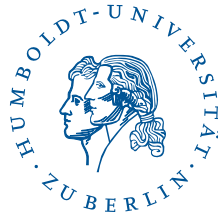


HUMBOLDT UNIVERSITY OF BERLIN

INSTITUTE FOR LIBRARY AND INFORMATION SCIENCE



SEEKING RESEARCH SOFTWARE. A QUALITATIVE STUDY OF
HUMANITIES SCHOLARS' INFORMATION PRACTICES.

MASTER'S THESIS IN THE POSTGRADUATE MASTER'S PROGRAM IN LIBRARY
AND INFORMATION SCIENCE BY DISTANCE LEARNING

RONNY GEY

SUPERVISORS:

PROF. REBECCA D. FRANK, PH.D.
PROF. DR. WOLFRAM HORSTMANN

LEIPZIG, NOVEMBER 10, 2020

Abstract

Within their own research process, researchers use a variety of software to collect, transform, analyze and present research data and to plan these steps. Seeking and using software that meets these needs has become an important aspect of researchers' scientific practice. This qualitative study examines the information practices of classical philologists in seeking research software. Nine interviews were conducted with young scholars of an institute of classical philology, which were subsequently evaluated using qualitative content analysis. The philologists studied seldom seek for software and prefer their own colleagues as a source of information in the active seeking process. Within the group of colleagues, the information needs of others are also taken into account; the colleagues often act as a proxy. Active seeking via information systems is often terminated early. Friends are another important source of information for the information practices of philologists. A very high value of trust in information sources and negative experiences with digital humanities have a significant influence on the information practices found. The results serve as a starting point for scientific libraries to show how researchers can be supported in their search for research software by an extended range of services.

Contents

List of Tables

List of Acronyms

APh l'Année philologique.

BMCR Bryn Mawr Classical Review.

CIS collaborative information seeking.

DFG Deutsche Forschungsgemeinschaft.

DH digital humanities.

ELIS everyday life information seeking.

LIS Library and Information Science.

LLTA Library of Latin Texts.

OSS open source software.

RSE Research Software Engineers.

TLG Thesaurus Linguae Graecae.

1 Introduction

Today, software is a central component of science. Throughout the entire research life cycle, researchers use software tools for data collection, transformation, analysis and presentation as well as for generating hypotheses, managing literature and writing scientific papers (??). Software has changed the way we actually do science. The complexity of the analyses carried out by researchers has increased, as has the amount of data that researchers can process. Software supports the documentation of the research process and enables reproducibility (??) and accuracy of results.

Due to the increased importance of research software for research (?) and the increase in the sheer number of software, it is all the more important for researchers to identify suitable software and select the one which best fits the research problem, the actual step in the research process, or the research data which has to be processed, and, in consequence, which satisfies the researchers information need (?). In addition to increased efforts, difficulties in seeking software can also endanger the scientific reproducibility of studies or even lead to multiple developments of software with identical functions instead of reusing existing software (?).

Information seeking of researchers is generally of great interest within the field of information, be it information behavior (e.g. ??????????) or information practices (e.g. ?????). However, seeking software is still rather challenging for researchers (?). In a recent study, ? surveyed scientists and engineers from several fields to better understand their approaches and selection criteria for seeking software. They found out that *"finding software suitable for a given purpose remains surprisingly difficult"*. Even in the domain of software development, the main challenge for software reuse are difficulties in finding software artifacts (??). These results are all the more surprising because the participants in the cited studies belong to a group with a greater affinity for software (software developers, engineers).

The information-seeking behavior of humanities scholars in general has been addressed widely (e.g. ?????????). While studies in information behavior draw on the cognitive viewpoint, information practice studies are guided by the ideas of social constructionism and collectivism (???). Humanities scholars information-seeking practices have also been addressed in several studies (????). In previous studies, however, the classic research process of humanities scholars has been examined, which is mainly based on literature research. Although the information-seeking in the humanities is also based on software tools, e.g. databases, web-based editions, search engines, or online journals (??), the search for software itself is rarely discussed.

In this thesis I want to investigate the information practices of humanities scholars in seeking research software. For this purpose, I interviewed classical philologists of a German university institute of classical philology in a qualitative study and

asked about their information practices. Data were transcribed and analyzed with the qualitative content analysis according to ?. In preparation of the study and to derive the research gap, I will present the theoretical concepts of this work in the next chapter. I discuss the concepts information seeking (??), information practices (??), research software (??) and the field of humanities (??). Afterwards I present the research design of the thesis in chapter ??. The subsequent presentation of the results is divided into two parts. I will describe the information practices of the subjects (section ??) and then interpret some of these descriptions in section ?? on the basis of the subjects' experiences in classical philology. In the following discussion (chapter ??), I will integrate the findings from the empirical study into the current theory. The conclusion (chapter ??) summarizes the essential aspects of the thesis and limitations. Ideas for further research are proposed.

2 Theory

2.1 Information Seeking

2.1.1 Introduction

Professional, personal, social and societal information problems face people with the task of searching for information and creating access to various information resources. There is a multitude of easily accessible information. In their daily work, people are interested in finding and using this information. This is accompanied by an interest in understanding how we collect, create and organize information, but also how we search this information. Library and Information Science (LIS) is particularly interested in how people proceed, what motives guide them and what challenges they encounter. Consequently, information seeking has been the subject of intensive research for many decades (?, p. 55ff).

Within information science, we distinguish between seeking, searching and retrieving information. The latter two are closely related to information seeking. To avoid confusion, the differences and the relationship between the concepts will be briefly explained before I continue the considerations on information seeking. Information seeking is the "*[h]uman information behavior dealing with searching or seeking information by means of information sources and (interactive) information retrieval systems*" (?, p. 21). Accordingly, information seeking is a sub-set of information behavior, the one which has been studied most often in information behavior research (?). Information searching is again a sub-set of information seeking (?, p. 263). The often synonymous meaning of both words makes it difficult to distinguish between seeking and searching. Information searching deals with the interactions between users and the often computer-based information system (?, p. 261). According to ?, information retrieval in turn deals with "*the purposeful searching for information in a system, of whatever kind, in which information - whether in the form of documents, or their surrogates, or factual material ('information itself'), are stored and represented*". The difference between information searching and information retrieval lies in the kind of information system a user applies for his information-related activities and its directionality. In retrieval, a user uses a specific retrieval system and purposefully searches within this system. However, in information searching a user can interact with any information system while it does not have to be necessarily designed for this purpose. Hence, information retrieval can be considered as a sub-set of information searching (?).

2.1.2 Information Seeking Research

Theoretical and methodological approaches in information seeking depend strongly on the scientific-theoretical orientation of the researchers themselves and their basic epistemological convictions regarding access to the world and how knowledge is created. Within information research, we distinguish between the dominant umbrella concept of information *behavior* and the critical alternative of information *practices*. Both concepts emerge from different discourses which offer alternative viewpoints on information seeking (?). I will describe the concept of information behavior in section ???. However, in the further course of the work and for the conception of the empirical study I will use the concept of information practices, which I will present in section ???.

For this reason, I will not discuss methodological or theoretical aspects of information seeking research here. These aspects will be analyzed in the following two sections with respect to the approaches of information behavior and information practices. Questions regarding the demographic or occupational role of participants in information seeking studies are also dealt with in the following two sections. At this point I will concentrate on different types or strategies of information seeking which can be described independently from the adopted approach (behavior or practices).

In his early work, ? noted that people also engage in non-work related information seeking. The everyday life of many people and groups of actors thus became the focus of LIS researchers. ? then coined the term everyday life information seeking (ELIS) and defined it as: "*[...] refers to the acquisition of various informational (both cognitive and expressive) elements which people employ to orient themselves in daily life or to solve problems not directly connected with the performance of occupational tasks.*". Since then, many attempts have been made to research the ELIS of different groups and under different circumstances (e.g. ?????). In contrast to previous studies, the ELIS studies have managed to emphasize the legitimate nature of the nonwork context (?, p. 266). In addition, studying ELIS is accompanied by the recognition that non-active information seeking is important because active information seeking does not account for all of information behavior (?, p. 21).

Besides ELIS, collaborative information seeking (CIS) is another prominent concept of information seeking (e.g. ??????). Collaborating means to cooperate, coordinate, contribute and communicate (?, p. 6). Working together is often necessary to perform activities that are too complex or too difficult for a single person. Many information-seeking situations require people to cooperate. This presupposes the description and design of methods, systems and tools for accessing information. ?, p. 26 points out that user collaboration can occur at different points in the information seeking process: "*(1) while formulating an information request, (2) while obtaining results, and (3) while organizing and using the results.*" He suggests that all three aspects should be considered when designing collaboration and the necessary technical support facilities. With the advent of CIS studies, attention in the field of information seeking shifted away from viewing the individual, the cognitive system, as an independent unit of information seeking, towards a recognition of the information

seeker as a social and collaborative being.

Furthermore, researchers have conducted important studies on the strategies of information seeking (e.g. [10]). By seeking strategy [11] refers to the plan for an entire search, while associating tactics with short-term goals and actions. There are seeking strategies of communication, avoidance, counter-strategies to certain information barriers, but also very concrete granular strategies, such as scanning, searching, learning, selecting, recognizing, specifying, information and meta-information ([12]). The investigation of seeking strategies allows to model and implement situationally appropriate tools and seeking strategies, thus relieving users both cognitively and temporally ([13]).

Another important aspect of research on information seeking is the differentiation of information seeking from other activities within the information behaviors. For example, incidental information acquisition can play a role in the detection of information ([14]). [15] distinguishes information encountering from purposive information seeking, whereby the author emphasizes the distinction between active and non-active seeking, similar to the ELIS studies. And again [16] try to integrate the process of the information seeker and communication, whereby not only the information seeker, but also the communicator or information provider is considered.

2.1.3 Information Behavior

Information behavior describes the ways of human interaction with information, specifically seeking and using information ([17], p. 2381) but also managing and giving information as well as their information needs ([18], p. 1). In the literature there are many definitions of information behavior. In my further explanations I will refer to the understanding of [19], p. 49, who defines information behavior as *"the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking and information use."* The interest of science in information behavior goes back to several information-relevant social needs: librarians wanted to better understand library users, governments wanted to understand how technical and social information is used to promote the implementation of new research results, and social scientists were interested in the effects of information use on society and its actors ([19], p. 2381).

Models represent the designer's ideas about how a certain phenomenon can be viewed at a certain time and from a certain perspective ([19], p. 142). In his prominent work on models of information behavior [20] describes the most common models of information behavior at that time. Among them there is Dervin's Sense-Making framework ([21]), [22] model on information seeking behavior, [23]'s ([24]) stage process model and both of Wilson's models. The model of [25] is a major revision of his first model ([26]). To date, a wide variety of models of information behavior with different orientations have been developed (e.g. [27]).

Just as the number of different models of information behavior has become unmanageable, so it is with the theories applied to information behavior. [28], p. 149 defines theories of information behavior as *"explanation of observed information*

behavior supported by a systematically obtained body of evidence". These theories are developed within (e.g. [10]) or outside (e.g. [11]) of information science and show different degrees of abstraction. [12] discusses different theories of information behavior and [13] analyzes the use of theories of information behavior across different disciplines.

[14], p. 50 considers information seeking, avoidance, evaluation, encountering, acquisition and use as typical types of behaviors in information behavior research. However, information seeking dominates among the studies of information behavior ([15]). The term information seeking behavior is sometimes even used synonymously for information behavior (e.g. [16]). Various factors influence the information behavior of individuals, both external (for example, roles, status, or social relationships) and internal (for example, cognitive, demographic, such as gender or age, or affective) factors ([17], pp. 99ff). The study of these factors makes it possible to design information systems that are better tailored to the needs of users, to provide more effective training and education in information behavior, and to enable information professionals to solve the problems of their clients. Studies of information behavior show a very high target group dispersion, which is mainly due to the sheer quantity of studies. There is probably hardly any occupational or demographic group that has not yet been investigated: students and scientists (e.g. [18]), nurses ([19]), transgender ([20]), patients ([21]), high school students ([22]), refugees ([23]), teachers ([24]), families ([25]), athletes ([26]), or the elderly ([27]) just to name a few (see [14], p. 277ff, for an overview of different groups or roles in information behavior research).

As mentioned earlier, information behavior is the dominant concept in information research. However, the concept of information behavior was repeatedly questioned and criticized. According to [14], p. 238 the concept still lacks "*of robust and applicable models and theories*" despite or because of the sheer amount of models and theories. Furthermore, he postulates the need for applicable and practical research that helps librarians or system developers to improve information systems or information delivery. Very direct and especially relevant for this thesis, the researchers of information practices criticize information behavior research ([14], p. 116ff) and propose an alternative concept based on these critics: information practices. In their eyes, studies of information behavior take a cognitive standpoint, while social, discursive, and interactive perspectives are not covered. It is criticized that participants in studies of information behavior with their ideas and motives are considered isolated from the social discourses and groups in which they are involved.

I agree with the criticism of the concept of information behavior and try to apply the approach of information practices to the information seeking of humanities scholars in need for research software and thus to consider context and discourse in my analysis. In the next section, I therefore present the approach of information practices as the research guiding theoretical approach of this thesis.

2.2 Information Practices

2.2.1 Introduction

As a critical alternative to the cognitive-oriented studies in information behavior, the approach of information practices relies on a social constructionist approach (?). According to ?, p. 102, the concept of "practices" can be traced back to the work of ? and his praxiological approach. With his idea of "praxis" Bourdieu criticized both, objectivism as well as the opposing subjectivism. Through the actions of a subject in a world, the subject builds this same world in a "praxiological" way (?, p. 102). ?, p. 120, on the other hand, sees the theoretical roots of the information practice approach in the work of Anthony ? on structuring theory. Similar to Bourdieu, his theory emphasizes the dialectic between structure and action. The actor is seen as a knowledgeable agent and the action is the starting point for further considerations. Giddens's structuration theory and Bourdieu's praxiological approach are based on the same idea of placing action at the center of the considerations. It is thus understandable that both serve as a theoretical foundation for information practices.

