A digital future for manuscript research

On the mutual influence between a digital information environment and manuscript research

Maaike de Koning
s1302442
F.E.W. Praal MA
mr. Dr. F. Kwakkel
23 May 2014
# Table of contents

Introduction ................................................................................................................................. 2

Chapter 1: Manuscript research ................................................................................................. 5
  1.1 General description .................................................................................................................. 6
  1.2 Common research questions ................................................................................................. 9
  1.3 Common research practices ............................................................................................... 13
    1.3.1 Researching the textual content .................................................................................... 13
    1.3.2 Researching the material aspects .................................................................................. 17
  1.4 Scholarly communication by means of grey literature online ........................................... 18
  1.5 Conclusion .......................................................................................................................... 21

Chapter 2: A digital information environment for manuscript research .................................... 22
  2.1 General characteristics, tools and services .......................................................................... 23
  2.2 Useful technologies considering manuscript research practice ........................................ 27
    2.2.1 Downloading and software .......................................................................................... 29
    2.2.2 Adding information .................................................................................................... 32
    2.2.3 Scholarly communication ........................................................................................... 41
  2.3 Conclusion .......................................................................................................................... 43

Chapter 3: Future prospects for manuscript research .................................................................. 45
  3.1 The influence of a digital information environment on manuscript research practices ... 45
    3.1.1 The influence of general tools and services ................................................................. 46
    3.1.2 The influence of downloading and using software ....................................................... 48
    3.1.3 The influence of adding information ........................................................................... 50
    3.1.4 The influence on scholarly communication ................................................................. 51
  3.2 The likelihood of uptake of a digital information environment ........................................... 53
    3.2.1 General factors in the likelihood of uptake ................................................................. 54
    3.2.2 The uptake of downloading and software ................................................................. 55
    3.2.3 The uptake of adding information ............................................................................... 57
    3.2.4 The uptake of tools for scholarly communication ....................................................... 59
  3.3 Conclusion .......................................................................................................................... 61

Conclusion .................................................................................................................................. 63

Bibliography ............................................................................................................................... 66

Appendix I: Illustrations .............................................................................................................. 72
Introduction

Research libraries have seen great changes in their workflows over the past few decades due to the growing importance of digitisation. A recent survey of the Association of Research Libraries showed that 72% of the member institutions surveyed has an active library-wide digitization program and 47% has participated in large-scale digitization projects. In their digitisation efforts these libraries digitise the materials that serve as primary sources for research in the humanities, for example archival materials and manuscripts.

This thesis will focus specifically on digitised medieval manuscripts as one type of material that humanities scholars can use as primary sources in their research. If scholars have digital access to these digital sources it will have certain benefits. They can, for example, access materials that are held abroad, or zoom in on the digital images to a greater extent than the human eye can. The platform created specifically to access these materials, from now on called a ‘digital information environment’, can provide even more benefits if it is equipped with the right kind of tools and services. These tools and services can, for example, allow scholars to easily search for and share digitised manuscripts. In order to become a useful new research tool, a digital information environment should be designed with the needs of manuscript scholars in mind. If a digital information environment does not meet the needs of manuscript scholars, it will likely not be used. Several institutions that digitise manuscripts, like the National Library of France (Bibliothèque nationale de France) and the National Library of the Netherlands (Koninklijke Bibliotheek) have also become aware of this notion. Their end-user studies have already uncovered some scholarly wishes regarding the

---


2 The term "medieval manuscripts" will be defined more clearly in the first chapter of this thesis.

3 The term "digital information environment" will be defined more clearly in the second chapter of this thesis.

4 Note for example the focus on user needs in the JISC Digital Media guide: ‘Selection procedures for digitisation’, *JISC Digital Media* <http://www.jiscdigitalmedia.ac.uk/guide/selection-procedures-for-digitisation/#sp2> (30 December 2013).
digitisation of manuscripts and presentation of the digitised manuscripts online.\textsuperscript{5} Besides end-user research, which usually focuses on a limited group of scholars, there is another method to uncover what scholars likely need in a digital information environment: investigating their research practices. Since scholars will use digital information environments for research, their needs are linked to their research practices. This thesis will therefore first discuss current manuscript research practices and then consider how a digital information environment can align with these practices.

Like many new technologies, the adoption of digital information environments will probably change scholars' research practices in turn. This will give rise to a upward spiral: the design of a digital information environment is influenced by the current research practices, but it can in turn influence these research practices and so on. The aim of this thesis is to uncover and discuss these two mutual influences by answering the following questions:

1) What characteristics, tools and services should an ideal digital information environment have considering the current research practices of manuscript scholars?
2) How can using a digital information environment influence these current research practices?

A third overarching question will be answered based on the discussion of these mutual influences:

3) Will a digital information environment change manuscript research practices, or should manuscript research practices shape a digital information environment?

The answer to this question aims to contribute to the current discussion about technological and social determinism, two concepts that will be discussed in the conclusion to this thesis.

Before answering these questions it is important to get a good grip on the academic disciplines that practice manuscript research.\textsuperscript{6} Most previous research concerning digital research tools focused on either the humanities in general or on one specific discipline within


The end-user research by the \textit{Koninklijke Bibliotheek} of the Netherlands was extensively discussed in an interview on 21 March 2014 with Rosemarie Pomp-Blangé, specialist in usability research. These documents are for internal use by the \textit{Koninklijke Bibliotheek} only, therefore all the information in this thesis about this end-user research is based on the aforementioned interview.

\textsuperscript{6} The term “academic discipline of manuscript research” will be used in this thesis as a container term for all research as defined in this chapter.
Both of these focus points present difficulties. On the one hand, discussing the humanities in general disregards the many differences between the academic disciplines it encompasses. Although the term 'the humanities' is often used, it is certainly not a homogeneous field of research. On the other hand, discussing only one discipline would disregard the fact that scholars from different disciplines in practice use the same primary sources, which form the basis of any digital information environment. In this thesis, it was therefore decided to focus on some of the most important humanities disciplines that practice manuscript research. This approach respects the differences between academic disciplines within the humanities, yet aims to account for most scholars who use manuscripts in their research.

The first chapter of this thesis will be devoted to a discussion of the various types of manuscript research. It will investigate some distinctive characteristics and discuss common research practices. The second chapter will consider the implementation of these characteristics and practices on a digital information environment. It will discuss several tools and services that can support manuscript research within such a platform and thus answer question one as stated above. The third chapter will consider how a digital information environment can influence the aforementioned research practices in turn. This chapter will provide and answer to the second question above. Finally, the conclusion to this thesis will answer the overarching third question.

---

Chapter 1: Manuscript research

The main and clearest characteristic of manuscript research is that its primary sources are manuscripts. Another clear characteristic is that all types of manuscript research will be considered to be within the humanities. However these two are the only characteristics that are distinctly clear. Scholars who study manuscripts come from different academic departments, have different backgrounds and formulate different research aims. Since manuscript research is apparently not a clear-cut field of research, it is important to first describe and define the academic disciplines of manuscript research that will be regarded in this thesis.

Four different aspects of manuscript research will be discussed in order to present a well-rounded description: 1) a general description of manuscript research with attention to some salient characteristics; 2) a discussion of the different academic departments in which manuscripts are used as a primary source for research; 3) a description of the research practices in these different departments; and 4) a discussion of scholarly communication by means of grey literature online in manuscript research. The decision to describe these features of manuscript research is partially based on research by Harley et. al. on the future possibilities of scholarly communication of seven different academic disciplines. In this report, Harley et. al. describe each discipline by its general characteristics, common research practices and common modes of scholarly communication. These descriptions set a nice standard to usefully and comprehensively discuss a scholarly discipline. The same basic design is therefore also used here to describe the research disciplines of manuscript research. However, since manuscript research is done by scholars from different academic departments within the universities and research institutions, this is inserted as the second characteristic to be discussed. A final important note: this thesis considers research practices and communication practices as two distinct things. The research practices will be the central focus point of this thesis, since these will be a key influence in the design of a digital information environment. The discussion of scholarly communication practices will be limited to grey literature online, since the other communication practices are not yet relevant.

---


9 Most universities do not have a separate department for manuscript studies. The scholars who study manuscripts are therefore necessarily part of a different academic department although they are all part of the academic discipline of manuscript research.
to the design of a digital information environment.

1.1 General description

As already mentioned, the most obvious characteristic of the academic discipline of manuscript research is its focus on one particular type of primary source. This thesis focuses solely on scholars who study manuscripts from the pre-print age since both the content and the materiality of these manuscripts can be researched. This makes it a diverse and interesting academic discipline. With modern manuscripts the focus is commonly on the textual content since the material aspects of these manuscripts are usually not intentionally chosen and thus not as relevant to scholarly research. Moreover, there are already several collections of digitised manuscripts from the pre-print age available online which provides actual examples of what might and might not work in a digital information environment. This also makes it easier to imagine the possibilities that can be offered by a digital information environment. In this thesis, a manuscript is thus considered to be a handwritten text from the pre-print age. All handwritten texts from the pre-print age can be primary sources for these scholars. This definition results in a diverse collection of materials, in respect to both the material dimensions, length of the text and nature of the contents. The range of materials varies from large and extensively decorated prayer books, but also single sheet letters and sober scholarly works. Handwritten materials from the modern age are excluded from this definition of manuscripts since there are some clear differences between these two types of written materials. For one, the materiality of older manuscripts can be an important research object, whereas it usually is not for modern handwritten materials. Secondly, before the emergence print, the kind an number of genres to be found in handwritten materials is much more extensive than after. For example, prayer books, legal documents and literary works were quickly being printed, while writing was reserved for more private genres like letters and diaries. When a scholar makes use of a pre-print manuscript, this research effort will be considered a part of manuscript research in this thesis.

Now that the manuscript is defined, some general characteristics of manuscript research will be discussed. Most of these characteristics are common to all humanities research and they set humanities apart from other academic disciplines. These general characteristics are important to consider since manuscript research, as a part of humanities scholarship, is shaped by these research characteristics.

Humanities research is traditionally qualitative with research output that is mainly
narrative and argumentative. These characteristics are also common to manuscript research. Qualitative research has always prevailed in the discipline of manuscript research, probably because it was difficult to find quantifiable evidence to support a manuscript scholar’s arguments. However, due to current developments in information technology it is increasingly feasible to extract quantifiable data from a manuscript or a collection of manuscripts. For example, a tool for automatic word counting can be used to analyse a digital full-text file of the content of a manuscript. This practice can be useful for linguists who are interested in the use of certain verbs over time. There seems to be a growing interest in such quantitative research within manuscript studies. It is however important to note that the quantitative data is commonly not viewed as well-rounded research output. The relevance of this data has to be explained in a narrative or it can be used to support an argument. Therefore, the analysis of a manuscript might increasingly be done with quantitative methods, but the research output will probably remain narrative and argumentative.

Referencing is another important characteristic of humanities research. Citations and references are common to all scholarly output, but two things set referencing in humanities apart from referencing in other scholarly disciplines. For one, humanities scholars use more references than scholars from other disciplines. This indicates that finding, reading and citing relevant research has a bigger role in the research practice of humanities scholars. Secondly, humanities scholars often provide an in-depth reaction to previous research and refute or confirm previous arguments. They have thus created interlinked threads of scholarly discourse that span over several articles. To find relevant secondary sources, humanities scholars are taught to make use of these links by a practice called footnote chaining. A scholar then starts with a trustworthy source and reads the relevant sources that it refers to and so on. By doing so he can get a good grip on what has been said about a specific topic or manuscript

---

11 See for example the abstracts of lectures during the conference ‘Easy tools for difficult texts’ that took place in april 2013: Huygens ING <http://easytools.huygens.knaw.nl/?page_id=39> (30 December 2013).
and how his own research can fit into this discourse. This practice can also be a useful for scientists, but it has a more prominent role in humanities scholar’s research practice. Since manuscript scholars are all humanities scholars, it is very likely that finding, reading and citing relevant research also has a prominent role in their research practice. It is also very likely that manuscript scholars use footnote chaining to find relevant secondary sources to refer to.

Flexibility is another common characteristic of humanities research. In order to illustrate this, it is useful to compare humanities research with scientific research. Scientific research has to adhere to certain conditions in order to be viable. For example, it commonly has to be based on a suitable sample-size, quantifiable methods should be used and the results should be generalizable. Humanities scholars are commonly much less limited in the way they can set up their research. They can use or combine quantifiable and qualitative research methods and study anything from a collection of primary sources to only a single primary source. This is also the case for manuscript scholars as clearly stated by J.H. Hexter:

Nowadays many historians casually and habitually pick up bits of the conceptual apparatus of this, that, or the other social science, try them for size on the data at hand, get what use they can out of them, and just as casually abandon them when they cease to serve their purposes. Manuscript scholars are also free to decide how many manuscripts they have to study to answer their research question. Even researching one manuscript is a common research design.

A final general characteristic specifically of manuscript research is the current shift towards research in teams. Traditionally manuscript research is more individualistic than scientific research. The explanations for this individualistic nature vary from a strong link between the quality of the research and the researcher, to the importance of individual

---

15 W.S. Brockman, et. al., Scholarly work in the humanities and the evolving information environment, p. 11.
18 Consider for example the research presented in the journal Manuscripta: Metapress <http://www.metapress.com/content/p37611120092/> (3 January 2014).
19 Consider for example the research presented in the journal Manuscripta: Metapress <http://www.metapress.com/content/p37611120092/> (3 January 2014).
characteristics in the production of papers.Regardless of the reason for individuality, there seems to be a growing interest in manuscript research by teams of scholars. This shift is often ascribed to the technological advancements in communication. Consider for example the VIDI project Turning over a New Leaf: Manuscript Innovation in the Twelfth-Century Renaissance or the project Dynamics of the Medieval Manuscript: Text Collections from a European Perspective. Both of these research projects are done by a team of manuscript scholars from several countries and both study a large number of manuscripts. The participating scholars will however not co-author one single research report, as is common in science. Instead, each scholar will write his own research output based on the jointly gathered data. Unfortunately there is no data available on the amount of manuscript research that is done in teams and the amount of research projects that results in co-authored output. However, there surely seems to be a growing interest in the possibilities of research in teams.

1.2 Common research questions

What kind of research question a scholar asks depends on his academic discipline. It is thus important to first consider the academic disciplines where manuscript scholars can be found, which are: philology, textual criticism, book history, history and art history. Scholars from other academic disciplines can also make use of manuscripts as a primary source for research, but within these five academic departments scholars are most likely to do so. This thesis will now present a brief description of these five research disciplines and the kind of research questions that manuscript scholars within each of these disciplines are likely to ask. The research practices will be discussed more in-depth later on in this chapter.

