Digital Images: A Valuable Scholar's Tool or Misleading Material?

Patricia Engel[†]

† European Research Centre for Book and Paper Conservation-Restoration, 3580 Horn, Austria.
ercbookpaper(at)gmail.com

Abstract: *The contribution looks at the various methods of digitization applied to our written cultural heritage today and the varying intents laying at the foundation of their application. It shows the different ways of using the data collected through the digitization process and points both at the benefits offered by digital images and the verifiable loss of information in a digital image compared to the original. The three case studies discussed in the paper seek to prove the author's principal thesis: "We are only able to preserve what we know and understand, while the original can hold its secrets still waiting for us to comprehend them."*

Keywords: *Digitization, Manuscript, Facsimile.*

I. INTRODUCTION

The Digitization is the transformation of information into a binary code and saving it in this form. In today´s library and archive practice, "digitization" is the procedure of taking and keeping images of historical texts and pictures. Digitization can serve various purposes and the resulting images can be used in various ways. The type of the digital image produced, the intended purpose of digitization and the intended use of the digital image can all influence the original.

The choice of original documents for digitization may be due to the fact that the artifacts are highly valuable or stored in a remote place, with scholars having virtually no access to them. Another typical reason for the choice is our fear that we might lose some artifacts completely due to their rapid decay. These are just three possible reasons for taking the digitization path, and, of course, there may be many more.

The types of digitization process, as well as the ways of handling the original before digitization and processing the data after the picture-taking proper reflect the perceived needs and intents underlying the decision to take the digital image of the object.

However, due to varying intended purposes of the digitization undertaken, as well as varying levels of awareness of the digitization risks (not only the physical damage to books which may occur as a result of their transportation, handling and reshelving, nor the fact that digital data also need conservation measures) digitization can often produce very diverse results.

II. MERE TEXT?

One of the core interests of Vestigia Manuscript Research Centre is to find traces of texts, save them and subsequently work with them. This idea is described on the Centre's webpage as follows[91]:

"*The object of our research are manuscripts and collections of manuscripts which are made accessible for scientific analysis. Because VESTIGIA cannot deal with all imaginable manuscripts and collections, three criteria guide our selection of material:*

• *Access to so far unknown or inaccessible collections, mainly in Southern and Eastern Europe. In this geographic area, smaller and larger collections have become accessible after the fall of political and ideological boundaries. Often, they are in private hands, but many are also owned by the Church. Anyway, a large section of the source material is theological in content.*

• *Indexing of individual manuscripts and collections that have been unnoticed because of their language.*

• *Indexing of palimpsests and fragments of manuscripts of mostly theological content which so far have gone unnoticed.*

The material is collected in the research centre, but also locally. Our excellent equipment makes it possible to work on manuscripts and fragments that are too delicate to transport."

(http://www.vestigia.at/e/erfassung.html)

While the last sentence concerns physical properties of the heritage, the three criteria mentioned before clearly indicate that images are made of texts which were unreachable so far due to geographical, political or language barriers. Thanks to the Centre's work, the material can now be read from digital pictures and from any place in the world. There is no doubt that this is a big step forward for researchers.

"*Apart from an analysis of the content of the sources, we are also concerned with their conservation (stabilisation, restoration) and digitization. Endangered sensible cultural values are thus conserved, and made accessible for research and scientific analysis.*"

The author of this paper knows that Vestigia Centre cooperates with some of the best Austrian conservators and really takes care of the physical state of the originals in the course of digitization. The digitization process, for example, is performed with a special device, which does not stress the books when texts are photographed. Furthermore training of local conservators is undertaken providing them with latest information and material, as well as the Vestigia´s own conservation research results.

As the linguistic competences of the VESTIGIA

---

[91] http://www.vestigia.at/e/verarbeitung.html

Manuscript Research Center covers Syrian and Armenian languages, one of the projects of Vestigia that we take as an example in this paper was entitled "Syrian Fragments in Armenian Manuscripts". Syrian manuscripts or fragments of Syrian texts belong to the category of "forgotten" material due to their language or their remote storage place. The project involved all manuscripts of the Matenadaran Research Centre in Yerevan and of the collection at Echmiatsin in Armenia. The project's result will be a catalogue of all Syrian texts held in the Matenadaran collection.

