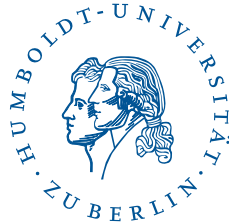


HUMBOLDT UNIVERSITY OF BERLIN

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, OCTOBER 8, 2020

Zusammenfassung

Zusammenfassung

Abstract

Abstract

Contents

List of Tables	iii
Acronyms	iv
1. Introduction	1
2. Theory	3
2.1. Information Seeking	3
2.1.1. Introduction	3
2.1.2. Information Seeking Research	3
2.1.3. Information Behavior	3
2.2. Information Practices	4
2.2.1. Introduction	4
2.2.2. Studies of Information Practices	4
2.3. Research Software	6
2.4. The Humanities	8
3. Research Design	9
3.1. Methods of Data Collection and Processing	10
3.2. Process of Data Analysis	12
4. Findings	14
4.1. Introduction of the Empirical Field	14
4.2. Software Use	15
4.3. Information Practices Towards Software	17
4.3.1. Introduction	17
4.3.2. Seeking	18
4.3.3. Scanning	21
4.3.4. Monitoring	21
4.3.5. Proxy	22
4.4. Domain Factors	24
4.4.1. Negative Experiences with Digital Humanities Projects	25
4.4.2. Information Seeking and the Limits of Language	27
4.4.3. Methodological deficits in classical philology?	28
4.4.4. Trusting the Source	29
4.4.5. Software and Data	30

5. Discussion	32
5.1. Theoretical Implications	32
5.2. Practical Implications	32
6. Conclusion	33
7. Zusammenfassung (German conclusion)	34
Bibliography	35
Glossary of Mentioned Software and Services	48
Appendix A. Interview Guide	51
Appendix B. Category System	55
Appendix C. Consent Form	56

List of Tables

1.	Transcription conventions	12
2.	Overview of interviewees	14

Acronyms

APh l'Année philologique.

BMCR Bryn Mawr Classical Review.

DFG Deutsche Forschungsgemeinschaft.

DH digital humanities.

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 (Kethers et al., 2017; Pan et al., 2016; Wolski et al., 2017). 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 (Dallmeier-Tiessen, 2016; Waltemath and Wolkenhauer, 2016) and accuracy of results.

Due to the increased importance of research software for research (Katz, 2017) 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 (Wilson, 1994). 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 (Hucka and Graham, 2018).

Information seeking of researchers is generally of great interest within the field of information, be it information behavior (Ahmadianyazdi and Chandrashekara, 2018; Barrett, 2005; Campbell, 2017; Catalano, 2013; Ellis, 1993; Hemminger et al., 2007; Korobili et al., 2011; Liyana and Noorhidawati, 2017; Rimmer et al., 2006; Rupp-Serrano and Robbins, 2013; Wang et al., 2008, e.g.) or information practices (Bøyum and Aabø, 2015; Bulger et al., 2011; Fry, 2006; Given and Willson, 2018; Roos, 2015, e.g.). However, seeking software is still rather challenging for researchers (Howison and Bullard, 2015). In a recent study, Hucka and Graham (2018) 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 as Bauer et al. (2014) revealed in a study on code reuse at Google. Grossman et al. (2009) identified users unawareness of specific software tools. These results are all the more surprising because the participants in the cited studies come from a group with a greater affinity for software (software developers, engineers).

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 (Struck, 2018). Standards and tools for citing software enable re-

searchers to identify, cite and reuse software (Niemeyer et al., 2016; Smith et al., 2016; Soito and Hwang, 2017, e.g.). Research funding agencies and research organizations (Haupt et al., 2018; Katerbow and Feulner, 2018; Scheliga et al., 2019, 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. For researchers with less experience in the use of software, e.g. from the humanities (Rimmer et al., 2006), seeking software remains a difficult undertaking.

The information-seeking behavior of humanities scholars in general has been addressed widely (Barrett, 2005; Bronstein, 2007; Bronstein and Baruchson-Arbib, 2007; Catalano, 2013; Ellis, 1993; Given and Willson, 2018; Korobili et al., 2011; Liew and Ng, 2006; Rimmer et al., 2006, e.g.). In his pioneering work on Grounded Theory in information-seeking, Ellis (1993) identified patterns of information-seeking of social sciences, sciences, and humanities scholars. In 2005, Barrett (2005) analyzed information-seeking habits of graduate student researchers in the humanities. Korobili2011 examined factors influencing information-seeking behavior at the philosophy faculties. While studies in information behavior draw on the cognitive viewpoint, information practice studies are guided by the ideas of social constructionism and collectivism (Savolainen, 2007; Talja et al., 2005; Talja and McKenzie, 2007). Humanities scholars information-seeking practices have also been addressed in several studies (Benardou et al., 2013; Bulger et al., 2011; Given and Willson, 2018; Palmer and Cragin, 2009). 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 (Barrett, 2005; Rimmer et al., 2006), the search for software itself is rarely discussed. One of these rare examples, however a non-humanities one, is Hepworths et al. (2017) study of journalism professors' information-seeking behavior. While seeking new online tools, journalism professors rely on other journalism professors, followed closely by media-related foundations, media professionals, and conferences.