Philologists who work with manuscript can be found in the academic departments of medieval literature studies and historical or medieval linguistics. Using a manuscript is inevitable for these scholars, since the materials they are interested in can only be found in manuscripts. These scholars often use philological research practices since the writing,
language use and cultural and social circumstances in the Middle Ages are very different from what we are currently used to. This presents obstacles when scholars want to read and interpret the textual content of a medieval manuscript. Philologist’s academic goal is to take away these interpretational barriers by providing all the contextual information the original readers had. Philologist thereby assume that all texts need to be read and interpreted with the appropriate cultural and historical circumstances in mind. Besides this contextual information, a philological study focuses on three other elements:

1) The writing, the symbol of the thing signifying.
2) Language, the thing signifying.
3) The thing signified, the knowledge contained in language.

By approaching a manuscript from a philological frame of mind, whether explicitly or implicitly, scholars can get a better understanding of their primary source. Even scholars who focus solely on specific elements of the text, like certain words or sentence structures, can benefit from such a broad approach as a start of their research. Their interpretative efforts can in turn also contribute to the philological study of the text as a whole.

The types of questions that these scholars ask all concern the textual content of a manuscript. Literature scholars most likely focus on interpreting the textual content of a manuscript. Common research questions concern story lines, the development of characters, motifs or uses of symbolism. Linguists are more likely to break the text down into words and word clusters and make these their research object. They can research, for example, the choice of words, verb conjugation and spelling in a certain manuscript or compare these to similar elements in other manuscripts.

Literature scholars and linguists can also use research practices of textual criticism to study the content of a manuscript more accurately. The final aim of textual criticism is to

---

24 H. Parker, ‘What is it that philologists do exactly?’, p. 160.
26 Consider for example the list of contents of Queeste, a Dutch journal on medieval literature: DBNL <http://www.dbnl.org/tekst/que002200701_01/> (3 January 2014).
reconstruct the text that most nearly represents the author’s original intention, or the ‘authoritative text’. The theory behind this method is that the textual content of a manuscript is always a compound of former versions of that text. When scribes received an exemplar of a text and copied it into a new manuscript they often changed it. Textual criticism provides a method to trace back these changes and either discover which manuscript holds the authoritative text or recreate this text if it no longer exists. So instead of taking a broad perspective like philology, textual criticism zooms in on the textual content of a manuscript and aims to analyse it on a word-by-word basis. To linguists who are researching the sentence structure used in a certain period it is important to research the authoritative text that is not clouded by changes made on a later date. Literature scholars can benefit from the overview of changes that were made throughout the years to describe how certain motifs changed.

Book historians also make use of manuscripts as a primary source for research. Book history is a relatively recent academic discipline that focuses on print culture and the role of books as material objects within that culture. While at first connected to the academic disciplines of bibliography and social history, its scope of research has broadened to include each phase in the life cycle of books and the relation of this process and its product to economic, social, political and cultural developments. Although a lot of research is dedicated to the current developments in new media, there are also many book historians who focus on the period before movable type. These scholars study for example the manufacturing stages of manuscripts, the way they were used and the different types of manuscripts that were produced. In order to answer these questions, a book historian mainly studies the material aspects of the manuscripts. These material aspects contain clues about, for example, the production methods and intended use of a manuscript. Book historians can use quantitative methods which have been related to this discipline since the beginning. However, qualitative studies of only one or just a few manuscripts and their production or cultural influence are also a common research design.

Medievalists can use manuscripts as a primary source for their research too. Medievalists are historians who focus on the history and culture of the Middle Ages. Their research concerns topics like medieval arts and culture, but also medieval politics and society. The textual and material remains of those times, most importantly manuscripts, are the only evidence these scholars can use in their research.\(^{33}\) However, accuracy is an important concern when using a manuscript as scholarly evidence of historic events. To assess this, both the authenticity of the author and manuscript, and the usefulness of its content are key pieces of information.\(^{34}\) If the source is judged to be accurate in all three aforementioned respects, the information within the manuscript can be used to answer a research question or support a hypothesis. The two most commonly asked research questions in history are ‘How did it come about that…?’ and ‘How did he (or they) happen to…?’\(^{35}\)

Art historians form the last group of scholars that is likely to use manuscripts as a primary source in their research, especially those who focus their research on the Middle Ages. Art historians who study manuscripts are mainly interested in illuminated manuscripts since these contain elaborate decorations and images. Their research efforts can focus on, for example, symbolism used in decorated capitals or how the illuminated images were created.\(^{36}\) Just like book historians, art historians will thus focus mainly on the material aspects of a manuscript, i.e. the images and decorations.

A clear pattern emerges from this review of the types of questions scholars can ask. Even though these scholars come from five different academic disciplines, they all focus either on the textual content of the manuscript, or on its material aspects. Evidently the scholars who research the content of a manuscript, i.e. philologists and medievalists, are looking for different elements within the content. Literature scholars and historians focus mainly on the narrative of the text, while linguists are more likely interested in the text on a word-by-word level. Book historians and art historians on the other hand, are more interested in the material aspects of a manuscript. Book historians can focus on all material aspects of the book, while art historians are probably only interested in the images and decorations. Of course this is a very broad categorisation, but it is a clear way to get a better understanding of

\(^{34}\) A. Tucker, *A companion to the philosophy of history and historiography*, p. 16.
\(^{36}\) Consider for example: J. Backhouse, *The illuminated page: Ten centuries of manuscript painting* (London: British library, 1997).
a research discipline that is as diverse as manuscript studies. However, this division does not
mean that scholars who research the material aspects completely disregard the textual content
of a manuscript and vice versa. For example, book historians need to know what type of text
they are handling in order to interpret its material aspects.

1.3 Common research practices
With these two types of research questions in mind, one can now consider the research
practices that are commonly used to answer these questions. This paragraph will first discuss
the research practices commonly used to answer questions concerning the textual content of a
manuscript and then those common to researching the material aspects of a manuscript.

1.3.1 Researching the textual content
There are two common research practices a manuscript scholar can use to research the textual
content of a manuscript: philology and textual criticism. Both of these research disciplines
have already been discussed briefly, but the following pages will discuss them more in depth.
However, there is an important distinction to be made first, between the textual content of a
manuscript and the text. In this thesis, the textual content (also referred to as ‘content’) of a
manuscript is the unique copy of a text as it is written down in a manuscript. The text is a
scholarly edition or corrected version of several manuscripts combined. Researching the
textual content of a manuscript and researching the text are therefore two different types of
research, both of which will be discussed here.

There is a lot of discussion about what the research discipline of philology entails and
how important it is when scholars study the content of a manuscript. Some scholars view
philology as a necessary start when studying the textual content or text within a manuscript,
while others see it as the product of several other disciplines and therefore not a scholarly
discipline on its own right. To complicate the matter, there is some overlap with the
disciplines of linguistics and literature studies. Although the discussion about this overlap is
on-going, this thesis will regard linguistics and literature studies as sub-disciplines of
philology since both research practices can contribute substantially to a philological study.
A more in-depth discussion of the relation between these disciplines goes beyond the scope of
this thesis. It is clearly difficult to find a dictionary definition of philology that is supported by
a majority of the scholarly community. However, for clarity purposes this thesis will present a

37 J. Ziolkowski, ‘”What is philology?” Introduction’, Comparative literature studies, 27
38 J. Ziolkowski, ‘”What is philology?” Introduction’, p. 6.
tentative description of philology, based on different considerations in scholarly literature.

De Saussure considered the final aim of philology to be ‘correcting, interpreting and commenting upon texts’.\(^{39}\) According to this definition, philology is basically the same as textual criticism or linguistics. However, instead of focusing on certain aspects of a manuscript and interpreting these, a philologist regards every aspect of the manuscript and of the time and place when it was produced as meaningful to the interpretation of its textual content. This makes philology a reconstructive method instead of a deconstructive theory.\(^ {40}\)

As already discussed, philologists assume that a modern reader will be better equipped to read and interpret the textual content of a manuscript when provided with contextual information and information about the handwriting and language use.\(^ {41}\) The information about the handwriting and language use is essential because writing styles and language use in the pre-print ages are different from our current typefaces and language use. At this point, philology clearly overlaps with palaeography, the research discipline that focuses on reading and studying script. Once a scholar can read the manuscript and view it with the information on the cultural and historical background in mind, he can interpret, correct and comment upon the text. It is often suggested that other disciplines of manuscript research such as palaeography, codicology and textual criticism are constituents of philology.\(^ {42}\) Although this discussion goes beyond the scope of this thesis, it is important to keep in mind that all these research methods can at least contribute to a philological study.

The final product of a philological study is indeed as De Saussure described it: a corrected text with comments and interpretations. This final product can help a scholar put the manuscript into a wider cultural and historical context and know its general structure and content. This is important to all manuscript scholars, regardless of their research discipline. However, a philological study can also be so extensive that it is a research effort in its own right instead of just the starting point of a research project. The rationale behind philology can also be applied to researching a text. In that case, a philological approach can help to compare several manuscripts while considering their different cultural backgrounds.

Textual criticism is another discipline of manuscript research that is commonly used to research the content of a manuscript. As already discussed, textual criticism aims to

\(^{39}\) J. Ziolkowski, ‘‘What is philology?’’ Introduction’, p. 6.
\(^{40}\) J. Ziolkowski, ‘‘What is philology?’’ Introduction’, p. 11.
\(^{41}\) H. Parker, ‘What is it that philologists do exactly?’, p. 160.
\(^{42}\) J. Ziolkowski, ‘‘What is philology?’’ Introduction’, p. 6.

The disciplines of manuscript studies that are mentioned here will be discussed later on in this chapter.
reconstruct the authoritative text, which most nearly represents the author’s original intention. In practice, scholars do this by creating a stemma codicum, which compares all manuscripts containing the text on a word-by-word level. This comparison will reveal several differences between the text in each manuscript. Based on these differences, a scholar can conclude which manuscripts are closely related and thus construct a genealogy of the text. This genealogy will trace back to the authoritative text, which is not necessarily an existing manuscript, but can also be an assembly of parts from several currently existing manuscripts.

The process of creating a stemma codicum has to be done meticulously and it usually takes a long time. When the stemma is finished it provides a scholar with an authoritative text and also a lot of information about the text. For example, the genealogy of the text and an overview of the changes that were made over the years. Manuscript scholars from each discipline can use this information, but each scholar will probably focus on a different aspect of the results. For example, linguists are likely interested in the changes in word use that scribes have made over the years, while medievalists can use a stemma to compare witnesses of a historic event and to eliminate those who only quote others.

Medievalists also have their own research practices to study the content of a manuscript. As with all research practices, it starts with selecting a hypothesis. It is however important to note that medievalists commonly formulate their research hypothesis based on the availability of evidence to prove or refute it. This is a common practice because only a limited amount of manuscripts are handed down to us, which makes certain hypotheses unverifiable and therefore not worth investigating. Although this practice is only explicated for historians, it seems logical that it applies to all scholars who use manuscripts as a primary source in their research.

Once the research hypothesis is formulated, a medievalist will use the evidence found in manuscripts to test this hypothesis. The following description on how to use evidence from manuscripts in a methodologically sound manner is wholly based on R.J. Shafer’s *A guide to historical method*. He describes how most scholars start with external criticism, or

---

authenticating the evidence.\textsuperscript{48} This practice includes dating the manuscript, obtaining information about the author and establishing the most accurate form of the relevant passages. If the manuscript is not dated, medievalists commonly look for dates mentioned in the content. Uncovering who wrote the manuscript can also help to roughly date a manuscript. Concerning the author, it is important to find out whether the text was written by one author or whether it contains forged passages. This can be done by tracing anachronisms in spelling, word use, language structure, script or references to other events. Establishing the most accurate form of the text is commonly done by comparing different versions of one text and trying to find the most reliable account. This is often the earliest account, before other scribes have added or omitted parts of the text. The research method of textual criticism can also be applied here although this will probably take more time.

After the external criticism, the internal criticism commonly follows to determine the credibility of the evidence.\textsuperscript{49} To assess the credibility of an account, a medievalist has to dive into the content and understand what the author tried to convey. Some of the methods discussed by Shafer overlap with the philological research method. For example, the first step is to understand the meaning of the words that are used, since the meaning of certain words has changed over time. Once a scholar becomes familiar with of the meaning of the words he can move on to interpreting them and assessing the veracity of the statements. In doing so, it is important to regard the author of the textual content and wonder, for example, what his intentions were and how well he could observe the thing he reports on.

The third and final aspect of the historical method concerns the synthetic operations, i.e. manipulating the evidence in order to create a comprehensive body of evidence that supports or refutes ones hypothesis.\textsuperscript{50} In practice, a medievalist will analyse the evidence he has found, compare this to other pieces of evidence, combine the pieces of evidence into groups and combine these groups into a final interpretation. The connections between the pieces of evidence will consist mainly of causal links, but inference can also be used. It is common practice to present one's final interpretation in the form of a historical narrative, since this helps to create coherence in historical facts.\textsuperscript{51} By continuous causal linking in the form of a story, a medievalist tends to make sense of what happened in the past and why it happened in such way.

\textsuperscript{48} R.J. Shafer, \textit{A guide to historical method} (Homewood, etc.: Dorsey Press, 1974), pp. 117-139.
\textsuperscript{49} R.J. Shafer, \textit{A guide to historical method}, pp. 141-162.
\textsuperscript{50} R.J. Shafer, \textit{A guide to historical method}, pp. 163-190.
\textsuperscript{51} J.H. Hexter, \textit{Doing history}, p. 27.
1.3.2 Researching the material aspects

The material aspects of a manuscript are also a viable research object, especially for book historians and art historians. There are two research disciplines devoted to researching the material aspects of a manuscript, namely codicology and palaeography.

Codicology is the discipline within manuscript studies that takes the most comprehensive approach to studying the material aspects of a manuscript. All material aspects can be studied, for example, the material on which the text is written, the collation of the manuscript, its pricking, ruling and binding. A scholar can study these and other aspects of one manuscript or compare them in several manuscripts to uncover patterns. Often the material aspects of a manuscript can provide a scholar with arguments to draw more comprehensive conclusions about the manuscript. For example, the amount of decorations and the type of script can indicate what the intended audience or the intended use of the manuscript was. These are the kind of conclusions that especially book historians are interested in.

There are several handbooks available that provide information on the aspects of a manuscript that can be studied and what these aspects signify. These handbooks also present the terminology commonly used in codicology so scholars can understand what their peers have observed. But not everything can be learned from these books. Codicology is also a ‘skill passed on through generations’ and scholars often acquire this skill through experience.

The research output of a codicological study is usually a manuscript description, which contains information on for example, the size of the leaves, the size of the writing area on an average page, the type of script, the decorations and the binding. A scholar has to study the manuscript meticulously to create a comprehensive description. However, not every aspect of the manuscript has to be analysed and described to create a methodologically sound manuscript description. A scholar can also choose to focus on just one or two material aspects and study only those. A manuscript description is a useful tool book historians to get an overview of a manuscript and draw further conclusions. It can also be useful for art historians.

---

54 R. Clemens and T. Graham, *Introduction to manuscript studies*.
55 R. Clemens and T. Graham, *Introduction to manuscript studies*, p. xiii.
since a manuscript description usually contains information about the illustrations and decorations within the manuscript.