In the information practice approach, research is sociologically oriented and focuses on the context of actions (?). ? summarizes that the information practices approach: *"recognizes the social dimension of disciplines as a primary influence on the information activities of scholars and scientists"*. Instead of motives and ideas of individuals as foundations of action, processes of information seeking are socially and dialogically constituted (?, p. 120). Rather than focusing on individuals' behavior, actions, skills, ideas and motives, practices regard the interactions between members of a community (?). As ?, p. 93 points out, the theoretical approach of information practices is used, among other things, in integrated studies of knowledge organization in various domains, in information seeking research as well as in the analysis of professional or scientific discourses in the field of information systems or technology.

The following overview of studies from the field of information practices refers exclusively to studies that are related to the concept of information practices (?) or to the constructionist approaches (?) linked to it. In the literature, many studies on information practices in general are found without any reference to the theoretical concept. According to ?, p. 126, the reason for the problematic delimitation of the term practices is that it is *"subject to multiple meanings in the vocabularies of philosophy, psychology, and sociology as well as in information-seeking studies"*. Further, the concepts lack of scrutiny regarding philosophical, psychological or sociological findings which often leads an understanding of information practices in an *"unreflective fashion"* (?, p. 126).

2.2.2 Studies of Information Practices

In this section, I will first present the work of Pamela McKenzie as a prominent proponent of information practices. Then, the research on information practices as a

whole is examined and presented according to various criteria.

According to ?, p. 121, with her constructionist approach to information practices, Pamela ??? is one of the most influential scholars of information practice. In her empirical studies, she focuses on a constructionist discourse analytic approach in practitioner-patient information seeking. ? analyses communication barriers and active seeking and scanning practices as counterstrategies in response. The patients seeking information presented themselves as active and alert. They deliberately asked questions and prepared lists to organize their behavior. They also looked for ways to intervene and receive feedback when a process was unexpected. The reports also revealed that the information seekers acted as indispensable partners without whose active participation practitioners could not work effectively.

In a study on cognitive authority decisions, (?) analyzed context-specific discursive techniques used by participants to strengthen or weaken the authority of peer and professional information sources. The discussion participants described cognitive authority in a very flexible and versatile way. Depending on the needs of interaction, they alternated fluently between claims to authority based on common, moral and those based on subjective criteria. The constructionist discourse-analytic methods used by McKenzie were particularly suitable for identifying the specific strategies used by the participants in creating their cognitive descriptions of authority.

Based on the experience gained during studies of everyday life information seeking in practitioner-patient settings, ? proposed an empirical model of everyday life information seeking. The model was a response to the then emerging knowledge about non-active information seeking. It contributes to broaden the analysis of information behavior by considering not only active but also non-active engagement in information seeking. Environmental scanning, chance encounters, lay referrals and connecting by proxy were the dominant non-active information practices in her study.

In the following paragraphs, I will analyze studies on information practices according to various criteria. In the first part, I will present the different demographic and occupational groups of actors that have been examined in studies on information practices. After that, I will focus on the different information practices used by the actors under study. Subsequently, I will present the methodological aspects of the studies with regard to data collection and analysis.

Since information practices are bound to concrete actors, I will start with the different occupational or demographic groups of actors described in the empirical studies. As in many other information studies, studies on scientists and students also predominate studies in information practices. ? in their study on use and non-use of e-journals and databases, for example, examined actors from nursing science, literature/cultural studies, history and ecological environmental science. And also ? in her study on information needs, seeking, searching, and uses within scholarly communities compared different sciences: physical sciences, applied sciences, social sciences and arts and humanities. ? again examined medical scientists during their research work. ? wiederum thought that it is very obvious to simply examine one's own group of scientists and examined information scientists in the discursive

construction of an author. (?) and (?), on the other hand, have put students in the focus of their work.

Nurses are another rather common professional group in studies on information practices (e.g. ???). In their discourse-oriented information seeking study, ? analyse nurses and the Swedish nursing profession. In further studies, ?? has in turn examined theater professionals and field archaeologists, among others. He interviewed 35 theater professionals from Canada, Finland, and the UK and 10 archaeologists and archaeology students while he also observed the group of field archaeologists. (?) conducted a three-year qualitative study of Australian firefighters and analyzed information literacy that was constituted through discursive practices. A vault inspector and its inspection practices were examined in the study of (?).

In addition to *professional* groups, there are other groups that appear in studies on information practices. Mainly, they can be classified according to the personal life situation of the participants of the studies. ?? for example, examined the health information practices of young Canadian parents between the ages of 16 and 23 in his studies. They were faced with a challenging assessment of a large amount of health-related information. The pregnant women studied by ??? have a similar need for health and situation related information. The immigrants studied by ??? exhibit a completely different life situation. In the basic tenor of the studies, for example, it emerges that the social exclusion experienced by immigrants is also an information problem caused by barriers in an unknown information environment. The refugees examined in the study by ? find themselves in a very similar situation. The results show that the refugees find complex and challenging information landscapes in their new communities, which present barriers to their full integration.

Another vulnerable group, the LGBTQ+ community, was investigated by ?. She studied qualitatively the South Carolina LGBTQ+ community and developed a conceptual framework where she describes protective and defensive types of information practices in reaction to experienced risks and barriers. In a grounded theory study, ? analyzed the interactions of queer individuals with entertainment media and described their identity-related information practices and their experiences to access queer-representative content in information institutions. Harlan's studies (???) on young authors of content focused on the information experiences of young people in specific social and digital contexts. The young people generally discussed information use and seeking and the process of information use to participate in communities.

In the literature there are now many different information practices, which vary according to context and situation. ? for instance described sharing, mapping, observing and listening as information literacy practices. In their studies on health information practices of LGBTQ+ communities, ? examined protective and defensive practices in relation to both the self and the community. The workplace information practices study by ? describes judgment, embodiment, educated perception, finding and navigating, and classification as information practices of a blue-collar worker. As a result of the grounded theory analysis by ?, learning community, negotiation aesthetics, negotiation control, negotiation skills, and the representation of knowledge as information practices emerge.

A very elaborate presentation of ELIS information practices is provided by ???. She divides the information practices of women pregnant with twins into four categories: active seeking, active scanning, non-directed monitoring, and by proxy. Active seeking practices are list making, actively asking questions, keeping the process on track, and persistence. Active scanning practices are opportune questioning, observation and listening, staying connected, and process monitoring. As a result of her work ? proposes an ELIS model based on a constructionist discourse analysis.

In the last two decades, studies of information practices have used a variety of theories to answer their research questions. One of the most prominent theories used for information practices is domain analysis (???). Domains are regarded as communities of thought or discourse that are part of a social division of labor. As such, they differ in knowledge organization, structure, patterns of cooperation, forms of language and communication, information systems used, and relevance criteria (?). Using domain analysis, the effect of factors such as domain size, degree of scatter in domain, or domain-specific relevance factors was investigated (?). Studies have remarkably frequently examined scientific communities as knowledge domains (?????), whereby it comes to the fore that scholars of information science in their studies often simply remain in their own domain, the science.

??'s Sense-Making theory has also been widely used in studies of information practices (e.g. ?????). Olsson can be regarded as one of the most prominent sense-making scholars (???). Dervin's model is particularly strong in its methodological consequences. In terms of both, information behavior and practices, it can lead to investigations that can unravel otherwise problematic situations and bridge the gap of uncertainty. It is hence considered as an instrument at the micro-level of analysis (?, p. 253). Further theoretical concepts have also been mentioned in studies of information practices. For example, ? or ? use the social practice theory (?). In contrast, ? or ? use activity theory to describe reality.

In the next paragraphs, I will present the methods applied in studies on information practices. In these studies, mainly qualitative studies were found, which focus on developing hypotheses rather than testing. This is not particularly surprising, considering the theoretical roots of the information practice approach, the social constructionism. ? points out that constructionism is clearly linked to the post-modern era of qualitative research. Hence, studies of information practices are often qualitatively oriented. The presentation of the research methods of information practices studies is distinguished into methods for data collection and data analysis.

Interviews are by far the most common form of data collection in information practices studies (????????????????????????????). The interview is the most frequently chosen data collection technique in qualitative social research. Since, as mentioned above, mainly qualitative studies of information practices were conducted, the frequent occurrence of interviews among the data collection techniques of information practices reflects the frequent occurrence of interviews in qualitative research in general. Several studies, however, drew on observations in her studies and very often combined them with interviews (???????). The only exception is ?, who observed various care simulation sessions and made additional audio recordings for later

analysis. Although studies of information practices rely primarily on qualitative methods, there are a few questionnaire surveys (e.g. ??) that are usually applied in combination with other data collection methods.

Among the data analysis methods used, studies using a discourse analytical approach (???????), a grounded theory approach (?????????) or approaches derived from the methods of grounded theory (????) predominate. Qualitative social studies often make use of the grounded theory methodology (?), so their frequent use in studies of information practices should not be surprising either. In terms of discourse analysis (?), ? have paved the way for subsequent studies of information practices with their work on information use as discursive action. In 2007, discursive approaches even received their own special issue in *The Library Quarterly* (?), which was intended to reflect the significance of discourse-theoretical approaches for the information sciences. In combination with grounded theory ? used content analysis to analyze the information practices of queer individuals. In their paper on renal nurses ? analyzed the information practices that nurses used to support everyday renal care practice. They applied a thematic analysis (?) to analyze interview data. In addition, there were some studies that gave no or very little information on how they proceeded with data analysis (e.g. ????).

The information practices examined in this study were derived from the accounts of seeking research software. Accordingly, the next section discusses theoretical considerations about research software in general and in relation to the concepts of information seeking and information practices.

2.3 Research Software

Research software is a central component of science. Throughout the entire research life cycle, researchers use software tools for data collection, transformation, analysis and presentation as well as for generating hypotheses, managing literature and writing scientific papers (???). Following the understanding of the Deutsche Forschungsgemeinschaft (DFG) (?) and the Helmholtz Association (?) I define research software as:

Software that researchers use during the research process to collect, transform, analyze and present research data and to plan these steps.

In formulating the definition, it was important to me to derive the distinction between research software and software in general through the reference to the researcher and the research process. Thus, even software tools that are perceived as trivial at first glance can be understood as research software because they are used in research activities. The term research software cannot be defined objectively and universally valid, but only the concrete application context makes a software a research software. What is considered as research software has to be differentiated according to the situation.

The lack of awareness of specific software tools among researchers has been addressed by several technical solutions. Code aggregators, specialized search engines, programming language package repositories, code and application repositories, research repositories (e.g. Zenodo or Figshare), and curated web lists and catalogues aid users in discovering software (?). Standards and tools for citing software enable researchers to identify, cite and reuse software (e.g. ???). Research funding agencies and research organizations (e.g. ???) adopt guidelines for the development and use of research software with the aim of increasing the reusability and quality of the software artifacts developed. In turn, the technical solutions presented are also aimed more at a technically experienced audience, often even at software developers directly. Despite these technical solutions, for researchers with less experience in the use of software, e.g. from the humanities (?), seeking software remains a difficult undertaking.

The topic of research software has gained increasing attention in recent years with a variety of perspectives from which research software was studied. One such perspective is the citation of research software (e.g. ???????). An important aspect for the findability and reusability of research software is to enable other researchers to cite the software. In order to consider software as an independent type of publication and to enable other researchers to cite it, for example, citation and metadata standards for software are needed. Closely related to citation is the discovery of research software (e.g. ???). Researchers with this focus ask how researchers discover new software, which systems they use for this purpose and, above all, how these systems must be designed to improve the process of software discovery.

The people behind research software, the Research Software Engineers (RSE), are also being increasingly studied in the recent years (e.g. ??????). Questions of interest here are the training and role of RSE's, their development paths and recognition in science, as well as networking and exchange among and within the scientific community. Closely related to the perspectives on research software mentioned so far is the concept of sustainability, which considers research software in the larger context of a sustainable and open science (e.g. ?????????). Here, questions about the sustainability of research software are asked from various perspectives (citation, discovery, open source, human resource, reusability, etc.). The topic open source software (OSS) is closely related to questions of sustainability of research software. For instance, various authors (e.g. ???????) have postulated that research software should be published under an open source license to meet the requirements of open science¹.

Furthermore, research software as a subtype of software is increasingly studied from the perspective of software engineering (e.g. ?????). In principle, all aspects of the discipline of software engineering or software-related computer science can be applied to research software (e.g. modeling, architecture, performance, testing, etc.). Finally, authors from a wide range of disciplines have approached the topic:

¹An opinion I fully agree with, which is why I used only OSS when creating the thesis.

e.g. astronomy (????????), biology (????), chemistry (????), humanities (????), medicine (???), and physics (???).

For the progress of the work, studies from the field of information seeking are of interest, preferably those dealing with information practices around research software. To date, I am not aware of any studies however that deal with the information practices of researchers while seeking software. In terms of seeking in general, ? found out, that seeking software is still rather challenging for researchers. ? surveyed scientists and engineers from several fields to better understand their approaches and selection criteria for seeking software. They found out that *"even among highly-trained people, the rudimentary approaches of relying on general web searches, the opinions of colleagues, and the literature were still the most commonly used"* approaches. Even in the domain of software development, the main challenge for software reuse are difficulties in finding software artifacts as ? revealed in a study on code reuse at Google. ? identified users unawareness of specific software tools.

The results of the cited studies indicate that researchers have significant problems in finding software. The authors can name some reasons for this, such as users unawareness of the existence of specific tools or, as (?, p. 171) call it, the recourse to *"rudimentary approaches"*. Based on little qualitative evidence, I want to investigate the problem of seeking research software with an approach that is open and broad and which investigates the context of seeking. I have chosen the approach of information practices over the project information behavior since it considers individuals as social actors who are involved in a larger social context such as groups or communities (?, p. 120). The information practices within such a group can help to answer questions about the reasons for the challenges, the unawareness or the rudimentary approaches to information seeking for software. The scholars I examined all belong to an institute for classical philology at a German university. As an independent community, classical philology shares many characteristics with other humanities. In the absence of studies on classical philology in particular, I will discuss information behavior and practices in the humanities in the next section.

