xylographa-sammlungen.

Some of the results can be seen in Figure 7 to 9:

Figure 7. Page 4 of the blockbook "Biblia pauperum - Nördlingen, 1470" (front side, left and back side, right) and its corresponding watermark (urn:nbn:de:bvb:12-bsb00038197-5).

Figure 7 shows the photo of the front (left) and the back (right) side of leaf 4 of the blockbook "Biblia pauperum" which was produced in 1470 in Nördlingen. Obviously the colored print was done only on the front side of a relatively thin paper with an accordingly qualitative good watermark.

Figure 8 shows the photo of the front (left) and the back (right) side of the blockbooks page 25 of the "Hartlieb, Johannes: Chiromantie" coming from Augsburg which was presumably produced in 1485/95. Obviously the black and white prints on both sides were done on separated paper sheets and later glued together which leads to a more or less good watermark.

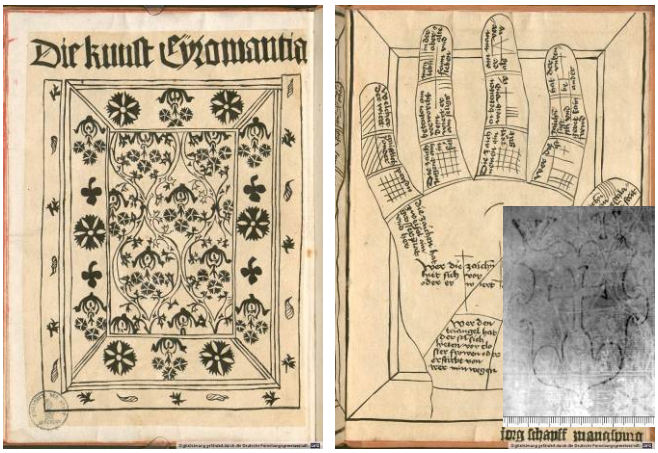

Figure 8. Page 25 of the blockbook "Hartlieb, Johannes: Chiromantie Augsburg, about 1485/95?" (front side, left and back side, right) and its corresponding watermark (urn:nbn:de:bvb:12-bsb00043466-3).

Figure 9 shows the photo of the front (left) and the back (right) side of the blockbooks page 19 of the "Biblia pauperum" coming from Germany which was produced in 1470. Obviously the colored prints were produced on separated paper sheets and later glued together which leads to a poor quality watermark.

Figure 9. Page 19 of the blockbook "Biblia pauperum" which was produced in Nördlingen in 1470 (front side, left and back side, right) and its corresponding watermark (urn:nbn:de:bvb:12-bsb00038196-0).

Conclusion

This paper gives an introduction into xylographic books of the 15th century (blockbooks), a very rare and therefore valuable item of the European cultural heritage. The workflow of the blockbook scanning project carried out at the BSB is explained in detail and both the procedures used in the digitization of the book pages as well as the infrared watermark detection are explained.

In conclusion, the quality of different digitized blockbook pages as well as its corresponding watermarks will show the efficiency of the project.

References

[1] Griebel, R.; Ceynowa, K.; Haller, K.; Information, Innovation, Inspiration, 450 Jahre Bayerische Staatsbibliothek, Berlin, New York 2008.

[2] Wolf-Klostermann, T. and Reiner, B. : How to Cope with 300,000 Scans a Day: Managing large scale digital collections in practice - the Bavarian State Library and the Leibniz Supercomputing Centre Approach the Next Level of Mass Digitisation; in: Archiving 2008. Final Program and Proceedings, Bern, 2008, p. 272-274.

[3] Internet site containing all digitized blockbooks and watermarks: http://daten.digitale-sammlungen.de/~db/ausgaben/uni_ausgabe.html?projekt=1236933450&recherche=ja&ordnung=sig or http://www.bayerische-landesbibliothek-online.de/xylographa

[4] Meinlschmidt, P.; Märgner V.: Multispectral Image Archiving of Watermarks in Historical Papers, In Proceedings of IS&T Archiving Conference, S. 92-96, Salt Lake City, USA, Mai 2011.

[5] http://www.digitale-sammlungen.de/index.html?c=digitalisierung-scanner&l=de.

[6] B. Wagner, R. Bacher, Blockbooks in Bavarian Collections: A New Project for the Digitization and Analysis of Early Printed Books, Microform & Digitization Review, MDR, Vol. 40, pp. 18–24, March 2011.

[7] Meinlschmidt, P.; Märgner V.: Erkennung von Dicken- und Dichteunterschieden in transparenten und halb-transparenten Materialien mittels Thermographie, Patentanmeldung.

[8] Meinlschmidt, P.: Untersuchung der Papierstruktur mit Hilfe der Infrarot-Thermographie. International DFG Conference „Zur Praxis der Bearbeitung von mittelalterlichen Handschriften“, Berlin Germany, Staatsbibliothek zu Berlin.

[9] Döhring, T.: Aus Rembrandts Kreis -Die Zeichnungen des Braunschweiger Kupferstichkabinetts- (2006).

[10] Laurentius, T.: Het onderzoek naar Rembrandt's papier -The investigation into Rembrandt's paper. In: Voelbaar papier : papierkunst in Nederland - Tactile paper. Houten : Ekspress-Zo 1996, 112-119.

[11] Brantl, M. and Schäfer, I., Challenges and Experiences in the Mass digitization of manuscripts and rare books at the Bavarian State Library, in: Eikonopoiia, Digital imaging of ancient textual heritage (Proceedings of the international conference Helsinki 28-29 October 2010), ed. Vahtikari, V. et al., in: Commentationes Humanarum Litterarum 129), ed. The Finnish Society of Sciences and Letters (Helsinki 2011) S. 223-249.

Author Biography

Peter Meinlschmidt studied physics at the universities of Hamburg and Oldenburg (Germany) and at Towson State University (USA). He received his M.Sc. degree in 1991. Since 1996 he has been with the Fraunhofer Institute of Wood Research, Braunschweig, working on NDE applications of thermography to wood and wood-based materials. His current research interests include optical metrology, thermographic techniques and spectral imaging.

Dr. Markus Brantl studied historical auxiliary sciences with focus in advanced computing in the humanities, PhD 1994, since 1997 programmer/project manager at the newly founded Munich DigitiZation Center (MDZ)/Digital Library of the Bavarian State Library; since 2003 director of the MDZ, particularly research interest in mass-digitization, digital long-term preservation, 3D, gesture-based computing.

Dr. Irmhild Schäfer studied historical auxiliary sciences, PhD 1994 in historical bookbinding materials and techniques, since 2002 director of the Institute of Book and Manuscript Conservation (IBR), Bavarian State Library, Munich, and chair of the BA/MA conservation program, section "Book and Paper", in an institutional cooperation with the Technische Universität Munich (TUM), since 2009.

Dr. Bettina Wagner studied mediaeval literature at the universities of Würzburg and Oxford, PhD 1993, and has been working in the area of rare books and manuscripts at the Bodleian Library, University of Oxford, and the Bavarian State Library at Munich, where she has been head of the section for manuscript cataloguing and incunabula since 2005.

Dr. Rahel Bacher studied historical auxiliary sciences and German at the universities of Tübingen and Oxford, PhD 2009, and has been working on the blockbook project since 2010.

Dr. Volker Märgner received his diploma and doctorate degrees in electrical engineering from the Technische Universitaet Braunschweig (Germany) in 1974 and 1983 respectively. His main areas of research are image processing and pattern recognition methods using visible light and infrared.