Palaeography is another discipline within manuscript studies that focuses on the material aspects of a manuscript. In short, palaeography is the study of scripts, but it also includes, for example, dating a manuscript based on the handwriting, classifying the scripts and reading the text.\(^5^6\) The handwriting in a manuscript contains a lot of information about the manuscript. For example, how many scribes have worked on a manuscript and approximately when and where it was written.\(^5^7\) This information is useful since it can help to define the origin of a manuscript and thus provide information about its cultural background. Therefore, palaeography is a valuable research practice for all manuscript scholars.

As with codicology, there are several handbooks that can guide a scholar in a palaeographic study.\(^5^8\) These handbooks usually present several examples of scripts and provide information about these scripts, for example, when and where they were commonly used. Experience is again key in learning to distinguish the subtle differences between scripts from different scribes, ages and scriptoria. Besides being able to distinguish between different scripts, palaeographers also have to be able to read the textual content of a manuscript in order to study the abbreviation styles and punctuation.

Palaeography is not practiced by all scholars who study manuscripts, since it is quite complex and requires experience. Fortunately, there is a common terminology that is used to record one’s research results. Scholars who are not able to do a complete palaeographic analysis themselves can thus understand what more experience palaeographers discovered. The terminology used in palaeography concerns both the names of different types of script and means to describe different types of abbreviation and punctuation styles.

It is important to note that all manuscript scholars can benefit from the aforementioned research practices in some way. This connects these research practices and makes it easier to view them as components of the overarching discipline of manuscript research. As such, these separate research practices characterise the discipline of manuscript research as a whole.

1.4 Scholarly communication by means of grey literature online

So far, this chapter has covered some general characteristics of manuscript research, the research questions that are commonly asked and the research practices scholars can employ to

\(^{5^6}\) S.A. Harvey and D.G. Hunter, *The Oxford handbook*, p. 140.

\(^{5^7}\) S.A. Harvey and D.G. Hunter, *The Oxford handbook*, p. 160.

\(^{5^8}\) For example the handbooks mentioned in footnote 52.
answer these questions. Sharing these answers with the scholarly community is usually the next step in the academic work flow. There are several incentives to share ones findings with the scholarly community, for example, validation of one’s research, communication with peers, career advancement, claiming a research finding and the stimulation of progress and knowledge in society in general. These incentives are relevant to all scholars, regardless of their scholarly disciplines. To publish one’s research, a manuscript scholar has roughly four possibilities: publishing in the form of a book, publishing an article in a journal, presenting ones research at a conference or presenting ones research in the more informal area of ‘grey literature’.

The use of a digital information environment as discussed in this thesis might bring about changes in scholarly communication too. Such changes will first be seen in grey literature which is currently still developing and probably more susceptible to change. Moreover, the area of grey literature has always been more flexible than formal publishing since it does not depend on publishers nor is it designed to obtain academic rewards. Changes in formal publishing would depend on scholarly publishers changing their work flow and universities changing their academic reward system. Such changes are not so easily brought about. Since the influence of a digital information environment will thus probably first become visible in grey literature, this thesis will limit itself to discussing this manner of publishing.

The term grey literature encompasses all modes of publishing that are not controlled by commercial publishers. Grey literature has become more important and influential since the World Wide Web provided scholars with user-friendly methods to publish work online. One can think of blogs, Twitter and Facebook. Scholars seem to be especially interested in blogs as an alternative way to draw attention to their research. For example, a recent study about scholarly blogging showed that 73% of the analysed blogs by academics were geared towards their professional peers and 40% of these blogs were written in a formal essay style. This indicates that most of these academics view their blog as an alternative way to reach their peers and possibly as a way to build a good reputation. These publications are however

---

60 Greylit <http://www.greylit.org/about> (13 August 2013).
not likely to benefit one’s professional career since they are not recognized as reasons for funding or promotion in academia.

Several scholars in the academic discipline of manuscript research also publish about their research in the form of grey literature online. The visual appeal of the manuscripts that they research, combined with the current possibilities to share images online have probably contributed to this development. Although there is no formal research regarding the frequency and impact of this kind of publishing, there are some interesting examples that are worth discussing here. For example the blog of several scholars participating in the VIDI project *Turning over a New Leaf: Manuscript Innovation in the Twelfth-Century Renaissance*. These scholars write blog entries about their research findings and general interest pieces concerning manuscript research. The scholars who work on the project *Dynamics of the Medieval Manuscript: Text Collections from a European Perspective* do the same thing. Contradictory to what was found in the study mentioned before, these two blogs seem to be catered towards general interested readers, instead of scholarly peers. However, both blogs contain links to information that can be interesting for peers too. Besides these project-based blogs there are several Twitter accounts dedicated to manuscript research. For example @BLMedieval, from the British library and @erik_kwakkel, from a book historian at Leiden university. These Twitter accounts mainly contain appealing images of manuscripts and messages that refer to interesting articles elsewhere. With respectively 9.428 and 5.661 followers, both pages seem to attract quite some interest. There are also two fora online dedicated to manuscript research, namely the *Digital medievalist* and the *Digital classicist wiki*. The *Digital medievalist* forum is dedicated to digital manuscript research. It features a section with articles on this topic and a section for discussion and questions. The *Digital classicist wiki* is not dedicated to historical research. Since this kind of research is often based on manuscripts, this platform also contains relevant information for manuscript scholars. This forum does not have a section with articles, but focuses solely on scholarly discussions on topics related to the academic discipline of history. Since there is no in depth research on this topic yet this thesis will limit itself to these examples. More in-depth research is required in

64 *The dynamics of the medieval manuscript* <http://everycodextellsastory.eu/> (3 January 2014).
65 *Twitter @BLMedieval* <https://twitter.com/BLMedieval> (3 January 2014).
*Twitter @erik_kwakkel* <https://twitter.com/erik_kwakkel> (3 January 2014).
order to draw any further conclusions.

1.5 Conclusion

Although the discipline of manuscript research encompasses scholars from different academic disciplines, this chapter has shown that it can be described by some general characteristics. The research output generated by the discipline of manuscript research is mainly argumentative and narrative, adding to existing threads of scholarly discourse. There seems to be a growing interest in quantitative research, although most of the research currently done is still qualitative. Compared to scientists, manuscript scholars are quite free in the way they want to set up their research. Both quantitative and qualitative research methods can be used to research a full collection of manuscripts, or a small set or even just one manuscript.

Since manuscript scholars come from different academic disciplines they tend to ask different research questions. In general, they can either ask questions about the content of the manuscript, or its material aspects. To answer these questions manuscript scholars can employ certain research practices. The most commonly used research practices are philology, textual criticism, history, codicology and palaeography. These research practices are all presented in several handbooks for scholars, but experience is always an important aspect in fruitfully applying them. The handbooks mentioned before usually also introduce the common terminology used to present research findings so they can be understood by other scholars.

Manuscript scholars have the same means to publish and share their research results as all other scholars; namely publishing a book, publishing an article in a journal, presenting ones research at a conference and publishing in the area of grey literature. This chapter only discussed the current state of publishing in the area of grey literature since a digital information environment will most likely influence this kind of publishing the most. Publishing in the area of grey literature is not controlled by commercial publishers and it has grown a lot in both size and relevance since the World Wide Web has provided scholars with easier means to publish online. Currently, the importance and practice of publishing in the area of grey literature for manuscript scholars has not been researched. However there seems to be a growing interest in its possibilities of which several examples were discussed.
Chapter 2: A digital information environment for manuscript research

The first chapter of this thesis discussed some general characteristics of manuscript research and the research practices of philology, textual criticism, history, codicology and palaeography. This second chapter will discuss the tools and services that should be available in a digital information environment, in order to make it align with these practices. Yet before discussing these tools and services, it is important to define what a 'digital information environment' is. First off, it is an online repository for digital images of manuscripts and derivative files, like a full-text file or a downloadable PDF-file of the images. Several aspects of these images and derivative files will be thoroughly discussed in this thesis since they are the core of the platform. Secondly, a digital information environment usually offers several tools to access and research these digitised manuscripts. These tools will be discussed more extensively over the course of this chapter. Thirdly, it enables discussion and collaboration between scholars, for example, by allowing scholars to share bibliographic information, to discuss new research practices and to co-author research output. The fourth and final important characteristic is openness and extensibility. The platform is open to all scholars who want to use it, but not to the general public in order to maintain a certain level of quality and academic validity. Extensibility enables the addition of new tools and services if need be, which allows the platform to develop alongside manuscript research practices.

When one unites all these characteristics with the tools that will be discussed in this thesis, it would be an ideal digital information environment. However, the actual creation of such a platform is currently not feasible because of both technological and financial reasons. This chapter therefore focuses on the theoretical possibilities of aligning a digital information environment with current research practices, thus supporting manuscript research. It will also present some examples of existing tools to illustrate the current possibilities. This chapter will start with an overview of general features of a digital information environment. These features are:

67 The term ‘digitised manuscript’ refers to the digital images of a manuscript. A full-text version of a manuscript is a digital record of the textual content of a manuscript that can be downloaded, searched and edited. These files are produced via either optical character recognition (OCR), transcription or a combination of these techniques.


69 D. de Roure, et. al., ‘my Experiment: Defining the social virtual research environment’, n. pag.
will align with the overarching characteristics of manuscript research as discussed in the previous chapter. It will then discuss several tools and services that will be useful to manuscript scholars, considering the research practices of philology, textual criticism, history, codicology and palaeography. These tools and services specifically concern downloading and software, adding data and scholarly communication.

2.1 General characteristics, tools and services

End-user research by the National Library of France (BnF) and the National Library of the Netherlands (KB) showed that humanities scholars have some general expectations of digitised sources and a digital information environment. Both of these end-user studies were based on focus groups of respectively seven and nine humanities scholars. This number of participants allows for an in-depth discussion where all participants get a say, which is the main goal of a focus group study. The aim of these studies was to find out what scholars expect from a digitised collection of paper heritage and the portal to access these collections. The results of these studies show many similar expectations, most of which are linked to scholarly research practices. These results and links will be discussed to explain why certain general characteristics and tools are necessary, or at least useful, in a digital information environment.

The first recurring theme in the end-user research by the BnF is a demand for high quality images. This demand is related to the fact that many manuscript research efforts are qualitative, with focus on a single manuscript or a small amount of manuscripts that is studied in great detail. It often involves the observation, description and analysis of a specific element within a manuscript. It is crucial that even the smallest details of such an element can be studied. Scholars will be better able to do this if they have access to high quality images of the manuscript. When using low quality images scholars will more likely overlook details because these might be blurred or displayed in black and white. The same goes for the derivative files. Some scholars will prefer to use a full-text version of the textual content of a manuscript instead of reading the textual content of digital images. This practice can be compared to using a scholarly edition instead of the original manuscript. It is in this case vitally important that the full-text version is highly accurate since any inaccuracy can lead to

70 P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
Interview with Rosemarie Pomp-Blangé, 21 March 2014.
71 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
flaws in their research.

Besides qualitative research there is also a growing interest for the possibilities of quantitative research of manuscripts, especially by means of computational methods. There are roughly two possibilities for quantitative manuscript research: a scholar can use quantitative methods to study either a single manuscript or a collection of manuscripts. A digital information environment can easily support current methods of quantitative research of a single manuscript. However, the option to download the images and derivative files would be essential. If a scholar is able to download these files he can continue his traditional research practice, or use computational methods to analyse the downloaded digital image or file. For example, the availability of full-text files enables frequency analysis. Quantitative research of collections can benefit from a tool to download the digitised manuscripts and derivative files in the same way. In addition, it is important that the digitised collections are as complete as possible. If manuscripts are missing from a digitised collection, this will present a bias in a scholar’s research output, or it forces him to look up the analogue manuscript. This explains the scholars’ request for completeness. Moreover, a digital information environment can support quantitative research on collections by digitally uniting manuscripts from collections that have been dispersed around the world. It is an ideal platform to achieve this since it can easily make these manuscripts accessible through one portal. Scholars can then research these collections without the need to travel. The end-user research by the BnF has shown that this research practice is indeed of special interest to current scholars, also for qualitative research projects.

The previous chapter already discussed the importance of finding useful secondary literature on the manuscript or collection that is being studied, most commonly done through footnote chaining. Since this is such an important aspect of manuscript research practice, it will be useful if scholars can also do this in a digital information environment. The practice of footnote chaining can probably be usefully combined with digital social reference management (SRM). This technology will be discussed more in depth under ‘2.2.3 Scholarly communication’.

Both individual research and research in teams should be supported by a digital information environment because manuscript research traditionally is an individual endeavour because manuscript research traditionally is an individual endeavour

---

72 Frequency analysis allows scholars to count how often certain words are used in a text. This is already done digitally, for example on the following website by Jonatha Reeve: Macro-etymological analyzer <http://jonreeve.com/dev/etym/etym.php> (5 May 2014).

73 Interview with Rosemarie Pomp-Blangé, 21 March 2014.

74 P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
but interest for research in teams is growing. Individual researchers can possibly benefit from a personal storage space to save items for further research. A storage space will provide an overview of scholar’s primary sources, which do not have to be retrieved repeatedly. The Early European books online website already offers such a personal storage space.

Unfortunately, user statistics are not available. The two aforementioned end-user studies show contradicting opinions on the merit of such a personal storage space - scholars participating in the research by the BnF showed an interest in a long term personal storage space. However, scholars participating in the research by the KB had no interest in such a feature since they all had their own method for storing useful sources. Research of the actual use of a personal storage space is thus needed to draw a definitive conclusion.

Scholars who work in teams will have to communicate on a regular basis and will probably share their research data and possibly co-author research output. Scholars can therefore benefit from tools that facilitate this kind of collaboration in a digital information environment. There are currently several platforms that facilitate online collaboration, for example Diigo. This online platform enables teams of scholars to share digital files, add comments to each other’s work or to shared secondary sources, tag important topics and engage in discussions. Its 7 million registered users attest its value for research practice.

Although Diigo thus already is a useful platform for collaboration available online, this thesis argues that it is still important to incorporate such a platform in a digital information environment because it will improve their extensibility. When scholars would be using a generic collaboration platform, like Diigo, this could not be completely catered to their needs since there would be users from other disciplines to consider. However, if the collaboration tools are only used by manuscript scholars with many similar needs, it is easier to adapt them to fit their needs. The tools that were mentioned above can thus all be usefully incorporated in a digital information environment.

Since it is common practice for all scholars to take research notes, it would be useful

75 P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
76 Early European books online <http://eeb.chadwyck.co.uk/access.authkb.kb.nl/info/demo.do> (4 February 2014).
77 P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
78 Interview with Rosemari Pomp-Blangé, 21 March 2014.
to also provide a tool to do so in a digital information environment. Digital note taking has certain advantages over taking notes with pen and paper. For one, a scholar can add digital notes directly to the digitised manuscript or derivative files, either in his local storage or online. A scholar can thus create a good overview of both the primary source he is using and the notes he has taken. This practice is already in use on the aforementioned platform Diigo (fig. 1, all images can be found in Appendix I). Secondly, it is easier to search and manage digital notes than notes on paper. Tools to search and sort digital notes are especially useful when a scholar has an extensive collectio. Although some scholars already have a satisfying system for taking notes, others can probably benefit from the ease of organisation that digital note taking offers. Thirdly, digital notes can easily be shared, which is convenient for scholars who work in teams. In most tools for digital note taking it takes only a minute to share these notes, which makes it more efficient than, for example, arranging a meeting or typing them in an e-mail. However, to reap these benefits scholars would have to learn how to work with tools for digital note taking and abandon the pen and paper they have gotten used to. Depending on one's personal preference this can be a reason not to use digital note taking.