2.4 Information Seeking and Digital Tools in the Humanities

The information seeking of humanities scholars has already been extensively investigated (e.g. ?????????). However, the high amount of studies should be regarded cautiously in view of the large number of different disciplines within the humanities. Already in the 1980s ? analyzed the literature on how humanities scholars work and what information needs and uses they exhibit. She focused on used materials, used technology, the role of libraries, the services they provide, and comparisons with other disciplines. Even then, she did not even attempt to define the concept of the humanities itself: *"definitions of the humanities are elusive"* (?, p. 293). Therefore, studies on information seeking should not lead to a generalization about

the humanities, but should be considered with regard to the respective sub-discipline within the humanities.

?, for example, have studied the information work of humanities scholars who work on an interdisciplinary basis, examining a fairly wide range of different disciplines. Their findings suggest that the scholars under investigation rely primarily on collaborative and informal relationships to explore the unknown in the context of their interdisciplinary work. ? in turn found out that the information seeking of humanities graduate students differs significantly in some points compared to undergraduate students and established scientists. In other respects there are strong parallels, leading Barrett to believe that graduates go through extremely formative years in their academic careers. Many studies of information seeking by humanities scholars were conducted with the intention to use the results to improve the design of information systems (e.g. ???).

The relationship of the humanities to digital tools and the discipline of digital humanities has already been extensively reflected in the literature (e.g. ?????????????). Researchers in the digital humanities are concerned with e.g. the development and use of technologies for understanding and interpreting digital artifacts of humanities interests (?), e.g. texts, images or audio and video recordings. Research on the digital humanities includes the needs of the discipline (????), developments in the tool landscape (?), requirements of digital infrastructure (?) new research practices of humanists (?), the influence of and the intersection with information science (??), the influence and role of libraries and librarians in the digital humanities project (??) and also criticism of previous projects and the perceived orientation of the digital humanities (??).

This work focuses on information seeking practices of humanities scholars in the search for research software. Literature from the field of information practices as well as information behavior provides only very few references. ?, p. 231 formulates it appropriately: "*Infrastructure for facilitating access to and reuse of research publications and data is well established nowadays. However, such is not the case for software.*" This concerns both technical infrastructures and the practices of their use. While infrastructures and the development of research software has gained momentum across disciplines promoted by the Research Software Engineering project (?), the situation is somewhat different for the seeking and use of research software. This is especially true for the humanities. Seeking tools or information about tool usage has not yet been covered in the literature to my knowledge. The only concrete example is ?, who investigated the information seeking behavior of journalism professors when searching for online tools for teaching. Yet it is particularly important for a sustainable digital infrastructure in the humanities (?) that the tools developed are also found by its current and future users. This is where the investigation of my thesis departs, focusing on the information seeking practices of humanities scholars when seeking software.

3 Research Design

Since "[u]nderstanding the nature of information practices and their relation to the production of scholarship is important for both theoretical and applied work in library and information science (LIS)" (?, p. 165) this thesis sets out to study information practices of humanities scholars and their seeking for research software to better understand humanists needs and future LIS services (??). I chose an exploratory study design (?) where the personal realm of experience of each interviewee lies in the center of the analysis. Through an explorative approach to the object of research, qualitative social research approaches the social and subjective reality constructed by humans. Theory development is usually inductive, using the individual cases of empirical studies. Acting in a sociocultural context is always an interpretative process. In this respect, ?, p. 20ff distinguishes six principles that qualitative research should be guided by: openness, communication, process-like, reflexivity, explication, and flexibility.

The aim of this work is to investigate the information-seeking practices of early-career philologists when seeking research software (RQ1). This research focuses on philologists problems, contradictions and barriers in finding information and their information sharing about research software. Special emphasis will be placed on the respondents' recourse to their own research process and the knowledge and practice structures in their field (?) which are socially constructed (RQ2):

- RQ1:** What information seeking practices humanities scholars engage in when looking for software to use with research data?
- RQ2:** How do context and domain specific structures shape the information practices of humanists?

With this thesis I want to comply with the principles of an open science¹. Hence, I have decided that all data generated during the concept, survey, analysis, and writing phase is publicly available on GitHub², as long as it meets research ethics standards of the ? : Interview audio records and unanonymized interview transcripts are excluded. Prior to the study, I obtained the consent of the interviewees regarding the interview recording and later regarding the transcribed interviews.

¹<https://ocsdnet.org/manifesto/open-science-manifesto/>

²https://github.com/geyslein/Masters_Thesis

3.1 Methods of Data Collection and Processing

According to ?, p. 329, the qualitative paradigm sees the interview as the ideal way to collect data. Compared to the participating observation, it is much easier to realize. Moreover, qualitative research on the interpretation of texts is very advanced. The qualitative interview can be recorded, intersubjectively traced, and arbitrarily reproduced. Consequently, the findings of the study are derived from accounts of information practices and not from observation of information practices as it happened.

The first step in data collection was to establish contact with the potential interview partners. It was necessary to consider criteria based on preliminary considerations derived from theory. The goal of this work is to investigate information practices of humanities scholars in the search for software. I assumed that software is now an integral part of scientific work in the humanities as well, so I did not make any further preliminary considerations in this regard. Since there is a very large spectrum of subjects within the humanities, I had to minimize the resulting side effects during the investigation. Practical criteria of field access helped with the selection, according to which I chose the area of classical philology. With this convenience sampling strategy, I gained access to two participants with little effort. The two participants then provided referrals for new participants. Through the snowball sampling strategy (?), I was able to achieve a desired sample size of 5 participants relatively quickly. None of the contacted persons refused the participation in the study.

Interviews are the main data gathering technique which are guided by semi-structured interview guidelines, and implemented in a face-to-face manner (?) in German language. With the interviews I obtain emotions, thoughts, and intentions of the participants and discover their perspective of the social world (?). Due to the corona pandemic situation, I originally planned the interviews as a virtual setup with the help of web conferencing software. Since a reduced attention span and lack of experience with web conferencing was assumed among the interviewees, I planned two interviews per participant. The first two interviewees already accepted this setting. In the meantime, however, the first relaxation measures took effect, which is why we agreed on face-to-face interviews. Such a setup with two staggered interviews (planned minimum interval of 2 weeks) increases flexibility, as it allows for adjustments in the second interview after the analysis of the first interview. In addition, I was able to refer to contents of the first interview. Furthermore, it enabled reflection and self-observation processes among the interviewees regarding the search and use of software. I conducted 9 interviews. Four interviewees were interviewed twice. The interview duration varied between 60 and 90 minutes. Due to practical reasons, the fifth interviewee agreed to conduct a longer interview (duration 125 minutes) instead of the two-interview-setup.

The interviewees agreed to the interviews, their recording and the publication of the anonymized protocols by signing a declaration of consent at the beginning of

the first interview³. In the beginning of the interviews, I briefly introduced myself and the background of the study. All interviews were conducted in the offices of the interviewees or in meeting rooms at the interviewees institute. During the interviews I followed the structure of the interview guidelines.

The interview guidelines support the thematic structuring of the interview, but should nevertheless leave enough free space for the qualitative process during the interview as well as for the ideas of the interview partners. At the beginning of the interview guidelines, there are questions about studies, doctorate, previous work experience and the current position. It is followed by questions about the contents, the methods used and the theories of the own research. Afterwards, the field of investigation is approached by discussing the search process in general, using the example of one's own literature research. And on the other hand, the not very familiar topic of software will be opened up by means of easy introductory questions (Section A). Section B focuses on the sources that are used in the search for software. The largest of the sections, Section C, focuses on information practices in the search for and use of software. Based on the literature on information practices, the section is divided into seeking, scanning, monitoring, proxy, context and avoiding. In Section D, the last part of the interview guide, follow-up questions for the second interview are provided, which are chosen depending on the respondent and the course of the first interview. The complete interview guide is listed in the appendix (??).

To ensure the reproducibility of the interviews, they were recorded with a standard voice recorder and, in parallel, with a voice recording app for an android smartphone. The interviews were recorded without disturbances as far as possible and could be transcribed completely due to good recording quality. The recording of an interview actually contradicts the demand for a natural communication situation and possibly causes rejection on the part of the interviewees. Guaranteeing anonymity and informing the interviewees about the necessity of recording should prevent this. By using tape or digital recording devices, it is also possible to conduct the interview situation inconspicuously and in its natural environment.

For a later analysis of the conversation content a transcription is necessary. Transcription means the writing of audiovisual material, in this case the audio recordings of the interviews. The transcript as a result of this process thus already embodies a first subjective selection and reduction of the interview recording (?, p. 321). Once the interviews were conducted, I orthographically transcribed the interviews recordings using EasyTranscript, an open source transcription software⁴. The transcription system used for this purpose should be based on the specific research objective (?, p. 331). The aim of this work is an evaluation of the discussions in terms of content and topic. A sophisticated transcription system, which is required for a conversation-analytical examination, was therefore not necessary. I chose a verbatim and partially annotated transcription system, which considers several para- and

³see [https://github.com/geyslein/Masters_Thesis/blob/master/_data/collection/consent%20form%20interview%20\(German\).md](https://github.com/geyslein/Masters_Thesis/blob/master/_data/collection/consent%20form%20interview%20(German).md) for the consent form

⁴<https://www.e-werkzeug.eu/index.php/de/produkte/easytranscript>

Table 1: Transcription Conventions

(lachen)	Laughter
ähm	Non-lexical utterance (uh, erm, um)
[Anony]	Anonymized parts of the transcript
(DESCRIPTIONS)	Further Explanations if text was anonymized or not considered (off-topic)

non-verbal aspects. Conventions are summarized in table ??.

I did not check the finished transcripts for correct spelling after the transcription for reasons of time economy. Further, I translated the relevant text passages for the final thesis into English anyway, whereby attention was paid to the correct use of spelling and grammar.

In the next step, the interview transcripts were anonymized from information regarding person, institutes, university, thesis and research topics, and criticism expressed during the interview. This was done at the request of all interview participants. They attached great importance to this step already before we conducted the interviews. Further, they intensively controlled the transcripts after the anonymization. On request, I then anonymized further aspects in the interviews.

The process finally resulted in 12 hours and 22 minutes of audio recordings and 162 pages of interview transcripts (Arial, 12pt).

3.2 Process of Data Analysis

After I received consent from all participants for the anonymized research transcripts, I analyzed the transcripts with a qualitative content analysis to explore qualified hypotheses (????). It enables the researcher to include textual information and to identify its properties systematically. In detail, I chose a qualitative content analysis according to ?. For the analysis process I used an open source software called QualCoder⁵, which supports coding, annotation and category building.

In a first step, I approached the material using systematic open coding (?) to conceptualize and categorize the interview data. For this purpose, I have deductively identified several categories from the knowledge about information practices, the field, and the process of compiling the interview guide as a starting point for the analysis. During the analysis, I inductively derived codes and categories and constantly revised the existing ones. The previously established category system is thus constantly modified and further developed, taking into account the demand for openness and flexibility of the research process.

?, p. 65 describes three basic interpretation techniques: reduction, explication and

⁵<https://github.com/ccbogel/QualCoder>

structuring. With the analysis I aim to *reduce* the material to the core statements. For the interpretation of the present interview data I have chosen the *inductive category formation* as a form of data reduction. In contrast to summarizing, the other and very extensive reduction method, inductive category formation considers only those parts relevant for the research question and the step of paraphrasing is skipped (?, p. 79). The final category system can be looked up in the Appendix (??).

During transcription, coding, analysis and even during the writing process, I was able to contact interviewees for field-specific questions. Private contacts with two interviewees and the lunch, which I promised in return for the study participation and which took place a few weeks after the interviews, were especially helpful. During these consultations, the interviewees answered my field-specific questions, such as correct descriptions and background information on the databases used or misunderstandings of certain workflows. I was also able to confront them with the findings and hypotheses I had developed and to obtain a qualified opinion and possibly a correction of the same.

4 Findings

The findings of the study are structured as follows. First, I will give a short overview of the field of classical philology under investigation (section ??). Afterwards I will briefly present the situation of software use in the research process of the participants (section ??). The main part of the results will be the information practices in search for and use of software (section ??), as well as field- or domain-specific factors regarding problems, barriers and contradictions in the search for software (section ??).

4.1 Introduction of the Empirical Field

Classical Philology is one of the oldest subjects at the interviewees university. It first flourished in the sixteenth century. In the nineteenth century, the institute became a center of classical philology. The dominant field of work at that time was edition philology. In the twentieth century, scholars opened up the great texts of antiquity in commentaries and interpretations. Similar to classical philology, the institute has declined in importance since its zenith in the 19th century and today counts about 20 scholars. Both Latin and Greek studies can be studied at the institute. In the present study I interviewed 5 members of the institute who are engaged as research assistants, PhD students or in post-doc positions at the institute. A brief profile of each interviewee is depicted in table ??.

During the interviews the participants attributed various characteristics to the field of classical philology or the way classical philologists think and work which

Table 2: Overview of the Interviewed Classical Philologists

Name	Education	Position	Interviews	
			No.	Duration
Sandra	Latin and Greek studies	Post-doc	2	2h25min
Peter	Latin and Greek studies	PhD student	1	2h09min
Marie	Latin and Greek studies	Post-doc	2	2h44min
Marco	Latin and Greek studies, teacher training	PhD student	2	2h56min
Christian	Latin and Greek studies, teacher training	PhD student	2	2h07min

participants perceived as typical. In Sandra’s opinion, classical philologist ”[...] *are quite conservative in their approach. They work the way they have always worked.* [Sandra, 1st interview¹]. This attribute was unanimously confirmed by the statements of the other interviewees. This means that they maintain a very analog way of working, use few digital tools and work mainly with the printed book. Furthermore, research is marked by an extremely critical and exact engagement with the available primary sources which is ”*a very central question*” [Peter, 1] for the whole field of classical philology. The sub domain edition science in which textual criticism takes place is ”[...] *actually the royal discipline [...] You just have to be extremely strict with yourself.*” [Sandra, 1]. In addition, a very friendly, helpful and open atmosphere is maintained among colleagues. This was mentioned in all interviews and also the way how easy the already mentioned snowball sampling (section ??) was supports this impression.