Once discovered in books, even small fragments of texts are digitized. The digitization is done relatively quickly. In Matenadaran, as well as in other places of storage, Vestigia often deals with collections of virtually holy character. Its experts are able to work in conditions where the permission to see books is given at rare intervals and the allowed time is short. Under these constraints, the picture taking process has to be as uncomplicated as possible. Metadata are always limited to the context of the particular find. No material survey (such as analysis of the sort of paper, parchment or inks) is undertaken, nor any background literature research conducted (in cases where only fragments of books, rather than entire codices, are concerned, as is often the case in Vestigia projects, no description of binding or the like is relevant anyway). Once back at their home base, the scientists have ample time and background literature to work on the images collected. Thus, the main idea is to collect as much material as possible in situ, go through books which might contain traces, find and identify interesting fragments and take pictures properly (which means, among other things, showing the utmost care in handling the respective books). The pictures must show the text, the context and contain comparatively few metadata. In most cases the latter are kept on the immediate image itself. For example, before a book is digitized, Vestigia expert takes a photo of its signature written on a plain piece of paper (the automatic order given by the camera allows to find each signature before each series of pictures of the particular texts and the signature is associated with the respective texts). Or, to understand the dimensions of the text, which may be important for identifying a certain handwriting, a measured ruler is photographed with the text.

The main idea behind the Vestigia's digitization enterprise is to record information which is only accessible for a short time and to help in subsequent deeper research.

The procedure developed also aims at minimizing the risk of loss of information and the most immediate metadata. Its current speed and efficiency help save time - and thus money. By an overall assessment, while ensuring safe handling of originals, the digitization process developed by Vestigia Centre promotes sharing of knowledge, which makes it an integrated conservation procedure[92].

Another technique of taking textual images and thus keeping the respective information is the so-called "Sicherheitsverfilmung". The technology involves filming newspapers and other types of text on wood pulp paper carriers, where the paper is so brittle that it can literally break down to smaller pieces at the touch, with the resulting loss of the textual information. In some countries conservators try to rescue these carriers themselves as well, in other countries they are just photographed and the original are then disposed of.

Allthough the approach used by these two techniques is quite different (while the Vestigia project is implemented in remote places, the Sicherheitsverfilmung technique is employed at some of the world's largest libraries), both of them rely on images with well-readable text. There are no special techniques of imaging used to bring to light more details but the text.

III. THE TEXT AND THE METADATA – THE WISH TO HAVE IT ALL

A somewhat different approach involves an attempt to retrieve as much information from an artifact as possible, not only by making series of different pictures but also by accompanying them with as much metadata as is deemed meaningful. All these possibilities – along with all the potential offered by the world of virtual technologies – were employed in the next example discussed below, the Codex Sinaiticus Project[93].

"Codex Sinaiticus, a manuscript of the Christian Bible written in the middle of the 4th century AD, contains the earliest complete copy of the Christian New Testament. The hand-written text is in Greek[94]. Four institutions are holding parts of the manuscript today, the Holy Monastery of the God-Trodden Mount Sinai (the Saint Catherine's Monastery), The British Library, the National Library of Russia in St. Petersburg and the Leipzig University Library. The virtual "consolidation" of the codex is one of the big chances for scholars provided by digitization.

Digital technologies can offer many more advantages in addition to the above. Close comparison of texts, transcription, translation, comments and finally the description of the physical aspects of the codex can be immediately linked to images – not to mention the new ease of Internet access for members of the scholars' community, no matter where they are in the world, or the fantastic opportunities for searches that can now be made on historical texts, allowing us to find all relevant passages in no time, something unthinkable on a handwritten text itself.

In the Codex Sinaiticus project, pictures were taken both in normal light and in transmitted and raking light,

---

[92] (Information partly taken from http://www.vestigia.at/syriacfragments.html)
[93] http://www.codex-sinaiticus.net/en/manuscript.aspx
[94] http://www.codex-sinaiticus.net/en/codex/

in different wave lengths, in inverted colours, and so on. Transmitted light shows the structure of the carrier, but also that of inks and colours. Dyes, being colours, allow light to get through much more readily than pigments (which are also colours, but more substantial, with more "body"). A trained conservator can "read" an image in transmitted light much like a trained medical expert can "read" an X-ray image. Raking light images reveal the surface structure characteristics, such as, for example, tensions created by binding media and thus causing the carrier to contract, individual surface features due to the particular methods of deleting, as well as many other types of valuable clues. Magnification offered by high resolution digital images reveals microstructures of material "by the click of the mouse" which would otherwise be accessible only with the help of a powerful microscope. In case of parchment, for example, it allows us to see the hair pattern, in case of an ink we can estimate the composition, and so on.