A digital information environment also needs some tools and services to help scholars answer their research questions. As discussed in the previous chapter, manuscript scholars can ask questions about the textual content (or text) of a manuscript, or its material aspects. Both of these approaches should be supported by some specific tools and services.

For scholars who want to research the textual content of a manuscript it would be convenient if an accurate full-text version of the manuscript content were available for two reasons. Firstly, it would make it easier for scholars to read the manuscript, especially for scholars who have trouble reading certain handwritings. Secondly, scholars can nowadays use computational methods to research a digital text and answer their questions. The application of frequency analysis was mentioned before, but the possibilities extend beyond that. There is, for example, more sophisticated software available to create a stemma codicum. These new research tools can be used more easily when the full-text files of manuscripts are available.


The topic will be discussed more in depth on pages seven and eight.
online. Using these tools can in turn allow scholars to ask different research questions, but this topic will be discussed more in-depth later on in this chapter. Although using a full-text versions has some clear benefits, it is important to note that scholars will want to consult the images of the manuscript to check if the transcription is correct.\textsuperscript{84} Providing only a full-text version of a manuscript is therefore not sufficient.

If a scholar wants to research the material aspects of a manuscript there is one condition that has to be met: the images of the manuscripts have to be of high quality.\textsuperscript{85} Moreover, tools to enlarge the images are essential according to the scholars who participated in the BnF end-user research.\textsuperscript{86} These scholars also desired a permanent display of the scale of the image, an electronic ruler and tools to change the contrast, luminosity and colour filters.\textsuperscript{87} Unfortunately there are currently no open digital collections that feature these tools to provide an example of how they can work. The Friedberg Genizah project announces that their collection has several of the aforementioned features, but access to this collections is restricted.\textsuperscript{88} One can however view part of its potential in an online video.\textsuperscript{89} A tool to search for the material aspects of a manuscript will also aide a scholar in answering his research question. Currently, most collections of digitised manuscripts can only be searched by bibliographic information and sometimes by keywords. Enabling scholars to search for manuscripts by the amount of columns on a page, the width of the margins or the material on which the text is written, will help them locate useful primary sources faster than before. This information could for example be encoded in XML mark-up or added by scholars in the form of tags, a topic that will be discussed more in depth later in this chapter.

### 2.2 Useful technologies considering manuscript research practice

This paragraph will discuss several technologies that will probably be useful to manuscript scholars considering their research practices in philology, textual criticism, history, codicology and palaeography. Both the benefits and downsides of these technologies will be considered and in some cases also the conditions that have to be met to benefit from them. The likelihood of adoption will be discussed in the next chapter.

\textsuperscript{84} A. Chassanoff, ‘Historians and the use of primary source materials in the digital age’, p. 470.  
\textsuperscript{85} P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.  
\textsuperscript{86} P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.  
\textsuperscript{87} P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.  
\textsuperscript{88} Genizah project \textless http://www.genizah.org/TheResearchPlatform.aspx\textgreater  (23 March 2014).  
\textsuperscript{89} The video can be found here: YouTube \textless https://www.youtube.com/watch?v=tkdht0LjczA\textgreater  (23 March 2014).
It is however important to first consider an act that is common to all aforementioned research practices: finding the manuscript that contains information that can help to answer one’s research question. This is commonly done by searching library catalogues. A digital information environment will also have such a catalogue containing record of the digitised manuscripts it holds. It is highly important that this catalogue contains complete and correct information on all manuscripts. If this is not the case, a record of a manuscript might not show up when a scholar searches for it and it will therefore be overlooked as a possible primary source.

Besides the coverage of the catalogue records, it is also important for scholars to know which manuscripts and collections can be searched. Ideally, all manuscripts in a library are available and searchable in a digital information environment, but this is currently not feasible for both economical and practical reasons. Therefore the scholars participating in the KB end-user study stated that they need information on the coverage of the digitised collection. By providing this information, scholars know ‘from which pond they are fishing’ when they look for primary sources.

The tools to search a collection can also contribute to the practical value of a digital information environment to find useful primary sources. If scholars are able to specify their search in detail it will help them find the right manuscript more easily. However, providing a scholar with too many search options will make the search tools less user-friendly. This friction is also discussed in the end-user research by the KB. Scholars in this focus group said to prefer easy and generic searching to start with and more enhanced search features to refine the results. This is probably the best way to combine a user friendly search tool with the more extensive search options that scholars need. The following examples will illustrate three different approaches for designing search tools in order to clarify the possibilities. For one, the Early Dutch books online project enables its users to search their collection based on just six parameters: title, printer/publisher, author, year, subject heading and language (fig. 2). Secondly, the Early European books online website provides its users with a much more extensive list of search fields (fig. 3). In addition to the parameters mentioned above it also

---

90 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
91 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
92 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
94 Early European books online <http://eeb.chadwyck.co.uk/search/initSearch.do> (5 February 2014).
allows its users to search for a bibliographic number, the country where the manuscript was created and a list of basic features of a manuscript page. Thirdly, a new platform of the KB, called Delpher, features a combined search engine of what was discussed before. The homepage of Delpher provides a simple search field, but the results from this simple search are accompanied by several options to narrow them down (fig. 4).

Once a scholar has located a useful manuscript for his research, it is common practice to gather more information about the manuscript. The BnF end-user study showed that scholars would like some of this information to be presented alongside the digitised manuscripts. They mentioned for example a scientific introduction to the corpus, a comprehensive bibliographic record of each manuscript and additional scientific data about the manuscripts, e.g. a summary of the work done on each corpus and manuscript. Providing this kind of information with the digitised manuscripts will aid a scholar in selecting a useful primary source for his research.

All in all, a digital information environment has to provide several tools to help scholars select a useful primary source for their research. After finding the right primary source, the following technologies can aid scholars in their research efforts.

2.2.1 Downloading and software

The technology to download a digitised manuscript is already commonly incorporated in many online collections. For example, the Early European books online website allows scholars to download the scans of digitised materials as PDF-file and save them on a local drive.

The greatest benefit of this technology is that it offers scholars the possibility to consult their primary source at any time or place. Scholars are no longer dependent on the opening hours of a library or the availability of an internet connection. Moreover, once the images are saved locally, scholars can zoom in on the images to a far greater extent than the human eye can. Depending on the file format, they can also change the contrast and luminosity of the images and add notes. Currently most libraries with a digitised collection offer their downloads as a PDF-file, which is open to zooming and the addition of notes.

An important overarching condition is that the libraries have to offer high quality

95 P. Chevallier et. al., ‘Consultation of manuscripts online’, n. pag.
96 Early European books online <http://eeb.chadwyck.co.uk.access.authkb.kb.nl/search/results.do> (18 February 2014).
images for downloading. Currently there are some online collections that present high quality images online, but only offer low quality downloads. For example the *Early Dutch books online* project only offers a black and white version of the images as a download (fig. 5). These images are commonly not useful for scholarly research and thus downloading them is useless.

Besides downloading the images and applying the traditional research practices, scholars can also choose to research the downloaded files with computer software. This software will probably not be incorporated in the digital information environment for technical reasons. However, analysing manuscripts with software offers some interesting new research possibilities that are closely intertwined with the development of such a platform. Therefore, two possibilities for research with computer software and its links with a digital information environment will be discussed here.

The first example is software that can generate a stemma codicum. As discussed in the previous chapter, making stemmas is a common research practice in textual criticism. In his article called ‘Making stemmas with small samples, and digital approaches to publishing them’, A. Hall describes how he uses a program called *Pars* to create a stemma of an Icelandic text. Although the *Pars* software is not specifically designed for language analysis, it can be used for this purpose. Hall’s research practice starts with selecting samples from the text. He then transcribes these samples and creates a spreadsheet of aligned readings which is converted into a file format readable for the *Pars* software. The stemma that is subsequently created by the software is deemed ‘very similar’ to a previously manually created stemma.

A digital information environment can be linked to this type of research practice in two ways. For one, it can contribute to this research practice by providing a scholar with the full-text files of the textual content of manuscripts. Scholars can then skip the transcription phase of the research practice as described above. Secondly, the digital files containing the raw data and the digital stemma can be linked to a digitised manuscript in a digital information environment. This practice will allow other scholars to see, reproduce and re-use the data that a scholar has generated. Such transparency and verifiability enhances the

98 *Early Dutch books online* <http://www.earlydutchbooksonline.nl/nl/view/image/dpo%3A3804%3Ampeg21%3A0007> (18 February 2014).
99 A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
All the following information on *Pars* was found in this article.
100 A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
credibility of scholarly output. The use of software is of course not required for sharing stemmas, but being able to share the raw data and have other reproduce the research is easier when using software.

An important benefit of using the *Pars* software is that it can handle larger amounts of samples than a human could in the same time span. Although sampling has always been a part of stemmatology, it is commonly accepted that a larger amount of samples provides a more accurate stemma.\(^{101}\) Since software can analyse more samples in the same time span, it can create a more accurate stemma without prolonging the research phase.

There are of course also some downsides to using software to create stemmas. For one, the stemma will be based solely on the textual content and disregard any other features. Secondly, there are always restrictions imposed by the software that is used.\(^{102}\) This is clearly the case with the *Pars* software since it is not designed for language analysis. This causes several limitations and the need to reformat the full-text files. But even software that is specifically designed for language analysis will have certain limitations. Overcoming these limitations would require reprogramming the software, something that scholars probably cannot do, thus making them dependent on software developers for their research tools. Thirdly, scholars will have to learn how to work with this software in order to produce the results they are looking for. Learning how to work with new research software can take up a lot of time and it is a valid question whether or not scholars will be willing to invest that much. This question will be discussed in the next chapter of this thesis.

The second example of software for manuscript research is the *System for palaeographic inspections (SPI)*. This software can do a basic palaeographic study based on digital images of a manuscript. It is designed to show, compare and categorise the graphical relationship between letterforms. It can also compare the letterforms in a manuscript to a corpus of dated and localised manuscripts to estimate when and where it was produced.\(^{103}\) To use the software a scholar has to separate the letterforms within an image and introduce several examples of the same letter to the software. By comparing these examples, *SPI* will create a standard pattern for each letter which can than be categorised and compared to that of

---

103 A. Ciula, ‘Digital palaeography: using the digital representation of medieval script to support palaeographic analysis’, *Digital medievalist*, 1 (2005), n. pag. <http://digitalmedievalist.org/journal/1.1/ciula/> (7 February 2014). All the following information on *SPI* was found in this article.
other manuscripts.

Software like SPI can be beneficial to scholars who are not trained to do a palaeographic study. Moreover, frequent use of this kind of software will make it more efficient for all users. Once SPI recognizes and saves more letterforms from dated and localised manuscripts, it will date and localise newly introduced letterforms more accurately. The fact that SPI does not rely on arbitrary terminology to categorise handwritings is another benefit. SPI considers the actual letterforms as means to classify handwritings, not the attributed terminology.

A downside of using SPI to analyse handwritings is that a scholar can only research manuscripts that are digitised and available in a file format that adheres to a certain standard. In the case of SPI this means that at least four digitised pages of script are needed in 300 dpi and in a TIFF file-format. These parameters should be taken into consideration when institutions digitise manuscripts and decide which type of files will be available in the digital information environment. As with the Pars software, scholars will also have to learn how to work with SPI and again the software is not as sophisticated yet that it can replace a trained human eye. However, the current results are called ‘promising’ and it can be a good start for further research.

2.2.2 Adding information

An online information environment can also offer tools to add information to a digitised manuscript, with the option to share it with peers. The term ‘information’ has a broad meaning in this thesis, from a description of the material aspects of a manuscript in one or two words, to a more elaborate discussion of social developments at the time when a manuscript was produced. This paragraph will first discuss how adding information aligns with current research practices in philology, history, codicology and textual criticism. It will then discuss the benefits and downsides of two technologies currently available to do so: XML and tagging.

Several scholars have written about the possibilities of online, collaborative scholarship by allowing scholars to add information to digitised primary sources. This

---

M.N. Smith, ‘Electronic scholarly editing’, S. Schreibman, R. Siemens and J. Unsworth, A
practice can also be useful for philologists considering its main aim: lowering interpretational barriers to older textual contents by providing readers information about the context in which the manuscript should be viewed, the handwriting and the use of language. Providing such information alongside the digitised manuscript will lower the interpretational barriers for scholars without them having to consult secondary literature. Adding information to the digitised manuscript thus helps philologists to achieve their scholarly goals more efficiently. It also allows for collaboration by letting each scholar add information from his own area of expertise. For example, the analysis of the handwriting can be done by palaeographers, while historians can add information for the social and political contextualisation of the manuscript. This aligns nicely with the research practice of philology which already combines several disciplines of research. Moreover, it has the potential to produce a high quality philological analysis.

For medievalists it can also be useful to add information to a manuscript. As discussed in the previous chapter, the historical research method commonly consists of external criticism, internal criticism and synthetic operations. Adding information to a manuscript can be beneficial to the practice of external and internal criticism. External criticism concerns the authentication of the primary source and establishing the most accurate form of the text. Most information that medievalists need to authenticate a manuscript can be derived from the bibliographic information provided by the institution that holds the manuscript. However, if this bibliographic record is incomplete, scholars need a tool to add information to it. This can facilitate easier external criticism for other scholars. Adding an accurate form of the text to the digitised primary source can provide the other component of external criticism. In this respect, medievalists can probably also benefit from the stemmas added by scholars from the discipline of textual criticism. Internal criticism, i.e. the assessment of the credibility of the textual content, can also be aided by added information. Once a scholar has made this assessment, it is useful for other scholars to read his considerations. Obviously there can be several viewpoints, but reading the assessments of peers can help to draw one's own conclusions. There is however a clear difference between the benefits of adding information for philologists and for medievalists. Adding information can help philologists to achieve their scholarly goals, but for medievalists it is mainly a way to help others with their research and in turn receive help from their peers. Medievalists will have to invest in this technology in order to really benefit from it. The same goes for book historians. Adding information from


their research is above all a good way to share this information and benefit other scholars.

Adding information will be beneficial to codicologists in yet another way: it can allow them to search databases of digitised manuscripts based on the material aspects of these manuscripts. Currently, most databases can only be searched by bibliographic information and sometimes by keywords. But for codicologists it would also be very useful to search for manuscripts by certain material aspects, for example manuscripts written on parchment or manuscripts with wide margins. This is possible by adding this kind of information to the digitised manuscripts and offering tools to search it. It would be most beneficial if scholars could add this information, since they know most about these materials. The technological possibilities of searching manuscript collections based on added information will be discussed more extensively below.