In addition to edition studies, there are also other sub domains of classical philology. In the interviews, for example, literary studies, philosophy, palaeography, papyrology, history, and translatology were mentioned as sub domains in which the participants are active or at least used to be active. Meanwhile there are also several professorships for digital humanities (DH) at the university, which will complement and continue the existing cooperation with the department of computer science and above all act as mediators between the humanities and the computer sciences.

4.2 Software Use

To get a first overview of the software used by the classical philologists examined, I will briefly introduce the tools mentioned by the participants in the following section. The conservative working style mentioned in the previous section based on few digital tools becomes visible in the number and scope of the tools used and mentioned during the interviews. The most frequently used tool is the text editor Word (Microsoft) from Microsoft, which the interviewees use for writing research papers and for managing their literature references:

” *Well, I actually work quite analog I must say. I have a Word document and there I have my literature list, which I partly keyword to find things using the search function. [...] But this is basically a very analog way of working.*” [Marco, 1]

Marco confirms that philologists work very analogously, i.e. with few software tools. Although there are many software tools dedicated to literature management, all interviewees stated that they also use Word for literature management. And all of them also mentioned that they heard about the Citavi² reference management

¹In the following, I use the convention [NAME, No. of Interview], for example [Sandra, 1] for the first and [Sandra, 2] for the second interview

²<https://www.citavi.com/en>

software and recognized its usefulness for that very purpose. Christian explained it as follows:

"I know that Citavi exists. But I was so used to work with Word from the beginning that I always manage my literature this way. I wanted to get used to Citavi, but it scared me a bit, because I had heard that it takes a long time to get used to it, so I thought, okay, now you do it as always."
[Christian, 1]

For most of the interviewees the learning curve to switch to Citavi was too big, they had progressed too far in their dissertation so that a change was no longer worthwhile or they had already learned in their bachelor or master studies to use Word and did not feel the urgent need to change it later. Two interviewees explicitly mentioned a small tool, *typegreek*³, that facilitates the writing of ancient Greek text with politonic, unicode-compliant characters. Marco was apparently the first to discover this tool and then recommended it to his colleague Peter who mentioned that *"If I had known this in my studies, many seminar papers would have been less painful."* [Peter, 1].

However, the largest collection of tools used was mentioned in terms of searching and accessing literature. Interviewees often use databases or digital libraries that contain primary sources of classical philology. For the Greek language, the most prominent database is the Thesaurus Linguae Graecae (TLG)⁴, which is a collection of all texts written in Greek from antiquity to the present era. A similar database exists also for the Latin language, the Library of Latin Texts (LLTA)⁵. Interviewees regard both tools as standard tools that are basically without alternative in the digital world. Sandra comments for the TLG:

"You can't build a more complete TLG now than the one that is there. This is not possible per se and nothing will change substantially. [...] TLG they really have a kind of monopoly on the data, on the texts." [Sandra, 2]

The same applies to the bibliographic databases. Here, the *l'Année philologique* (Aph)⁶ is generally considered the standard index for scholarly work on Ancient Greece or Rome: *"the l'Année philologique, that is the be-all and end-all."* [Marie, 1]. At the beginning of their research process, the interviewees would always seek in the Aph for research on a particular topic. In addition to bibliographic databases, review journals play an important role for the interviewees in their search for current research work and publications. The international journal *Bryn Mawr Classical Review* (BMCR)⁷ as well as the German journal *Gnomon*⁸ were frequently mentioned

³<http://www.typegreek.com/>

⁴<https://stephanus.tlg.uci.edu>

⁵<https://about.brepols.net/library-of-latin-texts/>

⁶<http://www.brepols.net/Pages/BrowseBySeries.aspx?TreeSeries=APH-O>

⁷<http://bmcr.brynmawr.edu/>

⁸<https://www.chbeck.de/buecher/zeitschriften/gnomon/>

by the interviewees. Sandra on the distribution and advantages of BMCR over other review journals:

"What actually almost all classical philologists have subscribed to as newsletter is the so-called BMCR, that is Bryn Mawr Classical Review. That is a review journal. The only difference, it's not better than all the others, but it just appears online. And they send it to you right away. Whenever someone finishes it, they send it to you, and that way it gets read a lot more than anything else. They just noticed very quickly in which direction digitization is going to. I assume that many more people will read it. Because it's just totally convenient, I get it via email." [Sandra, 1]

In particular, she highlights the immediate online availability and email notification as advantages over other services, which is why BMCR has quickly become the standard among philologists. Some interviewees also use Google Scholar⁹ for literature research. A more recent phenomenon among the participants is the use of Academia¹⁰, a social networking website for academics where they can share ideas, research interests, and publications. Marco thinks it is *"[...] sometimes very handy when people have just uploaded stuff that they have published."* [Marco, 1]. Further, Academia suggests publications to researchers based on their research interests and previous interactions on the website.

In conclusion, it can be stated that a large part of the tools mentioned by the interviewees are used for locating and accessing publications and that much of it is used on websites as server-side software. See the Glossary of Mentioned Software and Services (p. ??) for a list of all the tools that were mentioned by the interviewees during the empirical study.

4.3 Information Practices Towards Software

In the last section I introduced software tools that the interviewees use in their research projects. This section now focuses on the information practices that the classical philologists engage when seeking for software tools for research or for information on the use of such software. I will structure this section according to the investigated modes of information practices in the accounts of the participants. A general overview of the perceived information practices is presented in the introduction section ???. The very dominant mode, seeking, is then discussed in section ??. This is followed by explanations on scanning (section ??) and monitoring (section ??). Both information practices were rarely reported by the Interviewees. In contrast, Interviewees were more often referred to software or solutions to software problems by proxy (section ??) through colleagues or friends.

⁹<https://scholar.google.com>

¹⁰<https://www.academia.edu>

4.3.1 Introduction

Most interviewees were irritated when I asked them about the search for software. They had to think for a while about when they last searched for a specific tool. The reaction is not necessarily surprising, considering that classical philologists use very little software in their everyday research: "*[...] we don't use so much technology, so to speak, or many software.*" [Christian, 2]. Therefore, the information practices for searching for software are more difficult for the participants to retrieve in their accounts during the interview. In contrast to literature research, seeking software or information about software usage is rarely practiced.

4.3.2 Seeking

All Interviewees write their research papers using the text editor Word. Knowledge of how to use functions perceived as complex, such as the automatic creation of directories, the handling of special characters or the automatic formatting of heading hierarchies, is often gained by using Google's search engine and its search results:

"My knowledge comes from Google research, so according to the motto Google is my friend. And with Word I do almost everything, exactly with formatting stuff and almost everything alone.[...] I google all the time."
[Marie, 2]

The Google search engine is often the first source in such a search process. However, since the search for software enjoys only a subordinate priority in terms of available time resources, seeking is also aborted relatively early if the desired search success does not materialize:

"Yes, I am more likely to drop out than during literature research. Then I give up and think okay, then that's the way it is, then I have to get along somehow differently. [...] This is also due to time, above all. I think there are more important things, I just put my focus on my work, like literature research or writing whatever. And then I think okay, you are wasting too much time here, so I am too impatient." [Christian, 2]

The interviewees often break off the search relatively early due to time constraints. Instead, they invest their available resources in literature research and in writing research papers what they consider to be far more important. In addition to such cost-benefit considerations, their own ability to search also plays a role, as does their understanding of what they are looking for. Peter summarizes this insight:

"If you somehow don't really know what to query, then you stop searching. Or you have to understand something about searching, and if you understand little about it, then you search less and are then clearly only need-oriented in terms of the concrete. I really need this and that now, then I'll see if I can find it." [Peter, 1]

Following Peter’s explanations, he would only actively search for software if he can clearly outline the problem and the need and thus his own benefit is clearly recognizable. When it comes to digital resources, in addition to Google, the software itself is an important source for information about its own functioning. This practice consistently refers to the search for information about software in use. The software itself can serve as a source of information for its own use in two ways. Firstly, some tools offer extensive support, which is directly integrated into the software or which can be accessed via a link to the respective website. The TLG is a good example of this and, in Marie’s opinion, offers extensive help on topics concerning its own use:

” There are lots of tutorials. And for each function you can download a PDF that explains everything. [...] And yesterday I gave a recommendation to my students. There are several functions, it’s really very well done, [...] the parallel comparison of texts and it’s quite complex. I can show you later if you want.” [Marie, 2]

The practice of searching for information within the software itself was also confirmed by Marco [Marco, 2] for the LLTA database. Furthermore, the software can be explored in a kind of learning-by-doing or by trial-and-error. Marco explains how he discovered a specific function in the LLTA database:

” One day, I discovered a new function in this LLTA database or I found it out by trial and error. This is a function, which is a bit hidden, to check word frequencies of certain authors or works. [...] I looked to see if there was a function. [...] It wasn’t as easy as I thought, but you had to click on things you don’t usually click on and then you find it somehow.” [Marco, 2]

In an explorative engagement with the LLTA database, Marco discovered a functionality that was previously unknown to him. In his opinion the function he needed should be available in LLTA, an assumption he derived from other functions he had already discovered. To look for functionality within a software seems to be a common strategy for the interviewees as Christian confirms:

” So especially with TLG [...] I do a lot intuitively. I muddle my way through until it works. Well, I did it all myself, most of it is really self-taught.” [Christian, 1] *” And with me it is also a lot of trying, so intuitive. And you have to look, okay, where could the problem be solved now. And then I somehow try to solve it myself. [...] That is also difficult. And then you either give up or you intuitively come up with it.”* [Christian, 2]

Christian refers his statement among other things to his work with the TLG. He calls the way he proceeds to discover new functionality intuitive, i.e. instinctively and without conscious reasoning. An aspect that also shines through in his use of the verb ‘muddle through’, a non-systematic strategy of information seeking. He somehow

manages to reach the goal, i.e. the desired information, the desired functionality of the software. The concrete process remains hidden for him and is not further questioned by him.

An extremely prominent practice of searching for information among the participants is to ask colleagues and friends for help. This strategy was confirmed by all interviewees and is not limited to the search for software and software functions. In fact, this practice is also applied above all in relation to literature research, which is so important for classical philologists. Marie summarizes the importance of her colleagues very clearly:

"In fact, my colleagues are a source of information for me, the best, probably more important than other resources." [Marie, 2]

The information practice of asking colleagues for advice runs through all interviews. The help is mainly used when it comes to using the large literature databases such as TLG, LLTA, TLA or others. Christian and Marie have to work a lot with the TLG database and consult their colleagues frequently about its software functionality:

"Then I would ask colleagues who use it. Especially with the TLG, actually. So there I also asked Sandra." [Christian, 1]

"In fact the answer is a bit trivial, I ask colleagues frequently [...] the help in using the TLG portal came from my colleagues." [Marie, 1]

The participants would consult colleagues especially for questions on the use of subject-specific literature databases. However, friends are often consulted for more general but still research related questions about software and its use. This will be often the use of the text editor Word but also questions about tools for project management. Christian explains his experience with Word:

"Some things are not known at all. So now also in Word, I discovered things during my dissertation, which I learned from a friend. [...] She recommended various tools and also various functions. [...] Yes, it actually mostly refers to Word." [Christian, 2]

Many interviewees can fall back on people in their circle of friends who have sound knowledge about certain software topics. Sandra, who herself is frequently consulted by her colleagues on questions about the TLG, also seeks the help of friends for more complicated questions. Especially when the information need cannot be described so precisely that the search would produce fruitful results:

"The question is, what am I looking for? You just can't search for software for phonetic sequences or something like that. You won't find anything there. Even if you google it a little more precisely, you won't find anything. And I don't think there is anything that deals with this specific question. [...] I then asked a friend. I didn't get anywhere on my own and then I asked around to people who I think know more about these problems." [Sandra, 2]

Participants use various information practices to actively seek information about software. They actively seek for help via search engines, whereby the search is often aborted quite early. The software itself often serves as a source of information, either in the form of assistance or by exploring the software. Moreover, the interviewees very often consult friends and colleagues for questions about software and software functions.

4.3.3 Scanning

In the narratives of the interviewees, the information practice of scanning was very rarely referred to when they needed information about specific software or its use for their research projects. One of the rare examples of scanning was provided by Sandra. She was looking for a specific software for a very clearly defined purpose, a special kind of analysis of phonetic sequences:

"I also took a look at what is happening in the Digital Humanities. I looked a little bit, especially if there was something about phonetic sequences, because it is so convenient to do that, if something was published there and I didn't find anything that would help me. Maybe I didn't search systematically, but of course I have seen many other tools, a lot of software, but for completely different problems most of the time." [Sandra, 2]

In search of initial information to solve her problem, Sandra scanned through offers made by digital humanities, mainly journals and conference papers. And although she didn't find any suitable software in this way, dealing with the topic gave her a better idea of what kind of software she was looking for and how she could better describe its functionality. During her research, she also listened to a guest lecture that dealt with stylometric analyses in classical philology. It can be considered as an arena for potential information and she also scanned the lecture for information to her software need:

"It is basically a stylistic question. And then a friend of mine pointed out to me that there are different software, for example this Stylo¹¹. And later, by chance, through a lecture, that there are also other stylistic tools, which are not so suitable for this problem, because they work with word frequencies and such things." [Sandra, 2]

The lecture also enabled Sandra to further narrow down her problem, even though her scanning did not solve it.

Scanning as an information practice occurs very rarely in the field I observed and when it does, it occurs in connection with a very concrete information need.

¹¹<https://github.com/computationalstylistics/stylo>

4.3.4 Monitoring

Monitoring as an information practice seems to occur as rarely as scanning (section ??). The accounts of the participants allow only few conclusions about serendipitous information encounters. Due to the fact that concrete demands for certain software tools seldom occur, this will probably also rarely lead to accidental findings during monitoring.