Knowledge of the material composition leads us to better understanding of workshop traditions, the writers' philosophy, the trading habits, the ways of exchanging knowledge, etc. Where pictures are not sufficient, description can assist. Virtually everything can be described. Such is the material, the techniques, the state, structures of bindings partly under the surface etc.

As a general rule, where analysis is possible, results of such analysis of carriers and writing material (such as the data of spectroscopic, EDX or radiological surveys of colours, inks, coatings or the carrier itself, observations concerning the layer sequences and so on) can and should be added to digital images as metadata.

The difficulty of recording all the details that can be "brought to light" with respect to the material is shown by the fact that in the case of a survey of the books of Saint Catherine's monastery it took the team of well-experienced conservators three years to complete the forms they had to fill in for each codex, describing the binding techniques and other details on the codices' material at hand (Pickwoad, 2004). The complexity we face in real bindings is normally much larger than what we can imagine in advance.

However, the more we come to understand the structures of bookbinding worldwide, the more likely it will become for conservators not to forget important details. Bookbinding and preservation of bookbinding as conservation problems have been neglected over the centuries and have only recently been shifting into the research focus. Experts knowledgeable of this area and currently available would be urgently needed to create a sort of a standard or guidelines on how to take down notes on bookbinding both in the course of recording details coming with any particular conservation procedure and in the course of digitization. EU programmes supporting this development would be well-invested money and help understand an important niche in our heritage. This was also clearly demonstrated by the outcomes of the conference "New Approaches to Book and Paper Conservation" held in May 2011 in Horn, Austria. One of the most frequently mentioned and urgently needed research topics was "Book Binding/Archeology of the Book".

Along with papers like the E.C.C.O. Professional Guidelines95, a paper detailing a standard way of describing binding techniques with a relevant form to be completed, however, would then have to be translated into all European languages, including rare languages like Armenian or Georgian.

The Codex Sinaiticus digitization project seems to be a perfect digitization project in many respects: not only in the sense that the cooperation with all the involved parties is going on at the highest level of humanitarian and scientific research, but also in the sense that the fruits the scholars brought home will meet their needs for information in a substantial degree. Some of the spin-offs of the project included standardization of some technical book conservation terms, the increased awareness about the value of the books in question, as well as offering a good example of appropriate integration of all the collected data.

And still, there are limits to what all this vast information can amount to, as opposed to the knowledge that can be derived from an original codex. These limits can be summarized by an important observation: "we can only see what we know already".

The fact that we can only register what we know is emphasized by the latest findings concerning Islamic manuscript structures published by Scheper (Scheper, 2011). The seam on the spine created by the use of two pieces of leather being used to cover the upper and lower board separately has not been noticed so far as a characteristic feature of these bindings, although it was open to any eye to understand it for centuries. It was not only the fact that it is difficult to detect the joining of leather pieces, as the joint is made so well and is sometimes hidden under decoration. It was simply not understood that this detail might be a "signal element" for a peculiar type of binding.

Capitalizing on her ample knowledge about these bindings, Scheper understood what others before her might have seen but misinterpreted as something of no particular value in terms of information. On a photo, this area must be depicted separately and properly focused, as otherwise it will not be seen on an image with a general overview. Moreover, it can only be described if the person providing the description knows that this feature is of particular interest and is not just an accidental cut, as one seen on an individual item. Of course, a good database of digital images of this kind leaves space for adding in new information of the kind Scheper found. However, this means revisiting the collection, which, at least, means extra cost in terms of time and effort.

95(http://www.ecco-eu.org/about-e.c.c.o./professional-guidelines.html)

