

**SUSANNE STARNES**

**Digitization methods and their use for bookhistorical research, especially for analytical bibliography**

The thesis poses the question in which ways traditional book-historical research can be supported by digitisation and computational analysis. Not merely possibilities of support are explored. The crucial matter is whether completely new methods of research are evolving due to new potentials. How can modern methods of electronic storage, digital reproduction and computational analysis be employed for the investigation of medieval manuscripts or early printed material?

The theoretical part focuses on one specific method of book-historical research – analytical bibliography. This method is discussed in detail: It is shown how it originated in the descriptive bibliography of incunabula on the one hand and in the philological study of literary texts of the Elizabethan era on the other hand. The main principle of analytical bibliography is, that printed material supplies the body of primary evidence when trying to gather information about the production circumstances and process. By evaluating the printed pages carefully, by taking into account every– intended and accidental – trace the printing press left, details can be inferred, that are not revealed by any secondary documents. Studying this evidence can help solve technical and chronological questions about the printing process and can also further textual studies. For this latter use the term “textual bibliography” has been established, while pure “analytical bibliography” is not concerned with textual problems.

After laying out more clearly the distinction between textual and analytical bibliography, the thesis explains several specific terms: edition, impression, issue, state, cancellans, cancellandum, technique, procedure, practice, incidence – relying primarily on definitions by Lotte Hellinga and Martin Boghardt. The theoretical part concludes with an extensive discussion of procedures (based on the work of Martin Boghardt and Christoph Reske) of analytical bibliography. Examples, taken mainly from the study of Gutenberg-prints, illustrate this part. Finally, technical aids for this historical research method are presented. The Hinman Collator and its successors, as well as two different versions of composite imaging, one of them already making use of digital photography and computational manipulation, are described, focusing on their key advantages and disadvantages.

The second part of the thesis initially deals with different digitisation technologies, concentrating on aspects which are of importance specifically for medieval manuscripts and early prints, the central problem being how to enable perfect representation of details and - at the same time – guarantee the preservation of the material. Based on technical facts and on the experience of digitisation projects, criteria is developed for choosing the modes of scanning and processing and benchmarking rules are established. Following these preliminary technical considerations, three important research projects using digitisation in different ways are presented and evaluated.

The first, CEEC (Codices Electronici Ecclesia Coloniensis) is not concerned with analytical bibliography, but with the documentation of and research on one large manuscript collection. A system of digital cataloguing is presented, offering the possibility for researchers from all over the world not only to use the supplied information, but also to contribute their own results. Digital representations of – so far – 65.000 pages of manuscripts of the Cologne collections are exhibited in different resolutions, up to 4.491 x 3.480 pixel. Furthermore, paleographical tools were developed, translating traditional methods of analysing manuscript hands into digital methods.

The second project, which is discussed, is the Japanese initiative HUMI (Humanities Media Interface), whose broad aim it is to build up a digital research environment. One of the ambitions of HUMI is to digitise all remaining Gutenberg-Bibles (B 42) and to find out more about the typeset and other production circumstances with the help of computational analyses. So far, six copies of the B 42 have been digitized and partially made available via internet. Research has been done using methods of static and dynamic superimposition and line-by-line comparison, broadly speaking digital versions of collating machines.

The remaining project is that of Paul Needham and Blaise Agüera y Arcas, two Princeton researchers whose work is based on the digitisation and computational analysis of several prints, one of them the Calixtus Bulla. Their research is probably the most sophisticated in terms of digitisation and computer-supported, mathematical examination. Using their results, the two researchers have raised questions about some basic assumptions of the Gutenberg historiography and thereby caused a tremendous stir in the research community. So far no extensive publication has been brought forth by Needham and Arcas, making room for scepticism and criticism. The evaluation offered in the thesis is based on an unpublished article by one of the researchers and partially explains their method, but centers on showing perspectives and problems their – possibly groundbreaking – research poses.

The summary compares these three projects, differentiating between several digitisation aims: to make material widely available, to offer new grounds for international discussion, to transfer traditional modes of research to the new medium or to offer completely new methods, unthinkable by traditional means. It is demonstrated that there is a wide range of possibilities of using digitisation and/or computational analysis to – at the least – stimulate and facilitate research on historical manuscripts and prints, but that one must wait for further advances to judge, whether a wholly new form of research, which makes digital technology compatible with the study of historical material without decontextualizing it, will be developed.