Textual criticism can also benefit from tools to add information to a digitised manuscript. For example, P. Monella suggests that textual criticism can benefit from linking and aligning several full-text versions of one text. Instead of adding pieces of information, he envisions scholars will add a complete text to a digitised manuscript. This can facilitate comparison of differ versions of the text and possibly help to uncover the authoritative text. Such a body of linked texts can also allow a scholar to display the differences between these texts more easily and each scholar can verify these observations. Besides different versions of the text, this technology will also enable scholars to add their stemma codicum to the primary source material. Combining a stemma with its source will enhance the verifiability of the stemma as was discussed before.

Adding information to a digitised manuscript has three additional benefits that are relevant to all the aforementioned research practices. For one, scholars from all over the world can research the same manuscript, add information to it and react to each other’s findings. It will thus simplify collaboration between scholars from different institution. Secondly, it will lead to more manuscripts being disclosed and newly added information can immediately be used in research. And thirdly, scholars can more easily compare and contrast different interpretations, and use the one they think are correct or relevant to their research. This is especially useful in manuscript research where several interpretations often coexist.

---

106 P. Monella, ‘Towards a digital model to edit the different paratextuality levels within a textual tradition’, n. pag.
107 P. Monella, ‘Towards a digital model to edit the different paratextuality levels within a textual tradition’, n. pag.
Adding information also has two clear downsides for each of the aforementioned research practices. The most important downside is that the added information online is probably always one step behind on published information. All the aforementioned applications assume that scholars will share their research or part of their research online, but this kind of publication is not rewarded with tenure or professional recognition.\textsuperscript{110} It is therefore more likely that scholars will only add information that is already published and for which they have already received professional rewards. Another downside is the possibility that all these bits of information will form a cluttered whole. Scholars who research manuscripts come from different academic disciplines and thus have a different focus. Adding information from all these disciplines to one digitised manuscript can become confusing. To prevent this from happening there has to be a way to sort and search the added information.

There are also several conditions that have to be met in order to really benefit from adding information to digitised manuscripts. For one, it is important that not everyone can add information. Access to these tools should be restricted to scholars in order to ensure a scholarly level of quality.\textsuperscript{111} This can be achieved by only allowing people with a user account to add information. In order to get such a user account there should be some kind of basic screening, for example whether that person is enlisted with a university. Independent scholars should also be able to apply for access separately.

Secondly, there has to be an editorial staff to maintain the scholarly level of quality.\textsuperscript{112} A good example of a digital information environment with such an editorial staff is the Digital Classicist wiki.\textsuperscript{113} On this online platform scholars can share professional information like best practices, new software and recently published research. The editorial staff monitors the quality of the ongoing discussions and occasionally opens up new topics for discussion.

The option to legally protect shared research information from plagiarism is the third condition for this kind of digital and collaborative scholarship to work.\textsuperscript{114} Research information and new interpretations are the core merit of a scholar's research. Therefore, scholars will not share this information online without clear intellectual property laws. Until now there has been a lack of clarity about intellectual property rights online which prevents

\textsuperscript{110} C. Dallas, ‘Humanistic research, information resources and electronic communication’, p. 229.
\textsuperscript{111} S. Mahony, ‘Research communities and open collaboration’, n. pag.
\textsuperscript{112} S. Mahony, ‘Research communities and open collaboration’, n. pag.
\textsuperscript{114} Digital classicist wiki <http://wiki.digitalclassicist.org/Main_Page> (3 January 2014).
\textsuperscript{114} S. Mahony, ‘Research communities and open collaboration’, n. pag.
many scholars from sharing information. However, creative commons licenses (CC licenses) now allow scholars to legally protect their work online. The possibilities of these licenses must be explained to scholars in order to create a trustworthy environment in which they can comfortably share ideas and information. Several scholarly bloggers have already considered how CC licenses can support the kind of open scholarship that is also discussed here. Some scholarly initiatives even actively use the licenses, for example the MIT open courseware website which contains more than 2,200 courses all protected by CC licenses. Publishing research information under a CC license is however not yet a common practice. Possibly because not all scholars trust these licenses the way they trust traditional copyright on printed materials. Moreover, some scholars do not support the kind of open scholarship that is proposed here. These scholars are therefore unwilling to present their research online even if there are trustworthy intellectual property laws in place. But for those scholars who support open scholarship in a digital information environment, it is important to know how they can protect their contributions.

The fourth and final condition is that the tool to add information has to be user-friendly and it should add the information in a format that can be shared between platforms. Implementing a user-friendly tool will make it easier for scholars to adopt it. Since it is probable that digitised manuscripts will be spread over several databases and web environments it is important that these platforms can share their added information.

This final note aims to suggest a possibly useful additional feature, namely a tool to respond to the information added by peers. These tools already exist, for example in the platform Diigo where scholars can respond to each other’s notes. If used, such a tool will benefit scholarly communication beyond ones local network of peers. It also has the potential

116 Creative commons <http://creativecommons.org/> (28 April 2014).
117 For example:
to present a good overview of topics that are still under discussion. However, if longer
discussions start to form these might clutter the added information, which is the reason why
this tool is only suggested as a possibly useful feature. Only a practical application of this tool
can show whether it will be beneficial or only cause clutter.

There are two commonly used technologies to add information to a digital object:
XML and tagging. Both of these technologies will be shortly introduced here and their most
important benefits and downsides will be discussed.

XML stands for eXtensible Markup Language, a technology that basically provides a
way to explicitly state what kind of data a document holds.¹²⁰ For example, in a letter one can
encode who the author and receiver are and when the letter was written (fig. 6). This encoding
is done by adding a tag to the name of the author and receiver and the date on the letter. After
assigning these tags to several letters one can easily search for and sort all the encoded
information. XML mark-up can be used for much more than facilitating searches and sorting
information, but these aspects are most important in this context. It is commonly advised to
create a list of tags that can be used in the XML document, especially if documents from
several producers will be combined in one platform.¹²¹ Such a list is called a doctype
declaration (DTD) which can be based on an existing standard or independently designed.
The TEI-standard is currently the most commonly used standard for adding XML mark-up to
a manuscript.¹²² Chapter ten of the TEI-guide focuses specifically on encoding manuscripts
and provides a set of useful tags.

Using XML to add information to a digital object has several benefits. The amount
and kind of tags that can be used is not predefined, which presents scholars with enough
flexibility to encode manuscripts. It is however very likely that scholars will not be
completely free in their choice of tags, because they have to adhere to a DTD. This will
provide a controlled vocabulary which will benefit the regularity of the added information.
This will, amongst other things, make it easier to search the documents by preventing the use
of synonyms. Moreover, files that are enriched with XML and adhere to the same DTD can
easily be shared between different platforms. This would make it easier to add new encoded

¹²⁰ ‘Introduction to XML’, W3schools <http://www.w3schools.com/xml/xml_whatis.asp> (5
February 2014).
¹²¹ ‘XML DTD’, W3schools <http://www.w3schools.com/xml/xml_dtd.asp> (5 February
2014).
¹²² ‘P5: Guidelines for electronic text encoding and interchange’, Text encoding initiative, 20
2014).
materials to the collection or combine collections that were encoded according to the same DTD.

There are already several scholars who have discussed a practical application of XML in manuscript research. For example, E. Pierazzo and P. Stokes have described how XML mark-up can be used to encode the material aspects of manuscripts. As already mentioned before, this can benefit codicologists. Since XML uses a controlled vocabulary as defined in the DTD, the encoded manuscripts can easily be searched by using the keywords that were also used as tags. This would however require agreement on the terminology to describe the material aspects. P. Monella provides another practical application of XML, namely linking several layers of text together. As already discussed, this technique can clearly benefit textual criticism. However, reading Monella's article makes it clear that linking and aligning texts via XML is quite complex. This is a big downside of this application since it will probably prevent scholars from easily adopting this technology. Currently both of these applications of XML are not actually implemented on collections of digitised manuscripts yet.

Even though it has great potential, XML also has a clear downside. Commonly, manuscript scholars do not know how to work with XML and since it is quite a complex technology they will have to invest quite some valuable time to learn how to work with it. It is therefore unlikely that scholars will easily adopt XML to add information to a digitised manuscript. A more probable scenario is to let IT specialists do the actual encoding. These people will have the technical skills, but in turn lack the in depth knowledge that manuscript scholars have which can harm the quality of the added information. Moreover, it would be a costly affair since adding XML mark-up is time consuming. This thesis therefore favours a more user-friendly technology to add information and thus allows scholars to do it themselves: social tagging.

Social tagging is a less complex and thus more user-friendly technology to add

---

124 P. Monella, ‘Towards a digital model to edit the different paratextuality levels within a textual tradition’, n. pag.
information to a digitised manuscript. It allows scholars to add uncontrolled keywords, named tags, to digitised manuscripts. These tags can be used to organise information in a personal information space, they can be shared with others and they can be searched. Currently, tags are mainly one or two words long. Learning how to add tags is probably not a very time consuming task since most tools to add tags are very user-friendly.

Good examples of social tagging in the cultural heritage field can be found on the websites of The powerhouse museum (fig. 7) and The Brooklyn museum (fig. 8). Visitors of these websites are invited to add tags to the images of artefacts in the collection. Social tagging can be used in a similar way to benefit codicologists. The one or two word tags can easily describe the material aspects of a manuscript, for example tags like ‘parchment’, ‘two columns’, ‘illustration’. These tags can then be searched and thus scholars will create a collection of manuscripts that can be searched based on material aspects of manuscripts. It is currently not a common practice to add longer tags or tags with links to longer texts, but is technically possible. Social tagging can therefore also be useful in philology, historical research and textual criticism. The benefit of using tags in these disciplines is that any and all information can be added since there are no restrictions to the vocabulary that can be used.

There are however also some downsides to social tagging. For one, the manuscripts that have not been tagged will not show up in one’s search results. Moreover, if a scholar adds a faulty tag, a manuscript will unjustly show up in the search results. So unless the whole collection is enriched with correct tags, only part of the collection can be accessed by searching based on tags. The fact that tagging relies on an uncontrolled vocabulary is also often mentioned as a downside. This thesis will expand some more on this issue since it is a relevant discussion in this context.

An important benefit of an uncontrolled vocabulary is that it enables scholars to add the information they deem valuable in their own words. This allows for a more nuanced word choice that respects personal and cultural differences. However, this nuancing does create

---

more noise in the search results when compared to a controlled vocabulary.\textsuperscript{130} For example, the lack of control over synonyms often produces incomplete search results unless a scholar manages to search for all the synonyms that were used.\textsuperscript{131}

It seems that some kind of ontology or guideline is needed in order to make social tagging and searching these tags useful for scholars. Several scholars have thought of ways to implement a controlled vocabulary to social tagging. One suggestion often coined is the use of ‘common tags’, a list of suggested tags that have been used by others (fig. 9).\textsuperscript{132} When scholars re-use these tags, there will be less variation in the words used to describe the materials. Another interesting approach is proposed by C. van Damme, M. Hepp and K. Siorpaes, who discuss a way to integrate ontologies and folksonomies, i.e. the terms used in an environment without a controlled vocabulary.\textsuperscript{133} The details of this research are beyond the scope of this thesis, but this article shows that there are possibilities to overcome the downsides of an uncontrolled vocabulary. There are however also scholars who argue that we should not even try to control the vocabulary in social tagging. For example, C. Shirky argues that controlling the vocabulary, by for example collapsing synonyms, will lead to a loss of information.\textsuperscript{134} He argues that the internet finally provides ‘organic ways of organizing information’, instead of forced categories that are imposed by a single person or organisation.\textsuperscript{135} This newfound freedom would be lost if an overarching institution would control the vocabulary used in tags.

Clearly, these are two opposing views on the uncontrolled vocabulary in social tagging. This thesis would like to argue that there has to be some form of control over the vocabulary that can be used to describe the material aspects of a manuscript. An uncontrolled vocabulary would create too much noise in one’s search results, which is problematic since scholars possibly want to search a collection based on these tags. Although controlling the vocabulary might indeed lead to a more forced categorization, its benefits outweigh the

\textsuperscript{130} G. Macgregor and E. McCulloch, ‘Collaborative tagging as a knowledge organisation and resource discovery tool’, p. 295.
\textsuperscript{132} G. Macgregor and E. McCulloch, ‘Collaborative tagging as a knowledge organisation and resource discovery tool’, p. 295.
\textsuperscript{133} C. van Damme et. al., ‘FolksOntology: An integrated approach for turning folksonomies into ontologies’, pp. 1-14.
\textsuperscript{134} C. Shirky, ‘Ontology is overrated: Categories, links, and tags’, 2005, n. pag.
\textsuperscript{135} C. Shirky, ‘Ontology is overrated: Categories, links, and tags’, 2005, n. pag.
disadvantages. There is however less need for a controlled vocabulary in philology, history and textual criticism. The tags added in these disciplines will be more elaborate than one or two words which makes controlling the vocabulary impossible. Moreover, these tags will not be used as search terms to find certain manuscripts within a collection which makes standardisation unnecessary. For scholars from these disciplines it will however be useful if they can search within tags, especially if the number of tags per manuscript grows. It is important for all disciplines to make a clear distinction between personal and social tags. The tags for personal information management, like ‘important' or 'read later’, should not be shared since these will only create noise in the search results.

2.2.3 Scholarly communication

Manuscript scholars seem to be interested in the possibilities of sharing information via digital media. For example via blogs about research projects; twitter accounts about manuscripts and manuscript research; and digital fora like the Digital medievalist and the Digital classicist wiki. If such digital scholarly communication is supported by a well formed digital information environment, it can become a very useful platform for manuscript scholars. Three currently possible forms of scholarly communication in a digital information environment will be discussed here: social reference management (SRM), digital fora and publishing in the area of grey literature.

SRM has already been mentioned in this chapter as a useful technology to aid scholars in finding secondary literature. In short, SRM provides scholars with tools to share and index secondary sources online. There are already several platforms that facilitate SRM, for example CiteULike (fig. 10) and Bibsonomy (fig. 11). Both of these platforms offer tools to share bookmarks of digital sources and allow users can add tags to these bookmarks to facilitate searching. However, since much of the literature on manuscripts is currently not digitally available, it would be useful if scholars can also add references to analogue materials.