However, Sandra's search for a tool for the analysis of phonetic sequences provides another example, this time for monitoring:

"My acquaintance, who is also working a lot on data analysis but also has a greek background, told me that there is a good software. [...] And then we simply talked about it. [...] And then I asked him, is that possible, can you do that for me too, does it make sense for my phonetic sequence problem? And then he quickly gave me what I had been looking for for a very long time [...] an hour later I had the diagrams on my desk and it was really good. [...] But that was also very coincidental. If he hadn't had the problem of what someone else had given him, who is basically a classical philologist and has no idea about programming etc., he wouldn't have known it either." [Sandra, 2]

When her acquaintance told him about his current project, Sandra quickly realized that he talked about the tool she was actually looking for desperately (see section ??) for quite a while. Another example of serendipitous encounter was introduced by Christian, who regularly exchanges experiences with another doctoral student about the progress of their dissertations:

"That was more of a coincidence, that was more of a status update on her work process, so to speak. She is also writing her dissertation at the moment." [Christian, 2]

During one of these sessions, Christian's friend casually tells him that he has once again automatically sorted the bibliography. Triggered by this information, Christian directly intervened at this point and thus learned about this option. He further asked his friend to show him how to implement it.

Random encounters of information occur only very rarely among the participants. As in the case of scanning, this also depends on a concrete need for information about software. However, this is rarely the case among the interviewees.

4.3.5 Proxy

Information practices through a proxy occur when another person actively seeks or scans on behalf of the primary information seeker. This requires the proxy person to be aware of the information needs of the primary seeker. In the underlying empirical study, the colleagues often act as a proxy. Due to the close collegial ties at the

institute, the colleagues know each other as well as their research projects and current challenges very well. They would not only be contacted as sources during information seeking, as already shown in section ???. They also play an active role by taking the information needs of their colleagues into account in their own information practices such as seeking and scanning. In his comments, Sandra describes a very typical situation in which information practices through a proxy often occur:

"It's even common to talk over coffee: I'm working on that right now. And that the other person then says: and have you already checked that? Well, that's very important, I would say. That's right, I hadn't thought of that at all. It's just not the systematic approach, but it's part of it."
[Sandra, 1]

In his account, Marco later calls it *"this coffee break phenomenon"* [Marco, 1]. As we can see in Sandra's remarks, this form of information practice seems unsystematic in her eyes and is often overlooked by the interviewees. During the interviews, information practices about proxy or the colleagues were mentioned rather casually and I often had to ask specific questions. However, when the topic was then addressed, the importance of the colleagues as an important source of information repeatedly emerged in the narratives:

"So that's, in my opinion, I really believe that it's better to work together with other people, even if you're just sitting on your own topic, than just sitting alone in a quiet little room." [Marco, 1]

By having colleagues act as proxies, among other things, we often create what is generally understood by a collegial relationship. The other person has you on their screen just as you have them. Acting as a proxy in this study is based on this trust and mutual give and take. In his accounts, Marco explains how he learned about the LLTA database:

"I think Peter gave me the tip once but I can't tell it for sure anymore. [...] Well, he explained to me what the tool does when I remember correctly. And then I thought, this is actually quite practical. And then I just experimented with all the things you can do with asterisk and all the different parameters you can enter, I played around a bit." [Marco, 1]

The reference to the LLTA database triggered another information seeking for Marco. After he had access to the database, he explored its functionality and decided to use the database for his own project in the future. By passing on information about software or software functions, the information seeker receives information from a source that he or she trusts with regard to competencies in classical philology. This is one of the main advantages of such proxy connections compared to active information seeking. Marco tells how this proxy process meanders through the institute, from office to office:

"It's not that we just google and then we find something. Rather, it goes from office to office without anyone remembering how it happened. But somebody must have discovered it first. With the LLTA database it is relatively easy, the university has subscribed to it [...] and we can use it for free. And somebody here tried it and found it to be good and now everybody uses it." [Marco, 2]

For aspects such as software or digital databases, but also with regard to literature, the colleagues seem to form a common information group. As soon as an information enters this circle, it is obviously passed on to the colleagues, at least to those who need this particular information. Peter thus came across the typegreek tool already mentioned in section ??:

"And the degree of familiarity with certain practical software tools is quite variable. My colleague Marco once recommended a tool for Greek, with which you can type a Greek text in no time, which is usually a cramp with the keyboard layout." [Peter, 1]

In his narrative, Peter explains that his awareness of certain software tools fluctuates quite a bit and that he is not always informed about all possibilities. By sharing the burden of searching for information and then passing it on to his colleagues via proxy, the colleagues compensate for certain deficits in seeking software.

As already shown in section ??, when seeking for information about software, the interviewees also trust friends to whom software-specific knowledge is attributed. Marie summarizes the potential proxies among her colleagues and friends:

"A former colleague, my boyfriend, who also knows his way around, two other colleagues, these are usually the people that help me out." [Marie, 2]

In the present study, however, there was only little evidence of friends acting as a proxy. In her search for a suitable tool for the phonetic sequence problem (see section ??), Sandra falls back on the help of two friends. These friends also act as proxies from time to time, alerting Sandra to potentially helpful events or software tools. Furthermore, it is not always clear in the narratives whether a certain action is to be classified as active information seeking or as passive recording via a proxy. These two modes partly overlap and are sometimes not described precisely enough. ?, p. 35 describes those practical limitations as *"[...] much too complex to permit a clear-cut distinction between the three information-seeker-initiated forms of information practice and the agent-initiated 'being told'"*. The assumption is, however, very strong that there are cases of friends acting as proxy.

The proxy mode of information practices is very prominent among the studied participants. Due to the proximity of content to their own work, mainly colleagues who are also classical philologists act as proxies. Information about software thus rapidly reaches the colleagues in need for this information.

4.4 Domain Factors

The limits of my language mean the limits of my world.

Tractatus logico-philosophicus

LUDWIG WITTGENSTEIN

In the previous section, the main information practices of the classical philologists examined in the study were presented. On a descriptive basis, I used the narratives to work out how the interviewees pursue information about software and which sources they use. The following section picks up on some of these descriptions about information practices and interprets them in the light of the subjects' experiences in classical philology.

4.4.1 Negative Experiences with Digital Humanities Projects

Some interviewees reported quite clearly of their own negative experiences in project work in the field of DH. This applies both to the cooperation in the project and the results of the projects they themselves have participated in as well as the results of other projects in the field of DH. Sandra has worked in a DH project herself and has described her negative experiences very clearly during the interviews:

"It often goes wrong. I've already experienced that myself [...] You submit an application to the DFG or elsewhere. And those who come from classical philology actually have very concrete questions. And they don't want anything new to be invented. They just want the tools, they just want the hammer with which they can drive in their nails. [...] I once had the experience that people wanted to invent a new hammer, but then you couldn't drive the nail in with it anymore. [...] Since that happened to many people in classical philology, I can imagine that there is a certain frustration and that you think, oh no.[...] So there are an incredible number of projects within the Digital Humanities that are in ruins. There is a lot of money pumped into these projects, there is always something new and then half-finished websites are created that do not really work. And from this experience, that it often doesn't work so well after all, it may be that many in classical philology are not prepared to spend time with it. [...] but then it is also the case that it is often frustrating, that a lot is promised what can be done, but then it is not. [...] It's not as if something comes from the Digital Humanities and then you receive it with open arms." [Sandra, 1]

Often the developed tools were not used after the end of the project and the results were therefore not sustainable. Hence, a certain frustration arises among classical philologists with regard to the promises of and the experiences with the DH. For

this reason, and also from her own experience, Sandra believes that DH projects in classical philology are viewed with caution and have a hard time. Following the statements of the interviewees, the DH claim to be a science in its own right. This leads to the self-conception in DH projects to not only slip into the role of engineers for the classical philologists, but to master own research problems with own methods. According to Marco, this claim has yet to be proven:

"It is not obvious to me that a completely new way of working is now emerging where the methods of both sciences come together. I also don't know whether it would bring anything. I don't think it's really been tried out yet. If you define digital humanities in this way, then I don't think we're ready to say that. I don't know of any concrete project where I would say this is one of those cases where the methods of classical philology have really changed through computer methods. [...] That's what computer science explicitly wants, when they complain that they're ever regarded as assistants only. So you are claiming that your own methods can also have a value for classical studies in terms of content. I don't want to deny that either, but no one has ever tried it out or seriously tried to show it."
[Marco, 2]

Marco at this point and also Sandra in her previous statement suggest that they just need the right tools to deal with their own problems from classical philology. And this does not necessarily mean that a new science has to emerge. As a result of the negative experiences made and the not entirely friction-less relationship with the DH, the subjects are reluctant to include these actually obvious sources in their practices for information about software. Regarding her phonetic sequence problem, Sandra initially considered asking for advice within the DH, for example at the DH chairs at the university:

"Of course I could have done something with colleagues from the Digital Humanities, I could have called them and asked them what was going on there. But to be honest, I was a little shy about that. Simply because, as is sometimes the case with computer scientists, they say, well, we've been at it for a long time, but I'm 100% sure that the question itself was not so clear. [...] that is, it would be a bit like that without wanting to accuse anyone of anything, but that would have been such a stumbling block that I said I didn't really want to do that, it was uncomfortable. And that's why I didn't do it [...] although normally I wouldn't hesitate to contact anybody." [Sandra, 2]

Sandra avoids the interface to computer science or digital humanities, although she would actually have a very concrete and suitable request for computer scientists. As a consequence, the interviewees seek contact with those colleagues to whom they attribute knowledge from the field of computer science. Marie explains:

"My colleague, for example, is not a computer scientist in that sense, but he has a lot of experience with it. I think, in fact, I would rather ask him, who is also a classical philologist, than perhaps a computer scientist. Also a bit because he can perhaps better understand my concern, because we speak the same language, but he also speaks the other one. [...] I think digital humanities are also the future. Nevertheless we have the task of interpreting the texts. And it doesn't exhaust itself by designing great tools or anything. So no computer can take over the interpretation of the texts. I think that's also a bit of the problem." [Marie, 2]

The problem with the interface to computer science is certainly more complex than the interviewees have described in their accounts. Nevertheless, it can be deduced from this that the negative experiences made in the field of digital humanities or with computer science in the broadest sense probably led to the fact that the classical philologists examined did not adopt them in their information practice with regard to software tools. The representatives of the DH would be extremely suitable connections for software-related questions. However, the subjects lack confidence in these persons or in this branch of science, so they would fall back on colleagues or friends they know well for software-related questions.

4.4.2 Information Seeking and the Limits of Language

During the interview, when Peter quoted Ludwig ? to explain the problems he has while seeking for software, he brought it very symbolically to the point: what you cannot describe in your language remained and maybe remains hidden to you (also with regard to information practices):

"If one says, the limit of my language is the limit of my world, then this could be true. [...] that is indeed sometimes very difficult to get a description for what one wants. [...] Why is that so? Because of my speechlessness. (laughter) No, yes, it's because of the missing insights." [Peter, 1]

To start seeking for suitable software or software functionality, it is a prerequisite that the problem can be described sufficiently well. If, to stay with the metaphor, you come from another world, classical philology, speak another language and thus do not come from the world of 'software' with the accompanying understanding and language, then it is difficult to express your own information needs around software in such a way that the sources you use to do so can make sense of these formulations. In terms of information practices around software, Peter just seems to be speechless. This speechlessness presumably leads to the fact that the participants increasingly carry out their information practices among colleagues (see section ?? and ??). In connection with speechlessness stands also a statement of Christian, who uses offers made by the search engine that would support seeking the right terms:

"So I'll start with keywords, but if I don't get any further, I'll just ask a question about software. [...] or a phrase, there is sometimes something coming. Then you see in some suggestions or hits you see the different words, so then yes, sometimes you come up with things. [...] that is also kind of a snowball principle." [Christian, 2]

Christian uses the suggestions for completing the search query by the search engine as a means to approach the searched topic step by step, which he describes in analogy to the snowball principle.

In their search for software, the classical philologists studied often reach the limits of their own language or their own world. The lack of experience in the field of computer science is the main reason for this. This leads to an increased focus on information practices within their own world, their colleagues. But also offers from the other world, the world of software, are used, as the last case described clearly shows.

4.4.3 Methodological Challenges in Classical Philology

I gained a very important insight for the interpretation of the available interview data only in the last interview of the study, the second interview with Marco. It concerned methodological reflection in classical philology or even in the entire humanities. Asked about the use of methods and their reflection in his scientific work, he replied as follows:

"I know what you mean. The classical humanities usually don't do that. Well, I know this from the educational sciences, from a seminar, qualitative and quantitative social research. That you first describe everything you have done. Basically a diary of your own work. That is simply not established in the humanities. I've never seen anyone doing that before. You just present the results, but how you came up with them is something people have to think about for themselves. Of course you give secondary literature, where you used it and so on. And you don't write, I just read the sources for half a year and then spent half a year on secondary literature. But yes, maybe if (laughter) I think it would be quite a revolution, if somebody would do it like that." [Marco, 2]

Marco was also partially trained in educational science. He has attended seminars there on qualitative and quantitative social research and is therefore familiar with the intensive methodological reflection in the social sciences. In the humanities, he has not yet encountered such institutions as methodological chapters in research papers or entire seminars on methodology. In his opinion, the research process is reflected little, while the result is reflected all the more. In our conversation, we then spent a relatively large amount of time examining how methods are used in the humanities. Marco seemed very taken with his insight into methodological reflection in the humanities and made an attempt to explain humanities' practices:

"The term 'humanities' [RG: in German Geisteswissenschaft, literally: science of mind] is quite precise about that. These are things that often take place only in our brains and apart from that are not easily represented by language. [...] Okay, I do that kind of thing, not process-oriented but where I say: my corpus consists of these texts and I have compiled it according to these criteria. That would be in the direction of making it transparent. But it is not a retelling of my own work. So at the beginning I wanted to do it that way and then I did it that way. So this is also thought coming from the result. But it is already so that readers can understand it. So why did he now consider these passages as fragments of letters of Catul and why not this passage where it says Catul wrote a letter. And of course this must be clear to the reader." [Marco, 2]

In Marco's view, the process of gaining knowledge in the humanities is difficult to explicate because it takes place in the mind of the scientist and the process cannot be represented in a procedural way. He therefore finds it difficult to reflect on methods.