IV. AS CLOSE IMAGE AS POSSIBLE - BUT COULD IT BE MISLEADING?

Facsimile prints are intended to give their user – or, rather, their owner – the "feel" of the original itself. This is especially true for private owners. The techniques of picture-making and reproduction have been becoming more and more elaborate and having such facsimile books is not only entertainment and joy: they also comprise lots of information. Apart from the immediate availability of the text and the opportunity to see miniature paintings for stylistic comparison without any risk of damage to the original due to the need of handling and climate change between storage area and reading room, it is possible to draw a great number of conclusions from facsimile prints even about the material itself. Furthermore, they preserve perfectly well the state of the codex at a certain moment of its life and thus provide the conservator with a good tool to prove if the codex continued to decay or was stable. When the facsimile is compared with the original book after a decade or two, it can clearly be seen if ink or copper corrosion increased, if silver leaf tarnish continued or dyes faded, pigments flaked, and so on. On the other hand, how misleading this sort of information can be and to what degree some types of information can only be derived from the original and cannot be clearly understood from the facsimile image, is demonstrated by a comparison of the Reichenau Pericope book Cod. Guelf. 84.5 Aug. 2° with its facsimile edition by ADEVA. First, there are inaccuracies in the depiction of the binding: the original has got a protective zig zag fault, which is not reproduced in the facsimile. Secondly, the facsimile image misses information which goes beyond the obvious things, due to the impossibility to observe the material in transmitted and raking light, as the original is of parchment, while the copy is obviously printed on modern paper. Thirdly, with the facsimile there is no opportunity to use high magnification to see the thickness of layers, the pentimenti, which can only be seen in the transmitted light. The facsimile picture only reproduces the surface itself and shows the sort of parchment as hair holes are depicted. Among other things, the fact that the holes in the cover boards of the Pericope book Cod. Guelf. 84.5 Aug. 2° are very much like those of the Hildesheim manuscript Dombibliothek Hs 688 led to the conclusion that these manuscripts are closely related. The boards of the Wolfenbüttel manuscript carry recesses for fixing the clasps, which were definitely a later feature. However, nails from the older binding went so deeply into the wood that they left holes in the remaining board. They match in number, size, position and spacing against each other with what we find until today in the Hildesheim manuscript (Engel and Gallistl, 2009) While they bear essential information, these holes are not there in the facsimile, although other holes for the former metal covering are reproduced on it.

This might be treated as neglectable, however, such neglect would be fatal in any scholarly approach. The second detail is much more astonishing. Miniature paintings in the Cod. Guelf. 84.5 Aug. 2° are or were protected by veils. Fig. 8b on page 173 of the Wolfenbütteler Beiträge shows page fol. 4r with stitches on the reverse side holding the protective textile. There are threads and several unused stitch holes, possibly from early sewing, as the textiles might have become broken and were exchanged. In the facsimile there are neither holes nor threads, which means they must have been retouched after the photo was made, because it is hardly imaginable they were not on the photo.

To sum up, a facsimile provides us with a valuable tool up to a certain moment. We still cannot avoid the trip to the original and the study of it when it comes to a real survey. In a way, we could even say that facsimiles mean loss of information, namely, when they tempt us to believe we have a true copy of the original. But the fact that they are not true copies is not due to technical limits, but to editor´s thinking, as we have seen in the example above.

V. CONCLUSIONS

Digitization is performed, and the resulting images are used, for various purposes, which obviously influence the original.

The very fact that a particular text is digitized, while other texts are not, adds to both, gives new dimension and quality to both the original text and its digital replica. The decision to digitize (or otherwise) a text may be due to the fact that it is considered particularly important, to perceived danger of loss or other reasons. Whatever the reason behind it, any digitization procedure is preceded by interpretation and categorization. This type of categorization may influence the entire future of the original.

Furthermore, the sort of data taken as digital image and the amount and quality of metadata could be called "another interpretation" of the material and, therefore, of the human work, thinking and history behind the artifact. The virtually unlimited technical possibilities tempt us to think that we can "store it all". This is definitely a risky misinterpretation of our situation. Apart from the fact that the long-time durability of digital data is by no means satisfactory, there is one final observation to be made:

As previously with hand-made copies, which were a precaution against any cases where information could be lost due to complete destruction of the original because of war, fire or other circumstances (Fischer, 2009), digital images can also help convey information about things we have understood. But what we did not understand cannot be registered by us, neither in an image nor by description. This is true also for the image, because it is us who set the camera, choose the focus and thus make a certain and unavoidably subjective image of the material. For a digital image, with or without metadata, it holds true: we can only keep what we know. In contrast to that, the original keeps all its secrets and waits until we see them.

REFERENCES

Engel, P. and B., Gallistl, "Die Reichenauer Handschriften der Dombibliothek Hildesheim und der Herzog August Bibliothek Wolfenbüttel im Vergleich", Wolfenbütteler Beiträge, 15, 129-177 (2009).

Fischer, E., "Safeguarding Book Heritage, Research in Book and Paper Conservation in Europe – a State of the Art", P. Engel, ed., 255-292 (2009).

Pickwoad, N., "The condition survey of the manuscripts in the monastery of Saint Catherine on Mount Sinai", The Paper Conservator, 2, 33-61 (2004), and http://www.ligatus.org.uk/glossary/

Scheper, K., "Refining the Classification of Islamic Manuscript Structures", New Approaches to Book and Paper Conservation-Restoration, P. Engel et al, ed, 357-388 (2011).