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. The data were transcribed and analyzed with the qualitative content analysis according to Mayring (2014). In preparation of the study and to derive the research gap, I will present the theoretical concepts of this work in the next section. I discuss the concepts information seeking (2.1), information practices (2.2), research software (2.3) and the field of humanities (2.4). Afterwards I present the research design of the thesis in section (3). The subsequent presentation of the results is divided into two parts. I will describe the information practices of the subjects (section 4.3) and then interpret some of these descriptions in section 4.4 on the basis of the subjects' experiences in classical philology. In the following discussion (section 5), I will integrate the findings from the empirical study into the current theory. The conclusion section (6) summarizes the essential aspects of the thesis.

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. This is accompanied by an interest in understanding how we collect, create and organize information, but also how we search this information. Within information science, we distinguish between seeking, searching and retrieving information.

- Information Science briefly described - Seeking, Searching, Retrieval - Definition ISeeking

2.1.2. Information Seeking Research

- (Ingwersen2005) – Concepts: Strategies – Collaborative IS: Shah2013

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 (Savolainen, 2007). I will describe the concept of information seeking behavior in section 2.1.3. 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 2.2.

2.1.3. Information Behavior

- Bates2010 - information behaviour, Case2007 - information behaviour - models of IB: (Wilson, 1994) (Ellis, 1993), (Kuhlthau, 1991), (Niedzwiedzka, 2003) - (Fisher and Julien, 2009) - zielgruppen: jugendliche, akademiker, elis, ... – different conceptualizations: intra/inter/extrapersonal (Feinman – transgender ib: pohjanen2016

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 (Savolainen, 2007). According to (Araújo, 2019, p. 102), the concept of "practices" can be traced back to the work of Bourdieu and Pratique (1994) 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 (Araújo, 2019, p. 102). (Savolainen, 2007, p. 120), on the other hand, sees the theoretical roots of the information practice approach in the work of Anthony Giddens (1984) 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' 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 (Talja, 2005). Palmer and Cragin (2009) 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 (Savolainen, 2007, p. 120). Rather than focusing on individuals' behavior, actions, skills, ideas and motives, practices regard the interactions between members of a community (Tuominen et al., 2005). As Talja et al. (2005, 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.

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 Savolainen (2007, p. 121), with her constructionist approach to information practices, Pamela McKenzie (2002, 2003a,b) is one of the most influential scholars of information practice. In her empirical studies, she focuses on a constructionist discourse analytic approach in practioner-patient information seeking. McKenzie (2002) 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, (McKenzie, 2003a) 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, McKenzie (2003b) 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 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 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. Talja and Maula (2003) 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 Fry (2006) 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. Roos (2015) again examined medical scientists during their research work. Olsson (2007) 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. (Bøyum and Aabø, 2015) and (Ahenkorah-Marfo, 2016), 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. Bonner and Lloyd, 2011; Johannisson and Sundin, 2007; Mohammad, 2020). In their discourse-oriented information seeking study, Johannisson and Sundin (2007) analyse nurses and the Swedish nursing profession. In further studies, Olsson (2010, 2016) 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. (Lloyd, 2007) 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 (Veinot, 2007).

In addition to professional groups, there are other groups that appear in studies on information practices. Mainly, they are classified according to the personal life situation of the test persons. Greyson (2017, 2018) for example, examined the health information practices of young Canadian parents between the ages of 16 and 23 in his studies.

(Caidi et al., 2010) - Immigrants (Harlan, 2012) - Teen content creators (Kitzie et al., 2020) - LGBTQ (Lloyd et al., 2013) - Refugees (McKenzie, 2002, 2003a,b) - Pregnant Women

types of ip: seeking, connecting by proxy

Methods of the studies - Datenerhebung - Datenanalyse

Savolainen2007, Talja2007

The social constructionist paradigm puts emphasis on social practices, “the concrete and situated activities of interacting people, reproduced in routine social contexts across time and space” (Rosenbaum, 1993, p. 239). A focus on practices rather than on behaviour shifts the analysis from cognitive to social and is consistent with the study of information seekers within their social context (for examples, see Rothbauer (2002), McKenzie and Davies (2002)).