SRM assumes that it is easier to find useful sources when one can rely on the recommendations of peers. Research has indeed shown that scholars using SRM were able to find more useful sources for their research than scholars who do not use SRM. It is

---

136 CiteULike <http://www.citeulike.org/> (7 February 2014).
137 A more in depth discussion of SRM will follow on page twelve and thirteen.
138 H. Alhoori and R. Furuta, ‘Understanding the dynamic scholarly research needs and behavior as applied to social reference management’, p. 170.
therefore a useful addition to a digital information environment. However, scholars participating in the KB end-user study reported that they are only interested in recommendations from identified peers.\textsuperscript{139} They want to know who recommended an article or a book in order to assess the value of the recommendation. Coincidentally, both aforementioned websites give users the opportunity to add a profile with personal and professional information which allows scholars to check who recommended an article. Users are however not obliged to add this information; therefore some recommendations will come from unknown peers. In practice this does not seem to influence scholars’ willingness to use CiteULike and Bibsonomy. In 2009 CiteULike reported over 200,000 registered users and the website was visited almost 270,000 times a month by returning visitors.\textsuperscript{140} Bibsonomy reported over 190,000 users in 2010.\textsuperscript{141}

Besides sharing references to secondary literature it can also be useful to share more general information about one’s academic discipline, like best practices and links to interesting new software. This kind of information can be shared most easily in a digital forum. Good examples of this kind of fora are the aforementioned Digital classicist wiki and the Digital medievalist. One can of course question whether such a forum should be a part of a digital information environment since there are already independent fora that seem to be successful so far. However, integrating a forum would create a wealth of information all accessible via one platform. Both the primary sources for research and information about research practices can thus be gathered in one location. At the very least, the existing fora should therefore be referenced in a digital information environment.

Actual publications can also be shared in a digital information environment. This kind of digital publishing without the restraints and demands of a publishing house or journal can make the distribution of one’s research much easier.\textsuperscript{142} Especially when a growing number of scholars make use of the digital information environment and thus find and use research published there. Moreover, being able to publish research next to its primary source can enable peers to actually see what is described.\textsuperscript{143} Currently, publishing online is commonly not preceded by a peer-review process or a selection and editorial process as is common in

\textsuperscript{139} Interview with Rosemarie Pomp-Blangé, 21 March 2014.
\textsuperscript{140} CiteULike <http://www.citeulike.org/groupforum/1784> (7 February 2014).
\textsuperscript{142} R. Procter, et. al., ‘Adoption and use of web 2.0 in scholarly communications’, p. 4048.
\textsuperscript{143} W.G. Thomas, ‘Computing and the historical imagination’, p. 66.
formal academic publishing. Therefore, online publications are not officially recognised by the academic community as reasons for tenure or promotion. It is thus highly unlikely that publishing in a digital information environment will soon be viewed as an alternative to publishing in traditional scholarly media. It is more likely that scholars will view this kind of publishing as a form of publishing in the area of grey literature and possibly as a good means for pre-printing. Pre-printing is a way of publishing one's research or part of the research, often in preliminary form, before the actual publication. It gives peers an opportunity to react to the research before it is officially published. A digital information environment can provide scholars with an easy opportunity to do this, especially since it can also offer tools for peers to react to the publication.

2.3 Conclusion

This chapter has considered the question what characteristics, services and tools a digital information environment should ideally have to align with current manuscript research practices. The benefits and downsides of these services and tools were also discussed.

Concerning the general research practices of manuscript scholars it is imperative that the images of the manuscripts and the derivative files are of high quality. Any flaws in the digital reproductions can cause a bias in a scholar's research. To facilitate individual research it is probably useful if scholars are provided with a personal storage space. For research in teams, the digital information environment has to feature tools to share digital files and comment on each other’s work. Considering the fact that annotation is a common practice in both individual research and research in teams it would be useful to add tools for digital note taking with the option to share these notes. A digital ruler and tools to change the display of the images are required to research the material aspects of a manuscript.

The research practices of textual criticism and palaeography can benefit from tools to download the full-text files and images of the digitised manuscript so these can be analysed with computer software. The biggest benefits of using software to analyse manuscripts are the possibility to work faster and thus analyse more samples in the same time span, and the possibility to share the raw data and research results more easily to aid transparency and verifiability of the research. The biggest downside of this practice is that current software is

---

144 C. Dallas, ‘Humanistic research, information resources and electronic communication’, p.229.
C. Borgman, Scholarship in the digital age, p. 223.
often not specifically designed for manuscript research and therefore has certain limitations. The research practices of philology, history, codicology and textual criticism can all benefit from tools to add information to a manuscript. By adding information to a digitised manuscript, philologists will take away interpretational barriers of older texts more easily. Medievalists can organise and share their information on internal and external criticism. Codicologists will benefit from the fact that adding information about the material aspects of a manuscript allows them to search the manuscripts based on these aspects. Textual criticism can benefit from linking and aligning several versions of one text to aid comparison of these text and help to uncover the authoritative text. Moreover, adding stemmas to a primary source can enhance its transparency and verifiability. Two technologies, namely XML and social tagging, can help manuscript scholars achieve what is outlined above. The biggest benefit of XML is its controlled vocabulary, which makes it easy to search the encoded information and exchange enriched files so long as they are based on the same DTD. The biggest drawback of XML is the fact that it is a complicated technology and as such it can take quite some time to become proficient at using it. The biggest benefit of social tagging is its user-friendliness, but the uncontrolled vocabulary can create a lot of noise in one’s search results.

The final three technologies that were discussed in this chapter are all related to scholarly communication, namely SRM, fora and publishing in a digital information environment. SRM can aid scholars in finding useful secondary sources based on recommendations of their peers. However, scholars should be either willing to accept recommendations from unidentified peers, or they can only benefit from the recommendations of users with a personal profile. Fora can be an international informal meeting point for scholars who are interested in manuscript research, but its usefulness highly depends on the number of active users. Formal publication in a digital information environment are not likely to occur soon. However, it can be a useful platform for pre-printing research output since it can easily allow scholars to react to the research of their peers.
Chapter 3: Future prospects for manuscript research

The previous chapter of this thesis discussed what characteristics, tools and services a digital information environment should have, considering the current research practices of manuscript scholars. This chapter will consider the opposite influence: how adopting a digital information environment can influence manuscript research practices. This chapter thus assumes that using new research tools will influence scholarly research practices, a notion that is confirmed by several examples. The logic behind this influence is that new research tools commonly provide new ways to find, access and/or study scholar’s primary sources. In the case of manuscript research, one can think of facilitating access to materials that were previously hard to research, due to e.g. their size or the vulnerability of the material. It is difficult to predict the influence that new research tools might have, and the assessment of their added value is highly subjective. This thesis will therefore only discuss the influences that are most likely to occur and the ones that are based on the services and tools discussed in the previous chapter. Any value judgements are to be considered subjective. This chapter will also discuss the likelihood of uptake of a digital information environment and the tools that were discussed. This is an important factor in the amount of influence that a new research tool can have because tools that are easy to adopt will most likely be used more often and therefore have more influence on scholarly research practices.

3.1 The influence of a digital information environment on manuscript research practices

As discussed above, this chapter assumes that using a digital information environment will bring about changes in the research practice and output of manuscript scholars. This paragraph will discuss the influence of the general tools and services described in the previous chapter. Next, it will discuss the possible influence of tools and services for downloading and using software, adding information and scholarly communication.

---

But first, it is important to define the collection that can be accessed through the digital information environment as discussed here. The coverage of the collection will partly determine how much influence a digital information environment can have on manuscript research practice. A small collection from one institution can only be used for research that happens to focus on that collection, which makes its impact on manuscript research as a whole quite small. An extensive collection from several institutions can be used for a wider variety of research projects and will thus have more influence on manuscript research as a whole. This paragraph will presume that the collection accessible through the digital information environment is such an extensive collection. There are already several of these collections accessible online, for example, the collection of the *Early European books online* portal. This portal contains collections from five different institutions with works from the 15th, 16th and 17th century.

3.1.1 The influence of general tools and services

Following the typical academic research workflow, the influence of a digital information environment can first be observed in the way scholars search for a useful primary source. A digital information environment can improve and thus speed up the discoverability of useful primary sources, especially if scholars can search several collections with one search engine. This will increase their chances of finding a useful primary source. Moreover, they might find sources that would have been overlooked when only searching through familiar collections. Searching for useful primary sources can be further enhanced if the digitised manuscripts contain encoded information on the material aspects. This would be especially useful for scholars who are interested in manuscripts with specific material aspects.

Some plausible influences of a digital information environment are related to the fact that it contains digitised manuscripts instead of analogue manuscripts. When scholars use digitised manuscripts, they can research materials that are otherwise hard to access. For example, rare and vulnerable materials, or very large or small manuscripts. It can also allow scholars to work with geographically dispersed manuscripts since digital objects are not bound by physical distance, but can be brought together in a digital information environment.

---

148 M. Deegan and S. Tanner, ‘Conversion of Primary Sources’, p. 491.

---

46
environment.\textsuperscript{150} This kind of research is currently a costly affair due to the required travelling and it is therefore not feasible for many scholars. The way these materials can be studied also changes when the primary sources are digital instead of analogue. For example, tools to change the contrast, luminosity and colour filters of a digital image can enable scholars to see aspects of a manuscript that cannot be observed with the naked eye.\textsuperscript{151} Zooming in on the images can uncover the smallest details in decorations and images. Such tools can enable scholars to answer research questions based on more detailed information and, in some cases, it can allow them to answer different questions than before.

The previous chapter also suggested that a digital information environment can facilitate research in teams. However, currently available tools for collaboration, like Diigo and Google docs, do not seem to have any substantial effect on the amount of collaboration in manuscript research.\textsuperscript{152} Therefore, it might be too simplistic to expect a real increase in research done by teams simply because a digital information environment facilitates it. Moreover, team efforts are currently only a small part of manuscript research. It should probably first become a more common practice in real life before it can transfer to the digital realm. The same reasoning applies to the influence of a digital information environment on co-authorship.\textsuperscript{153} A digital information environment can allow several scholars to work on one text at the same time. There are already several platforms, like Google docs, that offer this service.\textsuperscript{154} However, this has not led to an increase of co-authored research output. It would thus again be too simplistic to assume that facilitating it will indeed cause an increase of co-authorship. One cannot disregard the fact that manuscript research has been an individual endeavour for a long time. Therefore the turn to co-authorship of research done in teams might not be made that easily, even if better tools to do so are available.

\textsuperscript{150} T. Burrows, ‘Applying semantic web technologies to medieval manuscript research’, p. 123.
\textsuperscript{151} M. Deegan and S. Tanner, ‘Conversion of primary sources’, p. 491.
\textsuperscript{152} There is not statistical data to support this observation, but a look at the research presented in journals on manuscript research (like the aforementioned Queeste and Manuscripta) shows that most research is still done individually.
\textsuperscript{153} S. Mahony, ‘Research communities and open collaboration: the example of the Digital Classicist wiki’, n. pag.
3.1.2 The influence of downloading and using software

Being able to download a manuscript and consult it digitally can influence a scholar's research practice in several small ways. For one, scholars can work on their research in any place, at any time instead of being bound to the opening hours of the library that holds the analogue manuscript. Secondly, scholars can add notes to the manuscript, which cannot be done with analogue manuscripts since it would harm the materials. Thirdly, tools to digitally zoom in on an image or change the display can influence the way scholars engage with a manuscript and the kind of information that can be accessed.

Scholars' research practices will change more drastically if they use software to study the downloaded file. The influence of using software on the research practices of textual criticism and palaeography will be discussed here. The software that will be considered is already discussed before, namely the Pars software to create stemmas and the System for palaeographic inspections (SPI) for palaeographic research.

The first influence stems from the fact that computers are able to process more data than humans can in the same time span. Using a computer with the right software can thus speed up manuscript research tremendously and help to increase its volume of output. In the case of textual criticism, using software can thereby make the research results more accurate because the stemmas can be based on longer samples without prolonging the research phase. Moreover, the increased speed can make quantitative research of larger collections more feasible than it currently is. There are however two factors that negatively affect the speed of computer analysis: the textual content of a manuscript often has to be reformatted and the scholar must invest time to learn how to work with the software. Reformatting the files will, for the foreseeable future, remain a time-consuming part of this research practice. However, learning how to work with the software is a one-time investment. Once a scholar becomes more skilled at working with the software and reformatting files, he can start to benefit from the enhanced speed of doing research with software. The benefits will be even more substantial once software development progresses and more specific software for manuscript research is created.

Secondly, the use of software to analyse manuscripts can influence the amount of exploratory analysis that is done. Exploratory analysis, i.e. research without a predefined

---

155 A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
156 A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
hypothesis, can be a useful tool, for example when a scholar is confronted with a manuscript or collection that has not been studied before. Even though it can produce unexpected results and give rise to new questions, it is currently not a common practice since it costs too much time. However, the use of software can speed up an exploratory analysis and thus make it more attractive for scholars to do.\textsuperscript{157}

A third influence of using software for manuscript research is an increased transparency of research.\textsuperscript{158} When scholars use software, they can more easily share the raw data and output with their peers since all these will be digital files. By sharing these files, their peers can get a clear insight into the research methods that were used, which increases the transparency and verifiability of the research. Scholars will however have to incorporate more technical information in their research output, for example, about the software settings since software can produce different results when different settings are used.

Fourthly, when scholars use software to analyse manuscripts it can open up a new class of evidence that was otherwise not accessible.\textsuperscript{159} One can think of stemmas compiled from larger samples. If a scholar has to create these by hand it would take a lot of time, but with software to analyse the text it can be done much faster. These stemmas will give a scholar access to evidence that he would otherwise not have access to. The same goes for the analysis of the material aspects of a manuscript. By changing the contrast, luminosity and colour filters on a digital image, a scholar can see aspects of a manuscript that cannot be observed with the naked eye.\textsuperscript{160}

Finally, the fact that currently most software is not designed for analysing manuscripts can also influence manuscript research. For example, the \textit{Pars} software only compares the textual content of a manuscript, disregarding other possibly relevant features like the layout.\textsuperscript{161} But even when software is designed specifically for manuscript research, it will always impose certain restrictions. Traditional research is also bound by certain restrictions, but scholars can often find a way around them with a bit of creativity. In the case of software, getting around the restrictions would mean reprogramming the software which is something

\textsuperscript{158} A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
\textsuperscript{159} H. Craig, ‘Stylistic analysis and authorship studies’, p. 280.
\textsuperscript{160} M. Deegan and S. Tanner, ‘Conversion of Primary Sources’, p. 491.
E. Collins and M. Jubb, ‘How do researchers in the humanities use information resources?’ p. 183.
\textsuperscript{161} W.G. Thomas, ‘Computing and the historical imagination’, p. 61.
most manuscripts scholars are not able to do. Scholars will thus remain bound by these restrictions and dependent on software programmers. Or they might decide not to use the software at all.

It is important to note that even though software can be a useful research tool, it cannot replace a scholar’s research effort. The results that the software produces on manuscripts will always have to be analysed and interpreted by a scholar, otherwise it is just raw data.

3.1.3 The influence of adding information

The previous chapter has already discussed how adding data can support research in philology, history, codicology and textual criticism. The influence of using XML and tagging on these research practices will be discussed here.