If we consistently follow this train of thought and apply it to the present problem of information practices related to software, the interpretation suggests that a shortened reflection on methods also implies a shortened reflection on the tools used in the own research process. Less reflection in this case also means that less thought is given to potential application scenarios and to consequently create an actual information need regarding software and its usage. Hence, in the humanities fewer opportunities are created to seek software. Instead, the knowledge about new tools or their application is passed on among colleagues in the classical way of working. The colleagues act as an important source during information seeking or as a proxy.

4.4.4 Trusting the Source

The classical philologists studied use a number of software tools in their scientific work, mainly digital databases or libraries. As already shown, some tools, such as the TLG, the LLTA, the BMCR or the APh, are very important for their work. They are the standard tools for a certain work step and were even described by the participants as without alternative for their research work. As standard tools without alternatives, the interviewees have great confidence in them. Sandra explains it:

"It's that what is really established is just pushed so extremely and these things are just very, very recognized and actually a lot of people access them. And it's not that there is one and the other. And one person prefers this and the other person prefers that. I don't really know that from experience, but one thing works and the other thing doesn't seem to work." [Sandra, 2]

As these tools have become the standard, they are used by the broad mass of classical philologists and are widely trusted in terms of their functionality and content.

Trust is also of enormous importance for a sub-field of classical philology, namely edition science. In this field, textual criticism is the main activity and represents the royal discipline in classical philology (see section ??). The interviewees described its working methods in a way that reminded me of the work of a detective:

” You can’t really have confidence in a source. Actually, as a historical humanities scholar one should not trust sources, but one should evaluate them, use them, classify them critically and so on.” [Marco, 2]

Marco describes an extremely critical handling of sources which leads to the fact that classical philologists really trust only a few of them. Mistrust is the first priority of the work, trust in itself has to be justified by the source first. The question now arises how this knowledge about trust can be used to interpret information practices on software. When the classical philologists studied are seeking software functions, this often happens in the software itself. The examples described here refer mainly to the TLG and partly to the LLTA. Both are sources that the participants trust. However, both are also sources that in their eyes have no alternative (see section ??) and thus to a certain extent they have no choice but to trust them. In section ?? it was described that the interviewees often abandon the search for software quite early. It can be assumed that the topic of trust also plays a role here. As already shown in section ??, they do not act in their own world, the classical philology, but in the world of software, whose concepts they do not know and therefore cannot trust. Thus, it can be assumed that this is one reason why they abort the search early. Conversely, this may mean that they turn to more trustworthy sources and consult colleagues or friends in search of help on software.

4.4.5 Software and Data

Among scholars of ancient Greek, the TLG is the standard instrument for digital access to texts. As already described in section ??, the TLG offers almost all existing texts of Ancient Greek. When Sandra reported on the TLG, she also spoke about its de facto monopoly on texts in ancient Greek:

” TLG, they really have a kind of monopoly on the data, on the texts. [...] I think the development will go into the direction that TLG will simply get better and better. [...] let’s say you would develop a tool for determining phonetic sequences, that would be quite nice, an extra tool, but it would make much more sense to implement it in TLG as just one more among the many functions you have there. That would make much more sense than building something new.” [Sandra, 2]

Sandra describes that new tools should be built into the TLG itself because of the data monopoly of the TLG. Only from inside the TLG, if you are a paying subscriber, you have unlimited access to the data. If an independent tool were to be developed outside the TLG, this would raise the question of data access. The terms

of use of the TLG website¹² as well as the license agreements¹³ indicate that the data monopoly of TLG has also been legally established by means of a strict copyright and by a limited use permitted on the website only. Accessing and analyzing the data with another software tool is not permitted.

The restrictive access to the data of the TLG therefore certainly has effects on the information practices of the participants. As already described by Sandra, functionality offered by software outside the TLG are impractical because of the problem of data access. As a consequence, the classical philologist of ancient Greek does not develop the need to seek for software next to the TLG. Consequently, the search for required software functions is shifted to the TLG itself, as already shown (see section ??). And the need for new functionality is also formulated by the subjects in the direction of the TLG, as Sandra's feature request for the analysis of sound sequences in this section demonstrates.

The LLTA in turn is the counterpart to the TLG for the Latin language in terms of completeness. A similarly restrictive policy¹⁴ can be observed here as well. As long as the data is bound by restrictive licensing requirements to use with the database or the digital library itself, the interviewees are more likely to resort to and seek for software functionality within the LLTA. They would rather not seek new software outside the LLTA.

The Perseus Project and its Perseus Digital Library is a possible alternative to the almost complete, but also very access restricted databases of the TLG or the LLTA. The freely usable data is licensed under a CC-BY license. However, the Perseus Digital Library is limited to canonical world literature, it provides only segments of a single work and complete texts are not available (?). Thus, the Perseus Digital Library does not represent an alternative for many classical philologists, as they have to access all texts on a certain topic for many fields of research.

The restrictive terms of use of the databases TLG and LLTA, which are very popular among classical philologists, obviously lead to the fact that the interviewees increasingly seek software functionality within the database interface instead of seeking outside the databases for independent software that could offer these functions.

¹²<http://stephanus.tlg.uci.edu/copyright.php>

¹³<http://stephanus.tlg.uci.edu/site.php>

¹⁴<http://www.brepolis.net/terms.html>

5 Discussion

In this study I investigated the information seeking practices of early-career classical philologists when seeking research software. I further wanted to explore contextual and domain-specific factors that influence these practices. For this purpose, I have divided the discussion chapter analogously in section ?? (??) and section ?? (??). Furthermore, in section ?? I will present the practical implications of the results of this thesis.

5.1 Information Seeking Practices of Classical Philologists

The examined scholars of classical philology fall back on different information practices. One of them is active seeking via search engines, which, however, is often aborted quite early. On the other hand, colleagues or friends are often actively asked for advice regarding the search for software or software functionalities. In addition, in the case of software that is already in use, the participants often fall back on help functions within the software or they explore the software in practical use. Furthermore, I could often find a practice where the information was brought to the participants via a proxy and thus the practice was initiated by other agents. Avoiding information or sources of information was not directly presented to me as an information practice, but could be interpreted from statements. Some participants seem to avoid potential sources from the field of digital humanities (see section ??). The participants also rarely reported on practices that can be assigned to active scanning or non-directed monitoring. In general, the classical philologists studied spend relatively little time seeking specific software tools or information on their use. Thus, the results of this study confirm the statement of ?, p. 808, which certifies that the humanities have a low need for specialized equipment or infrastructure.

Cognitive Barriers to Information Seeking

In the present study, the participants abandoned the active seeking of information about software relatively early. Reasons for this are the lack of linguistic means to express their own information needs (see section ??). ? has identified six different barriers in his conceptual analysis of cognitive barriers to information seeking. In addition to the four barriers during the final stage of the information seeking processes (unawareness of information sources, low self-efficacy, poor search skills, and inability to deal with information overload), there are also two barriers in

the initial stage of information seeking that are of importance for this study. The speechlessness of classical philologists in describing their own information needs around software functionality is a cognitive barrier ?, p. 616 refers to as "*inability to articulate one's information needs*". ?, p. 137 calls it "*non-specifiability of information need*", whereby this can apply to both the cognitive and linguistic level. The study participants justified their inability to specify their information needs by referring to their poor domain knowledge (?, p. 619) in the field of software. The problem of finding suitable keywords on that very domain knowledge has already been investigated by ?, p. 285 with regard to information seeking problems of young people and can be confirmed with the results of the present study. Here, too, information seeking is discontinued at an early stage. Further, the results do not contradict with ?'s ? study on journalism professors information seeking who reported "*frustration at the fragmented state of information sources about online tools*". However, I did not analyze the quality of the mentioned information sources of the philologists which is why I can not make any statements about a potential fragmented state of information sources.

Further, the cognitive barriers are not detached from the other influencing factors. The cognitive and situational relevance (?) or pertinence (?) of a certain information probably plays just as important a role for the termination of information seeking as the cognitive barriers mentioned above. The situational relevance of the sought information about software was not decisive for the classical philologists investigated in the respective moments, in contrast to e.g. the writing of the doctoral thesis, which is why the search for software was aborted early. The results of this study suggest that when studying the early discontinuation of information seeking, both barriers and relevance should be considered equally important. This coincides with the results of ?, p. 454 who claimed that information seeking failures are "*not caused by one single reason, but triggered by various interconnected contextual factors starting with the tasks, which could bring external barriers*".

?, p. 616 calls the second barrier of the initial stage of the information seeking process "*unwillingness to see one's needs as information needs*". This is a further indication for the early termination of the information seeking process, which is also associated with low relevance. As already mentioned (see section ??), the classical philologists investigated have gaps in the reception and reflection of one's own methodology, which could also affect the reflection of the use of tools. At this point, I would argue that the lack of reflection on the process and the tools translates into a lower relevance of the tool use for the classical philologists and the related information needs. The unwillingness to see the need to reflect on methods thus leads to the unwillingness to see the information needs regarding software. As a consequence, this means that a search for software is not even initiated because there is no information need. This also leads to the possible interpretation that there might be no information avoidance (?, p. 97ff) or intentional ignorance of problems (?, p. 616), but "*information is [...] simply not used*" (?, p. 100). Since I can only make very vague assumptions here, it is up to psychological scholars of e.g. cognitive dissonance theory (?) or related subjects, as ?, p. 616 suggests, to find out the exact

background of these barriers.

Collaborative Information Seeking and Sharing

The investigated scholars very often fall back on collaborative information seeking (??) within the group of colleagues. The colleagues are very intensively involved in the search for information, both in general and with regard to software problems. This happens both in active seeking and in seeking via a proxy. In the second case, the information needs of the colleagues have to be considered in order to act as a proxy for them. For this purpose, an awareness (?) of the topics that the colleagues are currently working on and the needs that can be derived from them is necessary. According to the motto: I know my colleagues so well that I know what they are looking for or need. ?, p. 149 describes this kind of information sharing as *social sharing*, ?, p. 774 again refer to it as *serendipitous altruism*, assuming that "in similar circumstances, you will in future return the favor." In the present study I regard the behavior of the philologists as altruistic. Especially the proxy practice indicates a very cooperative relationship among colleagues.

A well-functioning information practice by proxy within a group of people has a not insignificant side effect on the other information practices. Information that has reached a person via a proxy no longer needs to be actively sought by that person. This in turn reduces the proportion of active information seeking in the total information practices. Thus the practice of the proxy shifts the information seeking from the individual to the collective. The prerequisite for this is, as already mentioned, an awareness of the information needs of colleagues and the establishment of a common practice of information sharing and information grounds (?). On the one hand, this altruism increases the effort involved, as the information needs of colleagues are also scanned or monitored. On the other hand, the invested effort reduces the effort for the own information seeking.

Colleagues are also often involved in the active seeking process. There is always a colleague who is familiar with a particular system, such as a database, and can therefore help with problems in the use of the system. To gain an understanding of the work of others, it is important to be able to follow the work of the colleagues and for this it is a basic requirement to work in the same domain of classical philology or even at the same institute and to establish a common ground (?). This has the advantage of being in physical as well as cognitive proximity to colleagues. Thus, problems such as not being able to formulate certain facts due to a lack of domain knowledge (?, p. 616) recede into the background. In direct exchange with someone from the same domain, who therefore also speaks the same language, no translation problems occur and they avoid a struggle with terminology (?, p. 104). Participants also attribute to colleagues the necessary authority (?) to act as a reliable source for certain topics (e.g. using the TLG database). The results support the findings of ?, who found that journalism professors prefer peers as information sources and a face-to-face setting. However, the procedure is exhausted at the margins of the institute or the domain of classical philology, because only information that is already

in the circle of colleagues or within this domain can circulate via sharing. And especially information about software and its application is found relatively little among the studied scholars of classical philology.

The relationship between seeking and sharing has already been studied in this regard. ?, for example, has formulated three different approaches (indirect, sequential, interactive) with which he conceptualizes the relationship between seeking and sharing. For the present study it is worthwhile to take a look at the interactive approach, in which seeking and sharing are presented as activities that occur in varying temporal order. In this way, the dynamics of the cycles of seeking and sharing can be depicted. The group of scholars examined as a whole and their information seeking can be summarized well under this approach. The sequential approach assumes that information seeking precedes information sharing. An information has to be detected first to be able to share it afterwards. It is also possible to share only what was searched for and found before. This again shows the limits of the information seeking of the examined participants: if seeking for software is stopped early or if information is mainly searched for within the circle of colleagues, then there is relatively few information about software that could be shared with colleagues.

The frequent recourse to colleagues sometimes leads to what ?, p. 57 calls small-world related barriers, a socio-cultural barrier of information seeking. Here, social norms define what is to be regarded as acceptable and unacceptable behavior. Small-world communities dissociate themselves from certain sources of information and use among other things frequently proxy strategies of information seeking (?, p. 57), as this can be observed also in the present case. With respect to software – and I can only report on this here – the limits of this small world of the classical philologists under investigation become apparent. However, it is a somewhat fragile assumption to regard the situation at the institute as a small world, although there are first indications for that.

Friends as Information Source

When it comes to problems in using and seeking software, the classical philologists studied rely on a group of people in whom they also have special trust, namely friends. ?, p. 85 aptly put it this way: "*Relying heavily on informal collaborative relationships, scholars manage their interdisciplinary information work by developing strategies for exploring and translating information from unfamiliar domains.*" The situation found in seeking software by classical philologists can certainly be described as interdisciplinary information work. Friendship can be also regarded as an informal collaborative relationship. Friends are included in the information seeking practices of the participants, both in active seeking and in information by proxy. ? studied interdisciplinary humanities scholars who, in the course of their interdisciplinary work, have made connections beyond their own original discipline. Interestingly, their boundary-crossing information work does not contain any references to friends outside of their circle of colleagues who bring along certain domain knowledge. In the present study, no interdisciplinary humanities scholars were examined, but rather

classical philologists without greater or longer interdisciplinary experience. Where the experience was nevertheless made, in some DH projects, communication within the project was still additionally forwarded to the other domain via a proxy. Without this interdisciplinarity, the classical philologists studied lack the trustworthy source within the other domain during their boundary work ?. Instead, they fall back on trustworthy sources within their circle of friends that can help them in certain questions. Some of them stated that it is always good to know a computer scientist in their circle of friends. The results of this study suggest to consider friends in the context of boundary work in addition to the sources listed by (?).

Avoiding Information Seeking

In the present study, information avoiding could only be interpreted indirectly from the accounts of the participants. A very obvious source for information about software is the discipline of digital humanities. However, some participants are reluctant to consult this discipline or even avoid it altogether. The experiences made in previous DH projects have been cited here as decisive (see Findings in section ??) and are examined in more detail in section ??.

Missing Relevance - No Need for Scanning and Monitoring

Among the reported information practices, scanning and monitoring for information sources on software was rather rare. The explanation for this lies mainly in the subjective relevance (?) of the information sought for the participants. Since the need to solve certain software problems is not the top priority, but rather literature research and writing of the thesis, the topic does not have the necessary relevance to trigger a further scanning after a failed active seeking or even to include the topic in information monitoring. The relevance of an information need influences the activities scanning and monitoring significantly. In addition, accounts of scanning and monitoring are somewhat more difficult to recall by the participants or to bring out during the interview situation. The dominant idea among the participants regarding information seeking is still active seeking and less, the very broadly defined information activities emphasized by the approach of information practices, like especially proxy but also activities of monitoring and scanning.

5.2 Context and Domain Specific Factors

At this point, three context- and domain-specific (?) factors of the found information practices are discussed, which are assumed to influence the practices of the participants: trust, negative experiences with digital humanities projects and the cognitive barriers already mentioned in the previous chapter.

Digital Humanities and Sceptical Humanists

As already indicated in the findings (section ??), some philologists have reported negative experiences in DH projects. Either they themselves were involved in DH projects as assistants or they talked about the results of projects they knew about, which in their opinion were rarely used sustainably. The experience of working together within DH projects was characterized by the fact that different research cultures met. On the one hand the classical philologists, who prefer a very precise and complete approach to texts, especially in the area of textual criticism. On the other hand, computer scientists, who are satisfied if their algorithms process the majority of the available data set correctly. For philologists, however, it is important that exactly this special case is represented. The interest in the collaboration was precisely this, which repeatedly leads to misunderstandings in the pursuit of the project goals. Another insight is that computer scientists have a very strong self-confidence in the eyes of philologists regarding their own problem-solving abilities. In the eyes of philologists this sometimes leads to the fact that computer scientists dismiss complicated problems as something simple for them, thus making the necessary understanding of the problems more difficult. The problem was already reported by (?, p. 279f) with regard to the differences in formalization approaches of both disciplines: "*Experience shows that rigid prescriptive heuristics and mandatory explicit formalisations, if uncritically imposed from a computational paradigm, generally do not 'land well' in different humanities research communities.*". Same criticism was raised by ?, p. 198 who attested that the DH or modern technologies in general have a tendency to uncritically adopt neo-liberal approaches in such that "*humanities and digital technology do not entirely point in the same direction.*".

The often lacking sustainability of DH projects was also noted. Thus, various authors conclude that the average lifespan of digital project results is often quite short (e.g. ??). The authors also provide reasons for these developments which are based on a lack of institutional support, sustainable funding, and personnel support. The results of my study also point in this direction. Many philologists have had experience with project websites or databases that were either not publicly accessible after the end of the project or were no longer used, maintained and developed. Based on these experiences, it is hardly surprising that classical philologists have developed a strong reservation towards digital humanities in general. In the present case, there was also the additional fact that there was too little, and if so, no seamless cooperation with the DH chairs at the university. All in all, the negative experiences made have a direct or indirect effect on information seeking practices with regard to software. As a result, the DH as a potential source of information is less often accessed. Furthermore, the developing tool landscape is viewed rather skeptically. The usefulness of the tools is often doubted, which is why some philologists may not even start seeking for software or even stop seeking early on, even though the relevance for a tool and the search skills may be present. The negative experiences might lead to the assumption that there will be no tools for a certain use case. On

the other hand, the degree to which the experiences gained influence one's own seeking practices is probably only qualitatively and certainly very difficult to assess.

Trust - a Highly Esteemed Value Among Classical Philologists

During the investigation, the topic of 'trust' repeatedly came into the focus of argumentation. The classical philologists work very conscientiously. Trust plays an enormously important role in their research process in many respects. Textual criticism, for example, is a sub-discipline of classical philology (see section ??), which attempts to prove certain aspects particularly meticulously, but above all the authenticity of handed-down texts. A fundamental attitude towards the texts is to mistrust them and only by working with the exact sources can one slowly approach a secure interpretation of the text and thus slowly trust the text. A second time, trust appears as a motive in the negative experiences with the digital humanities just mentioned. These have led to the fact that the participants do not particularly trust the DH as a source of information. Trust also plays a very important role in the preferred information sources. The information seeking practices of the scholars include above all their trusted colleagues, but also other confidants, the friends, in their information practices as a source of information. Finally, trust is also evident in the use of certain databases, as certain databases (such as the TLG) are considered standard tools. This is accompanied by a high level of trust in the tool and its data. The high value of trust in certain sources leads me to the assumption that the working methods of classical philologists as well as their experiences with the DH lead to the fact that the observed information seeking takes the form I have just described.

Trust was also addressed by ?, p. 266 in their study on the tool search of journalism professors: *"It is likely that information authored by journalism professors, media professionals, and communities made up of these individuals is more sought after because it is perceived by journalism professors to be more trustworthy and relevant to their needs than information authored by organizations."* The professors investigated trust more of their peers over other sources. The same was also found out about face-to-face situations, which were considered to be more trustworthy than other, mainly digital sources. Even though Hepworth et al. do not speak for classical philologists, the results are quite similar to what I found out in the realm of this study. In a study about transdisciplinarity and digital humanities ?, p. 305 report a lack of trust and limited understanding between domain users and digital humanists, too.

Closely linked to the topic of trust is the concept of cognitive authority. In his seminal work on cognitive authority ? elaborated on aspects such as reliability, validity, and credibility of authorities and how cognitive authority emerges. It is precisely this emergence of cognitive authority that the philologists examined in the study failed to make in terms of digital humanities. Studies on cognitive authority certainly offer further clues as to why. In summary, trust is a very powerful concept, which could only be touched upon superficially in this study. The results, however,

suggest that the concept of 'trust' in the work of classical philologists should be examined more closely. Trust and its effects on the work and research process of classical philologists also influence their information seeking practices.

Software Seeking Humanists - Barriers and Missing Experience

A third domain-specific factor that significantly influences information seeking practices is the lack of experience in software seeking and the cognitive barriers (?) mentioned above (see section ??). The classical philologists work in a field where very little software is used in the research process (?, p. 808). As a result, they lack experience in software seeking, since such opportunities occur rarely. Moreover, the cognitive barriers that occur during information seeking have already been mentioned in the previous section. These findings can certainly be transferred to other sub-disciplines in the humanities. Whether and to what extent experience is missing in software searching or cognitive barriers are present certainly differs from individual to individual, but also from research discipline to research discipline. I assume that due to regular and more intensive work with software in disciplines like natural sciences or engineering, more experience and less cognitive barriers are encountered. Thus, the existing barriers and lack of experience may be considered a special, domain-specific factor of classical philology.

5.3 Practical Implications

Three different practical implications can be derived from the results of the study: a call for the disentangling of data and tools, a call for more methodological reflection in classical philology about processes and tools, and the sharpening of the role of intermediates such as librarians by providing access to research software.

First, in the findings in section ?? I have described how strongly access to data dominates over potential tool functionality. The access to the data (explained here using the examples of TLG and LLTA) is so restrictive in some databases that a use outside the provided interface is not allowed. This means that classical philologists are bound to the functionality provided by the respective interface or database. This practically prevents the data from being used by other tools with possibly interesting functionality.

Motivated by this insight, I would like to formulate the demand at this point to treat data and software as separately as possible and to enable the further development of open and innovative tools for classical philology through free access to the data. This demand is based, among others, on a paper by ?, who formulated similar demands on the relationship between data and software in corpus linguistics: "*it is essential that the researcher separates the actual data from the appearance of that data as seen through the observation tool.*" (?, p. 143). The requirement itself is not new and is one of the basic principles of software engineering (?), where it serves primarily to avoid complex dependencies between data and its representation. In order to avoid

the innovation backlog of proprietary systems and to accommodate a developing digital humanities, classical philology as well as the humanities in general should work on generating open data and thus enable a flourishing tool landscape.

Second, in section ?? I have shown that there may be weaknesses in the reception and reflection of one's own methodology at the institute under study or within classical philology in general. The findings have shown that institutions such as a separate chapter on methodology in theses, the research design in academic essays or seminars on methodology in classical philology are very unusual and would probably be considered revolutionary if any of these proposals were to be implemented. For classical philology in the German-speaking world, ?, p. 247 has above all identified methodology as a major weakness and formulated the following postulate: *"Im besonderen sollten sie [RG: die Forschenden in der klassischen Philologie] in der Lage sein, den Gegenstand ihrer Untersuchung in einer Weise zu perspektivieren, die es den Rezipienten ihrer Wissenschaft erlaubt, das Prozedere der Untersuchung methodisch nachzuvollziehen."* In essence, translated it means that researchers in classical philology should disclose their methodology in order to make it comprehensible to others.

I would like to agree with Schwindt's demand, both in general and in particular with regard to the reflection of one's own research tools. The discipline would certainly benefit from initiating a more intensive study of its own research process and the methods and tools used in it. This would also facilitate the formulation of its own information needs regarding software, if its application can be derived methodically. If the own research process is penetrated, the researcher can better describe the exact functionality of the searched software and thus explicate his information need (?). This is also the key to understanding the information seeking of the classical philologists studied, because *"information seeking behaviour cannot be explained by factors that lie outside the researcher and the researcher's subject matter."* (? , p. 200). The weaknesses in the reflection of the discipline's own research process can be considered as cognitive barriers (?) and, hence, also affect information seeking. As ? aptly analyzed, the user's information need evolves with the task itself while the user accumulates further knowledge over time.

And third, the role of information intermediaries, especially librarians, in the transfer of skills related to research software will be addressed. Already in 1978 ? formulated that the librarian should participate in the research process. This demand was raised constantly over the course of time (???????), especially with regard to the changing role of the librarian in the course of digitization. A few years ago, it was the research data management (?), which came into focus at universities and scientific libraries and was supposed to ensure good scientific practice in handling data. With regard to research software, scientific institutions now face a similar task. As professional intermediaries (?) it is of course the task of libraries to support their clients in seeking research software. In a study on the information practices of humanities scholars, ?, p. 73 found that there is a *"lack of awareness of tools"* and a *"lack of institutional training and support"* in relation to digital resources. The present study clearly shows starting points for this intermediary role of libraries and

why they should assist in overcoming these challenges. They represent a reliable source of information and could thus help to overcome the scepticism towards digital tools among their clients (see section ??), by acting as a trustworthy source (see section ??) in the information seeking process and by collecting, systematizing and reusing information about research software.

6 Conclusion

In the context of this master thesis, I have set out to investigate the information seeking practices of classical philologists in the search for software. I decided to take an explorative approach and collected data from nine interviews with young scholars of an institute of classical philology in Germany and analyzed them with a qualitative content analysis according to ?. The philologists studied seldom search for software and prefer their own colleagues as a source of information in the active seeking process. Within the group of colleagues, the information needs of others are also taken into account; the colleagues often act as a proxy. Active seeking via information systems, on the other hand, is often terminated early. Friends are another important source of information for the information practices of philologists. A very high value of trust in information sources and negative experiences with digital humanities have a significant influence on the information practices found. The results serve as a starting point for scientific libraries to show how researchers can be supported in their search for research software by an extended range of services.

The limitations of this work derive partly from the qualitative research approach chosen. For example, the results of this study on classical philologists at a single institute of a single university in Germany should not automatically be applied to other institutes, other disciplines within the humanities, and the humanities in general. The humanities disciplines differ greatly in their research questions, materials and methods used (?) and it can be assumed that this is similar for information seeking practices. As already noted and supported by ?, it is difficult to distinguish exactly between active seeking and proxy interactions. The dynamic nature of information seeking (?) could not be exactly reproduced by the participants in the interviews, so assumptions had to be made at some points. The statements on the low level of methodological reflection in classical philology are based on a rather thin foundation of data, since I only came across this fact in the very last interview. They therefore represent a starting point for future research questions regarding a domain analysis in classical philology.

The results of the work also provide further approaches for future research. The role of friends in information seeking could be analyzed more closely in future work, whereby a transition from information seeking in a professional context to everyday life information seeking (ELIS) can be observed. This intersection could also be examined more closely, especially since the border between work and private life is becoming increasingly blurred in the course of changing personal attitudes and new ubiquitous information technologies. Negative experiences with the digital humanities and the strong importance of trust in the information seeking process could be examined with regard to the effects and barriers for information seeking practices. Moreover,

future studies could examine other disciplines within the humanities and thus test the insights gained here for their transferability and generalizability.

7 Zusammenfassung (German Summary)

Die Nutzung von Forschungssoftware gehört für die meisten Forschenden mittlerweile zum wissenschaftlichen Alltag. Während des gesamten Forschungsprozesses verwenden Forschende Softwarewerkzeuge für die Datenerfassung, -transformation, -analyse und -präsentation sowie für die Generierung von Hypothesen, die Literaturverwaltung und das Verfassen wissenschaftlicher Arbeiten. Software hat die Art und Weise verändert, wie wir Wissenschaft betreiben. Die Komplexität der von Forschenden durchgeführten Analysen hat zugenommen, ebenso wie die Menge der Daten, die Forschende verarbeiten können. Software unterstützt die Dokumentation des Forschungsprozesses und ermöglicht die Reproduzierbarkeit und Genauigkeit der Ergebnisse. Aufgrund der gestiegenen Bedeutung von Forschungssoftware für die Forschung und der Zunahme der Anzahl genutzter Software ist es für Forschende umso wichtiger, geeignete Software zu identifizieren und diejenige auszuwählen, die am besten zu dem Forschungsproblem, dem jeweiligen Schritt im Forschungsprozess oder den zu verarbeitenden Forschungsdaten passt und folglich dem Informationsbedarf (information need) der Forschenden genügt. Neben dem erhöhten Aufwand können Schwierigkeiten bei der Suche nach Software auch die wissenschaftliche Reproduzierbarkeit von Studien gefährden oder zu Mehrfachentwicklungen von Software mit identischen Funktionen führen, anstatt bereits bestehende Software wiederzuverwenden.