An increased openness of scholarship will be the biggest influence that adding information can have on all three aforementioned disciplines. By adding information, scholars will be able to easily compare and contrast interpretations and read information found by peers. It can also influence the amount of collaboration between manuscript scholars of all aforementioned disciplines. If all scholars can add information and react to each other’s observations this can be considered a new form of collaboration. This kind of collaboration is easily accessible and can produce high quality information since scholars can choose to respond only to topics within their expertise. Moreover, adding information can have a big influence on a scholar's research practice. In addition to writing full research articles, they would dedicate part of their time to adding their research findings to a digitised manuscript. Lastly, adding information can benefit the disclosure of more manuscripts without the need to work with a scholarly publisher or go through peer-review process. This disclosure will not be in the form of a formal publication, but by making information about a manuscript available for peers next to the manuscript. One can of course question the value of this kind of disclosure since it did not go through a peer-review process. The establishment of an editorial board can possibly take away some of these doubts. But even without the validation of peer-review, the information can at least be valuable to spark one’s thoughts. Unfortunately, there

---

162 For example:
P. Monella, ‘Towards a digital model to edit the different paratextuality levels within a textual tradition’, n. pag.
are no examples of this kind of disclosure yet.

The research practice of philology can be influenced by adding information in a specific manner. Adding information to a manuscript can provide philologists with an easier way to achieve their scholarly goal, i.e. taking away interpretational barriers to older texts for current readers. If the information of a philological study is presented next to the manuscript content, the readers will be able to interpret the content more easily. Moreover they would not have to immediately search for and read secondary literature in order to understand the content of the manuscript. Adding information can thus actually help philologists to realise their scholarly goals.

The research practice of codicology will specifically be influenced by the possibility to search for manuscripts based on their material aspects. This will enable codicologists to find appropriate primary sources both faster and easier. However, scholars will have to dedicate some of their time to adding this descriptive information in order to make the manuscripts searchable. This application of adding information can also benefit scholars from other disciplines, even those who focus on the textual content. For example, the material aspects like the lay-out, rubrications and handwriting can indicate what kind of text is presented in a manuscript. This is relevant for the interpretation of the textual content of a manuscript. The textual content and the material aspects of the manuscript are clearly linked and should as such not be separated lightly. This is currently done with full-text versions of a manuscript content, but adding information about the material aspects of a manuscript can reverse this process.

The possible influence of adding information to a digitised manuscript on manuscript research is extensive. However, it is important to note that the practical value of this added information depends on the amount of participation by scholars and the amount and quality of the information they add.

3.1.4 The influence on scholarly communication

A digital information environment can also influence scholarly communication in several ways. The possible influence of social reference management (SRM), fora and publishing scholarly research in a digital information environment will be discussed here.

SRM is a digital tool to share and index bookmarks of digitised secondary sources or references to analogue secondary sources. It can help scholars to find useful secondary sources more easily, because recommendations of peers are an important way to be informed
on valuable contributions to literature. Research has shown that some scholars even expect that such reader ratings will complement peer review in the future. SRM can influence all manuscript research practices in several ways. For one, it would allow scholars to receive recommendations from peers both nearby and at other universities at any time of day. It can thereby help to extent one's network of peers. Secondly, SRM can become a new starting point of one’s search for secondary literature, but it is likely that scholars will still apply footnote chaining too. Thirdly, the way scholars manage their references will also be influenced by an active use of SRM. If a scholar did not manage his references digitally yet, he will be start to do so. Finally, scholars might consider to contribute to the SRM network by adding recommendations, which would add a new task to their research practice.

Discussing scholarly topics in fora would be a new form of informal scholarly communication, which traditionally happened mostly face-to-face at conferences, in ad-hoc discussions or via e-mail or telephone. Informal communication via fora will probably influence the scale and visibility of such communication. The scale is likely to grow, since scholars worldwide can join in the discussions. The visibility changes due to the fact that discussion in a forum can be re-read and stored, something that is commonly not possible with spontaneous face-to-face discussions.

Formal communication, i.e. publishing research, can also be influenced by a digital information environment. For example, a digital information environment presents scholars with the possibility to connect one’s research to the primary source it discusses. Readers will then be able to view the evidence on which the research is based which increases its transparency. This would allow for a networked environment of scholarly work where interpretation, narrative, evidence and commentary are all interconnected. Moreover, it would become easier to get an overview of the research that has already been done and find relevant secondary literature on a manuscript. There are currently no examples of research output that is linked to the primary source it discusses, but there is already an online platform

165 E. Collins and M. Jubb, ‘How do researchers in the humanities use information resources?’, p. 178.
165 S. Mahony, ‘Research communities and open collaboration’, n. pag.
166 R. Procter, et. al., ‘Adoption and use of web 2.0 in scholarly communications’, p. 4049.
where scholars can present their research on digital humanities, called the *Digital Medievalist*. The contributions to this platform for online publication are reviewed by authorities from the field which is a form of peer-review that is commonly not related to online publications. The way scholars can present their research in a digital information environment might also be different from the presentation of research in printed media. This is however a separate discussion which goes beyond the scope of this thesis.

Although these are interesting possibilities, there is an important condition that has to be met in order to make this kind of publishing attractive for scholars. Professional rewards are one of the most important incentives for publishing one's research, but publishing in a digital information environment is currently not rewarded as such. If this does not change, scholars are not likely to publish their research in a digital information environment and thus the aforementioned influences will always remain theoretical.

### 3.2 The likelihood of uptake of a digital information environment

Over the course of this thesis an ideal, yet imaginary, digital information environment has been discussed. Some of the aspects of this platform are currently in use, like downloading and tagging. Other aspects can seem somewhat futuristic, like publishing scholarly research in a digital information environment. This final paragraph will consider how likely it is that the services and tools discussed in this thesis will be used by current scholars and in some cases what the conditions for uptake would be. In general, the uptake of digital research tools and environments depend on their perceived value.\(^{170}\) A new platform will likely be used if scholars think it will make it easier to do their research. The time it takes to learn how to work with a new platform or tool is also an important factor. Scholars are commonly not prepared to invest a lot of time in learning how to work with new research tools, especially if these tools work not as good or just as well as their traditional research tools.\(^{171}\) Both of these factors will be taken into account in the following discussion. The services and tools discussed in this thesis will be evaluated in terms of barriers that can prevent scholars from using them, and attractive features that are likely to promote uptake.

---


\(^{171}\) R. Procter, et. al., ‘Adoption and use of web 2.0 in scholarly communications’, p. 4046.

E. Collins and M. Jubb, ‘How do researchers in the humanities use information resources?’, p. 184.
3.2.1 General factors in the likelihood of uptake

The discoverability of information in a digital platform is an important factor in the likelihood of its uptake. There are currently many separate collections online with their own interface and search system which reduces the discoverability of useful primary sources. Scholars have to search several databases and sometimes even finding the right database can be challenging and time-consuming. Making these, currently separate, collections accessible through one platform is an attractive feature since it can speed up and simplify locating the right sources. Using a digital catalogue to find the right materials should not be a barrier for current scholars since most university libraries already have a digital catalogue which means they are used to working with them.

A related factor in the likelihood of uptake is the amount of search parameters in the search engine of a digital information environment. This should be catered towards scholarly searching and thus go beyond only author names, dates and titles. However, too many search parameters would make the search engine less user-friendly and thus present a barrier. This friction has already been discussed in the previous chapter. A currently achievable solution is providing a simple search to start and more extensive search parameters to narrow down the simple search results.

Trust is also an important factor in the uptake of a digital information environment. A scholar has to trust that the digitisation was done accurately, since any flaws in the digitised manuscript can lead to a bias in his research. If a scholar starts to research a globally dispersed collection, trust is even more important since it will be harder to view the original versions held abroad if need be. Research has shown that distrust of the quality and completeness of a digitised manuscript is indeed a reason for scholars to consult the manuscripts in a library instead of using the digitised manuscript. Moreover, the end-user research by the KB has shown that scholars would like to be informed about the decisions that were made during the digitisation and presentation of the materials. Sharing this information can create trust, which means that it is not necessarily a barrier. However, these

T. Burrows, ‘Applying semantic web technologies to medieval manuscript research’, p. 118.
C. Dallas, ‘Humanistic research, information resources and electronic communication’, p. 228.
175 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
considerations needs to be taken seriously when designing a digital information environment.

A final possible barrier to the uptake of a digital information environment is the coverage of the digitised collection. If a collection is incomplete, scholars will probably rather view the entire collection in a library than part of it in a digital information environment.\footnote{A. Chassanoff, ‘Historians and the use of primary source materials in the digital age’, p. 470.} It is therefore important that scholars know which materials they can access through a digital information environment, so they can make an educated decision about the usefulness of the collection for their research.

The digital tools to study the material aspects of a manuscript are probably easily adopted. Examples of these tools are an electronic ruler and tools to change the contrast, luminosity and colour filters. If these kinds of tools are incorporated in the document viewer and they are easy to use, they will present no barriers to uptake. Concerns of quality can be a barrier to the uptake of full-text versions to study the textual content of a manuscript. A full-text file can be created automatically by using Optical Character Recognition (OCR), manual transcription or a combination of these technologies. Although much progress has been made with OCR, most of these full-text files are still flawed. This means that words within the text are misrecognized by the software and thus presented as different words.\footnote{E. Klijn, ‘Ingezonden: ‘Het digitale drama’, NIOD blog, 15 September 2011 <http://niodbibliotheek.blogspot.nl/2011/09/ingezonden-het-digitale-drama.html> (5 April 2014).} For example, the word 'bold' can be presented as 'bald' since the letter a was misrecognized as an o. This barrier can possibly be circumvented by presenting the original images alongside the full-text files. This will allow scholars to check the transcription and possibly correct any errors to improve the quality of the full-text files. The actual use of full-text files in manuscript research has not been measured yet.

### 3.2.2 The uptake of downloading and software

Downloading a digitised manuscript for research has three general attractive features. For one, a scholar can access his primary source at any time of day instead of having to adhere to the opening hours of a library. Secondly, he can easily access and research manuscripts that are held abroad, which would otherwise be an expensive undertaking. Thirdly, downloadable derivative files, like a full-text file, can make researching a manuscript easier and it can facilitate the use of computer software. A current barrier to downloading a file and using it for research is the quality of the downloadable files, which is not always good enough for
research. Moreover, some platforms offer downloads that are of inferior quality compared to the images displayed online. This barrier can be overcome by allowing scholars to choose between high quality images and lower quality images for downloading. The files containing the high quality images are likely bigger and will thus take more time to be downloaded.

Analysing a downloaded file with appropriate software has two clear attractive features. For one, software can handle larger amounts of data in the same timespan than a scholar can. This allows scholars to process more data which benefits the validity of the research without prolonging the research phase. Secondly, the raw data and output of the software can easily be shared, which makes this research method highly transparent.

Despite these attractive features, the uptake of software for manuscript research is currently quite low. This might be due to the lack of reputation and esteem of software as a research tool. New research practices and tools commonly take some time to be accepted as valid. Since software is still in its infancy as a research tool, it lacks the esteem of more traditional research methods. It will likely take some more time before this changes. A more clear barrier to using software is the time it takes to learn how to work with it. Scholars are commonly not willing to adopt new technologies with a long learning curve unless they clearly see its benefits. Current software for manuscript research is probably still too complex and not clearly beneficial which prevents its uptake. Moreover, the software that manuscript scholars currently use is often not designed specifically for manuscript research. It is therefore probably flawed and imposes more restrictions than software which is specifically designed to analyse manuscripts. It would thus be beneficial if manuscript scholars help to design software specifically catered to analysing manuscripts instead of using generic software. However, this would however require scholars to invest their time in designing software and it is possible that more specialised and sophisticated software will be less user-friendly.

Considering the barriers for uptake, it is probably too soon to expect frequent use of

---

178 For example on the Early Dutch books online platform: <http://www.earlydutchbooksonline.nl/nl/edbo> (5 February 2014).
179 A. Hall, ‘Making stemmas with small samples, and digital approaches to publishing them’, n. pag.
software to analyse manuscripts. The creation of a digital information environment would provide more primary sources that can be researched with software. However, it will not take away the barriers described above and it will therefore most likely not increase the use of software as a tool in manuscript research.

### 3.2.3 The uptake of adding information

The uptake of tools to add information to a digitised manuscript highly depends on the kind of tool that is chosen: tagging or XML. This thesis will therefore proceed to discuss the attractive features of and barriers to uptake of both these technologies.

The most important attractive feature of XML is that it relies on a controlled vocabulary. This makes it easy to search the encoded information, which is especially relevant to scholars who want to search a digitised collection based on the encoded information. Moreover, it makes it easier to exchange encoded files with other platforms so long as they adhere to the same DTD. The biggest barrier to prevent uptake of XML is the long learning curve. Since XML is quite complex, it will probably take scholars a long time to learn how to work with it. This makes it highly unlikely that scholars will easily adopt this technology.

The most attractive feature of social tagging is its user-friendliness. Many websites already use tagging which makes it likely that scholars are already familiar with it, although not in a scholarly setting. Tagging can also be used for personal information management, to help a scholar search and manage his research notes more easily than before.\(^{183}\) Thus, even if scholars are not willing to share their notes, tagging can thus still be a useful feature in a digital information environment. The biggest barrier to prevent uptake of social tagging is the fact that the uncontrolled vocabulary can create too much noise in the search results. The uncontrolled vocabulary also requires that a scholar looks for several synonyms of a search term in order to find all relevant manuscripts. Moreover, the manuscripts that are not enriched with tags will not be listed among the results. Searching a collection based on tags is thus more laborious and not comprehensive until all manuscripts are enriched with tags.

Disregarding the current technological possibilities, there are several attractive features to digitally adding information to a manuscript. Adding information can be beneficial to manuscript scholars since it can facilitate their current research practices. For example, philologists can more easily take away interpretational barriers of older texts; medievalists can profit from the internal and external criticism done by their peers; codicologists can

---

search collections based on the material aspects of the digitised manuscripts; and textual
criticism can benefit from linking aligning texts to compare them. Moreover, scholars do not
have to change their research practice to accommodate this technology since annotation has
always been connected with scholarly practice. 184 Another benefit of digital annotation is that
scholars can easily sort, search and share their notes.

There are however also some general barriers to prevent the uptake of this technology. For one, some scholars are inevitably attached to their own system of annotation and will therefore not switch to digital note taking. Secondly, scholars from different academic
disciplines will focus on different aspects of a manuscript and thus add different information. When all this information is added to one manuscript, it will probably become cluttered and
unusable, which will prevent its uptake. 185 A tool to sort the added information and search it
can possibly help to overcome this barrier. A third barrier is the extra time it takes to add
information, combined with the fact that it is not rewarded so far as tenure or professional
recognition. 186 It will indeed take scholars some extra time to get used to digitally adding
information to a manuscript. However, this practice can be viewed as an alternative way to
annotate a manuscript, which is already a common practice in manuscript research. The time
formerly spent on manual annotation, can now be dedicated to digital annotation and possibly
sharing these annotations. It might thus not be more time consuming than manual annotation
was before. The fact that this practice is not rewarded professionally might be less of a barrier
when scholars see the value of the added information to manuscript scholarship in general.
Moreover, adding information in a scholarly digital environment can generate a good
reputation and awareness of one’s work among peers. 187 This might not directly advance one's
career, but a good reputation can influences how one's research is received in the scholarly
community. These two barriers are thus currently relevant, but possibly not permanent.

The possibility to share one's notes with other scholars has an additional barrier to
prevent uptake, namely the perceived lack of intellectual property rights online. 188 A scholar’s

184 S. Mahony, ‘Research communities and open collaboration’, n. pag.
185 H. van Vliet and E. Hekman, ‘Enhancing user involvement with digital cultural heritage’,
n. pag.
186 C. Dallas, ‘Humanistic research, information resources and electronic communication’, p.
229.
188 M. Greenhalgh, ‘Art history’, S. Schreibman, R. Siemens and J. Unsworth, A companion
C. Borgman, Scholarship in the digital age, p. 223.
C. Dallas, ‘Humanistic research, information resources and electronic communication’, p.231.
research notes are a big part of his intellectual merit. Based on these notes he draws new conclusions and writes his articles. Therefore, scholars might be unwilling to share their notes and risk other scholars using their ideas before they publish them. This is a clear barrier since a claim on research findings is important, also in the humanities.\(^{189}\) There are however already several methods of protecting one’s intellectual property online. Especially the creative commons licenses provide a user-friendly and legally binding option to show if and how online information may be reused.\(^{190}\) However, the creative commons licenses do not seem to cause a big increase of online publication of research output at the moment. Moreover, some scholars do not support such open scholarship and will therefore not share their research notes, even if they are legally protected.

All in all, it seems to be too early for general adoption of sharing research information since there are still quite some barriers to overcome. A digital information environment is not likely to change that. Although the tools are currently available and there are some clear benefits to using them, the adoption will be highly dependent on the ease of use, the perception of its value and the acceptance of creative commons licenses or other ways to legally protect online information.

3.2.4 The uptake of tools for scholarly communication

There are several ways for scholars to communicate in a digital information environment. This thesis has considered social reference management (SRM), discussions in a forum and publishing research. The barriers to uptake and the attractive features of these three technologies will be discussed here.

It seems likely that SRM will be easily adopted when one considers the current success of websites like CiteULike and Bibsonomy. Research has confirmed that scholars using SRM find more useful sources than scholars who do not use SRM, which is a good incentive for adoption.\(^{191}\) Moreover, a network of peers is an important tool in raising awareness of significant contributions to literature.\(^{192}\) SRM provides scholars with a worldwide network of peers which another attractive feature. An important condition for the adoption of SRM is that it should be easy to use. Most scholars already have a system for


\(^{190}\) Creative commons <https://creativecommons.org/> (2 April 2014).

\(^{191}\) H. Alhoori and R. Furuta, ‘Understanding the dynamic scholarly research needs and behavior as applied to social reference management’, p. 170.

\(^{192}\) E. Collins and M. Jubb, ‘How do researchers in the humanities use information resources?’, p. 178.
reference management and if the digital SRM tools are hard to use, these scholars are not likely to turn away from their trusted system. A possible barrier is the fact that some recommendations will come from unidentified peers. The scholars who participated in the KB end-user research indicated that they were only interested in recommendations from peers they know or can look up. It is of course questionable whether this is true for the majority of scholars, especially since the aforementioned SRM platforms are quite popular. Another barrier to adoption is the fact that it is a new technique and a new way of searching for and organising secondary sources. It commonly takes a while before scholars realise the value of new technologies and replace traditional methods with new methods. This barrier will likely be overcome with time.

A digital forum offers scholars a new possibility to communicate about their research discipline. There are currently several forums concerning scholarly topics, for example the aforementioned Digital classicist wiki and the Digital medievalist. The adoption of this kind of fora probably depends mostly on the quality of information gathered there, the amount of active users and the accessibility. By connecting the forum to a digital information environment it will be more easily accessible to scholars who already use other services within the platform. This will, in turn, increase the amount of users. The quality of the forum should be controlled by an editorial staff, to ensure scholars that it will be worthwhile to browse the gathered information and discussions. Since current students have probably participated in non-scholarly fora, it is likely they will easily adopt this technology to communicate with their peers.

The previous and current chapter have already discussed some possibilities for publishing one's research in a digital information environment. This practice has three attractive features for scholars. For one, it allows scholars to publish their research without the restraints and demands of a publishing house or journal. This will speed up the publishing process while still allowing for communication with peers, a claim on research findings and the stimulation of progress and knowledge in society in general, which are three common incentives for publishing. Secondly, the option to link one's research to the primary source it discusses makes the research more transparent since all readers can see what a scholar has observed. Thirdly, scholars can immediately react to the research that is published. A digital

193 Interview with Rosemarie Pomp-Blangé, 21 March 2014.
195 S. Mahony, ‘Research communities and open collaboration’, n. pag.
information environment allows for immediate two-way communication which is not possible in traditional publishing.

However, the adoption of publishing one’s work in a digital information environment still has a lot of barriers to overcome. For one, the perceived lack of protection of intellectual property rights online. Again, creative commons licenses can provide a solution, but the uptake and general acceptance of these licenses is not very high yet. A second barrier specific to humanities, and thus to manuscript research, is the very low uptake of digital monographs, even though this is a common format for research output. A monograph is a lengthy work, commonly written by a single author. Reasons for the scarce use of digital monographs are an aversion to reading lengthy texts from a screen and the lack of availability of digital monographs. The same scholars who do not read monographs online do tend to read articles online, so there is no general aversion to digital research output. This is, at the very least, a barrier to the adoption of publishing monographs in a digital information environment. It seems highly unlikely that a digital information environment will change this. However, the fact that scholars read shorter research output online might encourage scholars to publish these materials in a digital information environment. A third barrier to the uptake of publishing in a digital information environment is the fact that it is currently not academically rewarded. This is one of the most important reasons why scholars publish their research. It is thus highly unlikely that scholars will publish their research in a digital information environment if this does not change. Such a change depends on the universities and it cannot be brought about by launching a digital information environment.

### 3.3 Conclusion

This chapter has discussed the potential influence of a digital information environment on manuscript research practices and the likelihood of uptake of this environment. The potential to influence manuscript research is substantial. A well-designed digital research environment can support new research question, e.g. concerning geographically dispersed collections; new research practices, e.g. by means of software; and new ways of publishing one's research results, e.g. linked to the primary source it discusses.

However, the scholarly community is currently not ready to embrace all these possibilities. Moreover, the digital information environment will have to measure up to high standards. The collection of digitised manuscripts has to be of high quality with good

---

coverage; the institutions responsible for digitisation have to win scholars trust; the search engine has to be fit for scholarly searching; the tools have to be easy to work with; scholars have to be willing to invest time and information; and for the adoption of publishing in a digital information environment it has to be professionally rewarded. Meeting these conditions is essential, since it will determine whether or not scholars will use the digital information environment. When scholars do not use the digital information environment, it will not have any influence on manuscript research. However, if scholars do use it, its influence on manuscript research can be substantial.
Conclusion

This thesis aimed to uncover how current research practices in manuscript research can influence the design of a digital information environment and how this platform can influence scholarly research practices in turn. This conclusion will not recite all the research practices, technologies and examples that were mentioned over the course of this thesis. It will rather take a step back to discuss the relation between research practices and a digital information environment in manuscript research.

When talking about a new research tool like a digital information environment, it is important to keep in mind that scholars will not use new research tools if they do not see the added value it brings to their current research practices.\(^{200}\) They will not adopt new research tools simply because they are available. Manuscript scholars have always done research without a digital information environment and if the platform does not function to their satisfaction, they will most likely favour their traditional research practice.

It is also important to note that adoption of a digital information environment will be much easier if the platform aligns with current manuscript research practices, as discussed in the second chapter of this thesis. These practices have been shaped by years of research and they have a certain esteem. It is therefore unlikely that scholars will easily abandon these practices in favour of new practices in a digital information environment. The third chapter of this thesis indeed confirmed that tools which align with current practices are more likely to be adopted. In order for the digital information environment to be adopted, it should thus support current research practices and not aim to force new practices on scholars. But even if the platform aligns with current research practices, it will probably still take some time before scholars completely adopt it.

Once scholars start to use a digital information environment it will probably change their research practices, as discussed in the third chapter of this thesis. A digital information environment allows scholars to ask different questions, use different tools to access and view their primary sources and allows for a new way to communicate and share information. It might sound contradictory that a digital information environment has to adhere to current research practices first, to eventually change them anyhow, but this is more logical than it may seem. A digital information environment can only influence manuscript research if it is commonly used. In order to be commonly used, it has to align with current research practices,

\(^{200}\) R. Procter, et. al., ‘Adoption and use of web 2.0 in scholarly communications’, p. 4046, 4050.
for reasons mentioned before. Once it is in use, it offers scholars new possibilities that were
impossible in a non-digital research environment. If scholars start to make use of these new
possibilities, it will change their research practice. This will however be a natural evolution
due to new possibilities, not a forced change as a condition to use the digital information
environment.

Considering the discussion above, this thesis fits in with a discussion that has been
held over the past decades on technological and social determinism. Technological
determinism is most commonly defined as a theory which consist of two ideas: 1) ‘that
technological developments take place outside society, independently of social, economic,
and political forces’, 2) ‘that technological change causes or determines social change’. 201
Social determinism argues the exact opposite, namely that ‘wider cultural and social structural
patterns of specific societies’ determine how technologies are used. 202 These theories are also
applicable to the relation between research practices and a digital information environment in
manuscript research. Technological determinism would suggest that a digital information
environment can be developed outside the research environment and that it will change
manuscript research practices. Considering what is discussed in this conclusion so far, this
thesis does not support this reasoning. A digital information environment has to be designed
with current research practices in mind, since both theory and practice have shown that the
platform will otherwise not be used. Subsequently, a platform which is not commonly used
will not cause or determine any change. Social determinism would argue that the current
research practices of manuscript scholars will determine how a digital information
environment will be used. This thesis supports this reasoning, since scholars will probably not
easily abandon their traditional research practices just so they can use a new research tool.
Therefore it should be the other way around: the current research practices will determine
how the platform will be used, and ideally these practices should determine the design of the
platform.

One final note has to be made. The development of a digital information environment
should not stop after its first launch. If the research practices of manuscript scholars change,
the platform will have to develop alongside these changes. Otherwise it will fall into disuse as

201 R.C. Scharff and V. Dusek, *Philosophy of technology: The technological condition - An
202 M.R. Jones, ‘Information systems and the double mangle: Steering a course between the
scylla of embedded structure and the charybdis of strong symmetry’, T.J. Larsen, L. Levine,
a tool that was once useful, but is now outdated. As such, a digital information environment and manuscript research practices will continue to influence each other and will ideally develop together. The first step in this cycle was discussed in this thesis, namely creating a useful digital information environment that will likely be adopted by current manuscript scholars and how this platform can influence their current research practices. The following developments will depend on the actual uptake of the platform and the ways in which it will influence manuscript research practices.
Bibliography

Primary sources

CiteULike <http://www.citeulike.org/> (7 February 2014).


Early European books online <http://eeb.chadwyck.co.uk.access.authkb.kb.nl/info/demo.do> (4 February 2014).


The dynamics of the medieval manuscript <http://everycodextellsastory.eu/> (3 January 2014).


Twitter @BLMedieval <https://twitter.com/BLMedieval> (3 January 2014).

Twitter @erik_kwakkel <https://twitter.com/erik_kwakkel> (3 January 2014).

Dynamics of the medieval manuscript <http://dynamicsofthemedevalmanuscript.eu/> (3 January 2014).

Secondary sources


McGann, J., A critique of modern textual criticism (Chicago etc.: University of Chicago Press, 1983).


Shafer, R.J., A guide to historical method (Homewood, etc.: Dorsey Press, 1974).


**Websites**

Creative commons <http://creativecommons.org/> (28 April 2014).

DBNL <http://www.dbnl.org/tekst/que002200701_01/> (3 January 2014).


Metapress <http://www.metapress.com/content/p37611120092/> (3 January 2014).

Turning over a new leaf <http://www.hum.leiden.edu/lucas/turning-over-a-new-leaf/project-manuscript-innovation/manuscript-innovation.html> (3 January 2014).

YouTube <https://www.youtube.com/watch?v=tkdht0LjczA> (23 March 2014).
Appendix I: Illustrations

Figure 1: Adding notes using Diigo. <http://idoe29dayweb20challenge.blogspot.nl/2012/02/day-22-diigo-social-bookmarking-and.html> (1 April 2014).

Figure 2: The extended search options in the Early Dutch books online platform. <http://www.earlydutchbooksonline.nl/nl/edbo> (1 April 2014).

QR Codes in the Classroom

For the classroom teacher, they are valuable for three reasons:
1. They can save us time;
2. They can save paper;
3. They provide a link to mobile devices that help students do their homework and follow along.

Back in March, we discussed QR codes in our Lesson in Simple Augmented Reality. We used them extensively in our eighth grade portfolios this year and it has cut assessment time in half. Many of you have asked that this post be written to help you on your journey.

QR Codes

The first step of a teaching journey is to embark on learning it yourself.
Figure 3: The extended search options in the *Early European books online* platform. <http://eeb.chadwyck.co.uk.access.authkb.kb.nl/info/demo.do> (1 April 2014).
Figure 4: The blue arrow indicates the simple search feature on the homepage of the platform *Delpher*. The orange arrow indicates the means to narrow down the simple search results. &lt;http://www.delpher.nl/&gt; (1 April 2014).

Figure 5: The image on the left shows the quality of the image displayed on the website of the *Early Dutch books online* project. The image on the right shows the quality of the image that is available for downloading. &lt;http://www.earlydutchbooksonline.nl/nl/edbo&gt; (1 April 2014).
Figure 6: Encoding with XML. <http://www.w3schools.com/xml/xml_whatis.asp> (1 April 2014).

Figure 7: A list of tags added to a digital image of an item in the collection of The powerhouse museum. <http://www.powerhousemuseum.com/> (1 April 2014).
Deborah Hall
The portrait of sixteen-year-old Deborah Hall demonstrates the rich and multifaceted language of symbols at the disposal of portraitists in British America. William Williams portrayed his young subject in a fictional, carefully designed landscape standing alongside a relief sculpture of Apollo and Daphne, who escaped the god’s unwelcome advances by turning into a laurel

Figure 8: A list of tags added to a digital image of an item in the collection of The Brooklyn Museum. <http://www.brooklynmuseum.org/home.php> (1 April 2014).

Figure 9: An example of a ‘tag cloud’ that displays common tags. The size of the words indicates how often a tag is used. <http://www.powerhousemuseum.com/> (1 April 2014).
Figure 10: A display of search results in CiteULike. The orange arrow indicates the screenname of the person who added this article to CiteULike. The green arrow indicates the tags that have been added to this article. <http://www.citeulike.org/> (1 April 2014).
Figure 11: A display of search results in Bibsonomy. The orange arrow indicates the screenname of the person who added this article to Bibsonomy. The green arrow indicates the tags that have been added to this article. <http://www.bibsonomy.org/> (1 April 2014).