Forschungen zum Prozess der Informationssuche (information seeking) sowie dessen Einflussfaktoren und Barrieren liefern wichtige Erkenntnisse darüber, wie die Informationssuche Forschender sowie die dafür nötige technische und organisatorische Infrastruktur effizienter gestaltet werden können. In der vorliegenden qualitativen Studie wurden die Informationspraktiken (information practices) von klassischen Philolog*innen bei der Suche nach Forschungssoftware untersucht. Es wurden dafür neun Einzelinterviews mit fünf Nachwuchswissenschaftler*innen eines Instituts für klassische Philologie geführt, welche im Anschluss transkribiert, offen codiert und danach unter Verwendung der qualitativen Inhaltsanalyse nach ? ausgewertet wurden.

Die untersuchten Philolog*innen suchen selten nach Software und bevorzugen die eigenen Kolleg*innen als Informationsquelle bei der aktiven Suche. Innerhalb des Kolleg*innenkreises werden auch die Informationsbedarfe der anderen berücksichtigt, die Kolleg*innen agieren häufig als Proxy. Die aktive Suche über Informationssysteme wird dagegen häufig aufgrund fehlender Relevanz für den eigenen Forschungsprozess

und aufgrund kognitiver Barrieren frühzeitig abgebrochen. Freunde stellen eine weitere wichtige Informationsquelle für die Informationspraktiken der Philolog*innen dar. Ein sehr hoher Stellenwert von Vertrauen in Informationsquellen und negative Erfahrungen mit den Digital Humanities beeinflussen die vorgefundenen Informationspraktiken wesentlich.

Die Ergebnisse dienen als Ausgangspunkt für Wissenschaftliche Bibliotheken, wie Forschende bei der Suche nach Forschungssoftware durch ein erweitertes Dienstleistungsangebot unterstützt werden können. Sie liefert außerdem einen Anstoß für die Geisteswissenschaften, die eigene Methodologie und die eigenen Methoden vor allem hinsichtlich der Entwicklungen in den Digital Humanities stärker zu reflektieren. Vor allem aber vermitteln die Ergebnisse einen ersten Einblick in die Informationspraktiken von klassischen Philologen bei der Suche nach Forschungssoftware und liefern Anhaltspunkte für weitere Forschungsvorhaben.

Acknowledgements

This work was made possible by the voluntary work of many people who were involved in the development of the following open source software, which I was thankfully allowed to use in the creation of this work: TeXMaker¹, JabRef², L^AT_EX³, Atom⁴, LibreOffice⁵, NextCloud⁶, git⁷, okular⁸, ReColl⁹, pandoc¹⁰, QualCoder¹¹, EasyTranscript¹², Voice Recorder¹³, and not to forget Arch Linux¹⁴.

¹<http://www.xmlmath.net/texmaker/>

²<https://www.jabref.org>

³<https://www.latex-project.org/>

⁴<https://atom.io/>

⁵<https://www.libreoffice.org>

⁶<https://www.nextcloud.org>

⁷<https://git-scm.com/>

⁸<https://okular.kde.org/>

⁹<https://www.lesbonscomptes.com/recoll/>

¹⁰<https://pandoc.org/>

¹¹<https://github.com/ccbogel/QualCoder>

¹²<https://www.e-werkzeug.eu/index.php/de/produkte/easytranscript>

¹³<https://www.simplmobiletools.com/>

¹⁴<http://www.archlinux.org>

Glossary of Mentioned Software and Services

L^AT_EX <https://www.latex-project.org/>.

Academia <https://www.academia.edu/>.

AntConc <https://www.laurenceanthony.net/software/antconc>.

Bryn Mawr Classical Review <http://bmcr.brynmawr.edu/>.

Catullus online <http://www.catullusonline.org>.

Citavi A reference management software. <https://www.citavi.com/en>.

DeepL <https://www.deepl.com>.

Dropbox <https://www.dropbox.com>.

Excel (Microsoft) <https://products.office.com/excel>.

Germanistik.net <https://www.germanistik-im-netz.de/>.

Gnomon <https://www.chbeck.de/buecher/zeitschriften/gnomon/>.

Google Scholar <https://scholar.google.com>.

HSozKult <https://www.hsozkult.de/>.

JSTOR <https://www.jstor.org/>.

Library of Latin Texts <https://about.brepolis.net/library-of-latin-texts/>.

Liddel-Scott-Jones <https://lsj.gr>.

Mozilla Thunderbird <https://www.thunderbird.net>.

Perseus <http://www.perseus.tufts.edu/>.

Stylo <https://github.com/computationalstylistics/stylo>.

Thesaurus Linguae Graecae <https://stephanus.tlg.uci.edu>.

Thesaurus Linguae Latinae <https://db.degruyter.com/databasecontent?dbid=tll&dbsource=%2Fdb%2Ftll&language=en>.

Word (Microsoft) <https://products.office.com/word>.

Appendices

A Interview Guide

Interview questionnaire

Personal questions

- Study, doctorate
- Work experience, current position
- Field of expertise: Contents, methods, theories

A Entry questions

Search in general

- Describe the process of your own literature research
 - Procedure Which sources are used? (search engines, databases, colleagues, publications, events, service providers, social media)
 - Which tools are used?

Software usage

- Which software do you use in your everyday scientific work?
 - Since when?
 - Motivation - For what?
 - Experience?
- Which software was last integrated into your own research process?
 - When?
 - Motivation - For what?
 - What was the impulse for it?
 - Experience by now?
 - What has changed as a result?
 - satisfaction?

B The search for software - sources

- Describe the process of searching for software tools
- Which sources are used? (Friends, Colleagues, Within the institution: working group/professorship/faculty, Outside the institution: national vs. international, subject vs. non-subject, special catalogues, search engines, service providers: library, subject information services, subject specific websites, social media, software repositories)
- Which tools are used? (special catalogues, software repositories, social media, search engines)
- What are the hurdles?

C Search and use of software - information practices

Seeking

- What kind of software have you already been looking for for your work?
- What sources have you ever consulted when searching for software? (once/regularly)
- What sources have you identified that could help you in your search but do not have it yet?
- How do you proceed with the active search? (strategies, lists, preparation for talks)

Scanning

- Where do you come across references to used software in your everyday work (journals, conferences, meetings, students...)?
- Which relevant source have you already come across and how did you deal with it?
- If applicable, how do you observe other researchers during their work (looking over your shoulder while talking, explanations)

Monitoring

- What sources of information (newsletters, websites, newspapers, social media) do you regularly monitor?
- In which everyday situations could you encounter information about interesting software or hints for its use?
 - Within your workplace?
 - Outside?
 - If difficult: using literature as an example

Proxy

- What is the role of information providers? (Connecting)
 - Someone with the necessary knowledge about software identifies you as an information seeker (once/regularly)?
 - Someone refers you to a potential information source?
 - Someone makes recommendations (once/regularly)?
- How does this someone interact with you? (interacting) Informing through stories or experience
 - Advising, instructing, guiding

Avoiding

- What sources do you trust?
 - Why?
- Which sources would you not consult for literature/software?
 - Why?

- What are the alternatives to a source (database, website, search engine etc)?

Context

- What role does context play in the search for software?
 - Field: Classical Philology
 - Institute, Colleagues
 - Personal life situation: family, friends, leisure groups

D Follow-up questions for the second interview

Follow-up first interview (interval: approx. 2 weeks)

- What topics have you been thinking about until today?
- What do you want to add to the last interview?
- Depending on the context: (Add at the point)
 - ...

B Category System

Codebuch für masters_thesis.qda

=====

_Thesen

- Antriggern von Kollege/Freund als Proxy
- Die Grenze meiner Sprache ist die Grenze meiner Welt.
- Digitales wird nicht gern zitiert
- Geringe Methodenreflektion
- Geringere Prio => SW Suche frühzeitiger Abbruch
- Kollegen sind die Primärquelle bei Schwierigem
- kP war in der Vergangenheit teilweise weiter als heute
- kP-Kollegen begeistern mit DH Projekten
- Neg Erf mit DH => Reserviertheit ggüber Proz und Werkz
- Technologie hilft beim Präzisieren des InfoNeeds
- Wenige

Feld - Klassische Philologie

--Eigenschaften

---- analog

----- genau

----- IT-fern

----- kollegial

----- Konservativ

----- kritisch

----- vorsichtig

----- ästhetisch

----Methoden

---- Bildung

----- Corpus

----- Kommentar

----- Methodik

----- Stilometrie

----- Textkritik

----Teilbereiche

---- Editionswissenschaften

----- Literaturwissenschaften

----- Paläographie

----- Philosophie

----- Translatologie

----Werkzeuge

----Literaturrecherche

----- Academia

----- allgemein

----- BMCR

----- BTL

----- Catul

----- DeepL

----- Facebook

----- FID

----- Franzosius

----- germanistik.net

----- Gnomon

----- Google [Scholar]

----- JSTOR

----- l'année philologique

----- LLTA

----- LSJ

----- Michael Erler

----- Pauly

----- Perseus

----- Prepolis

----- Rheinisches Museum

----- Schmidt-Herzog

----- TLG

----- TLL

-----Software

----- Citavi

----- Dropbox

----- Excel

----- Latex

----- Thunderbird

- typegreek
- Word
- Digital Humanities
- Publizieren
- Unterschied zu and. Fachbereichen
- Verbesserung
- Vergangenheit / Zukunft
- Information Practices
- Kanal
- Browser [www]
- Chip.de
- Email
- face-to-face
- Software selber
- Erfolg
- Ignoring
- Information Need
- IP_allgemein
- Kontext
- Literatursuche
- Monitoring
- Proxy
- Scanning
- Schneeballprinzip
- Seeking
- Technology
- Ungeduldig
- Unsicher
- Vorbereitung
- Zufall
- Institut/Uni
- Bibliothek
- Forschungsunterstützungsinfrastruktur
- Inhalte
- Kollegen
- Mentoring/Stiftung
- Oberseminar
- Studierende
- Person
- Berufserfahrung
- Projekt
- SFB
- Studium
- Finanzierung
- Freunde / Familie
- Frustration
- Vertrauen

C Consent Form

Einverständniserklärung zum Interview

Forschungsprojekt: Masterarbeit „Seeking Research Software. A Qualitative Study of Humanities Scholars’ Information Practices“

Institution: HU Berlin

Masterand/Interviewer: Ronny Gey

Interviewdatum: xx.xx.2020

Nachname, Vorname (Proband*in):

Ich erkläre mich dazu bereit, im Rahmen des genannten Forschungsprojekts an einem Interview teilzunehmen. Ich wurde über das Ziel und den Verlauf des Forschungsprojekts informiert. Ich kann das Interview jederzeit abbrechen und meine Einwilligung in eine Aufzeichnung und Niederschrift des Interviews zurückziehen, ohne dass mir dadurch irgendwelche Nachteile entstehen. Ich bin damit einverstanden, dass das Interview mit einem Aufnahmegerät aufgezeichnet und anschließend vom Interviewer in Schriftform gebracht wird. Für die weitere wissenschaftliche Auswertung des Interviewtextes werden alle Angaben, die zu einer Identifizierung der Person führen könnten, verändert oder aus dem Text entfernt. In wissenschaftlichen Veröffentlichungen werden Interviews nur in Ausschnitten zitiert, um gegenüber Dritten sicherzustellen, dass der entstehende Gesamtzusammenhang von Ereignissen nicht zu einer Identifizierung der Person führen kann. Personenbezogene Kontaktdaten werden nicht gespeichert.

Ort, Unterschrift

Ich bin damit einverstanden, dass das anonymisierte Interviewtranskript unter einer Creative Commons CC BY-NC-SA 4.0 Lizenz steht und unter https://github.com/geyslein/Masters_Thesis zur besseren Nachvollziehbarkeit der Forschungsergebnisse und zur Weiternutzung durch Dritte verfügbar ist.

Ort, Unterschrift



Name: Gey Vorname: Ronny

Matr.Nr.: 599655

Eidesstattliche Erklärung zur

- ☐ Hausarbeit *
- ☐ Bachelorarbeit *
- ☒ Masterarbeit *
- ☐ Abschlussarbeit im Bibliotheksreferendariat *

* Die eingereichte PDF-Datei ist mit den Printexemplaren identisch.

Ich erkläre ausdrücklich, dass es sich bei der von mir eingereichten schriftlichen Arbeit mit dem Titel

Seeking Research Software. A Qualitative Study of Humanities Scholars' Information Practices.

um eine von mir erstmalig, selbstständig und ohne fremde Hilfe verfasste Arbeit handelt.

Ich erkläre ausdrücklich, dass ich *sämtliche* in der oben genannten Arbeit verwendeten fremden Quellen, auch aus dem Internet (einschließlich Tabellen, Grafiken u. Ä.) als solche kenntlich gemacht habe. Insbesondere bestätige ich, dass ich ausnahmslos sowohl bei wörtlich übernommenen Aussagen bzw. unverändert übernommenen Tabellen, Grafiken u. Ä. (Zitaten) als auch bei in eigenen Worten wiedergegebenen Aussagen bzw. von mir abgewandelten Tabellen, Grafiken u. Ä. anderer Autorinnen und Autoren (Paraphrasen) die Quelle angegeben habe.

Mir ist bewusst, dass Verstöße gegen die Grundsätze der Selbstständigkeit als Täuschung betrachtet und entsprechend der Prüfungsordnung und/oder der Fächerübergreifenden Satzung zur Regelung von Zulassung, Studium und Prüfung (ZSP-HU) geahndet werden.

Datum 20.10.2020

Unterschrift Ronny Gey