- and Talja

- further examples of Information Practices: Savolainen 2007 - LitReview

Health Care: (Kitzie et al., 2020) Immigrants: (Caidi et al., 2010)

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 (Kethers et al., 2017; Pan et al., 2016; Wolski et al., 2017). Following the understanding of the Deutsche Forschungsgemeinschaft (DFG) (Katerbow and Feulner, 2018) and the Helmholtz Association (Scheliga et al., 2019) 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 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. Allen et al., 2020; Cosmo et al., 2020; Druskat, 2020; Katz et al., 2020; Nielsen and Van De Sandt, 2019; Niemeyer et al., 2016; Soito and Hwang, 2017). 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. Howison and Bullard, 2015; Hucka and Graham, 2018; Murphy-Hill et al., 2015; Struck, 2018). 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. Anzt et al., 2020; Baxter et al., 2012; Brett et al., 2017; Jay et al., 2016; Katz et al., 2018; Richardson, 2017). 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. Imran and Kosar, 2019; Badreddin et al., 2019; Cohen et al., 2018; de Souza et al., 2019; Fehr et al., 2019; Katz et al., 2019; Löffler et al., 2019; Lubbock, 2019; Saputri and Lee, 2020). 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. Herraiz-Tabernero et al., 2007; Hong, 2019; Ince et al., 2012; Kisilinskiy and Goch, 2016; Oishi et al., 2018; Pianosi et al., 2019; Tennant et al., 2020) 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. Johanson and Hasselbring, 2018; Hasselbring, 2019; Hasselbring and van Hoorn, 2020; Kanewala and Bieman, 2018; Zirkelbach et al., 2020). 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: e.g. astronomy (Allen, 2019; Borne et al., 2009; Gray and Mann, 2011; Grosbol and Tody, 2010; Muna et al., 2016; Oishi et al., 2018; Shortridge, 2009;

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

Weiner et al., 2009), biology (Lee, 2018; Mangul et al., 2019; Markowetz, 2017; Waltemath and Wolkenhauer, 2016), chemistry (Álvarez-Moreno et al., 2014; Dryden et al., 2017; Goh et al., 2017; Grimme and Schreiner, 2017), humanities (Anthony, 2013; Brian et al., 2020; Edmond et al., 2017; Heyer et al., 2019; Wolski et al., 2017), medicine (Kislinskiy and Goch, 2016; White et al., 2019; Ye et al., 2019), and physics (Bobra et al., 2020; Shamir et al., 2013; Zwart, 2018).

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, Howison and Bullard (2015) found out, that seeking software is still rather challenging for researchers. Hucka and Graham (2018) 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 Bauer et al. (2014) revealed in a study on code reuse at Google. Grossman et al. (2009) 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 (Hucka and Graham, 2018, 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 (Savolainen, 2007, p. 120). The information practices within such a group can help to answer questions about the reasons for the deficits, the unawareness or the rudimentary approaches to information seeking for software.

2.4. The Humanities

- short: humanities, long: philology - definition - characteristics: subjects, work procedures, tools, ... - class. philology - Information Practices of Humanists - examples of studies, what has been studied yet - bisher nicht viel gefunden, practices of other scholars, but humanists seldom

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)" (Palmer and Cragin, 2009, 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 (Case, 2008; Cunningham, 2010). I chose an exploratory study design (Rinsdorf, 2013) 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, Lamnek (2005, 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 (Hjørland and Albrechtsen, 1995) 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 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 Deutsche Forschungsgemeinschaft (2019): 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 Lamnek (2005, 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 (Biernacki and Waldorf, 1981), 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 (Bryman, 2004) in German language. With the interviews I obtain emotions, thoughts, and intentions of the participants and discover their perspective of the social world (Patton, 2002). 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 back-

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

ground of the study. All interviews were conducted in the offices of the interviewees or in meeting rooms of the 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 (A).

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 (Edwards, 2003, 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 (Edwards, 2003, 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 non-verbal aspects. The following conventions were applied:

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

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)

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 results 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 (Kohlbacher, 2006; Krippendorff, 2012; Mayring, 2000, 2014). It enables the researcher to include textual information and to identify its properties systematically. In detail, I chose a qualitative content analysis according to Mayring (2014). 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 (Corbin and Strauss, 1990) to conceptualize and categorize the interview data. For this purpose, I have deductively identified categories from the knowledge about information practices, the field, and the process of compiling the interview guide. 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.

Mayring (2014, p. 65) describes three basic interpretation techniques: reduction, explication and 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,

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

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 (Mayring, 2014, p. 79). The final category system can be looked up in the Appendix (B).

During transcription, coding, analysis and even during the writing process, I was able to contact Interviewees for field-specific questions. The 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 4.1). Afterwards I will briefly present the situation of software use in the research process of the participants (section 4.4.5). The main part of the results will be the information practices in search for and use of software (section 4.3), as well as field- or domain-specific factors regarding problems, barriers and contradictions in the search for software (section 4.4).

4.1. Introduction of the Empirical Field

Classical Philology is one of the oldest subjects at our interviewee's 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 2.

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

Table 2.: Overview of interviewees

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

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

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>

"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 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

³<http://www.typegreek.com/>

⁴<https://stephanus.tlg.uci.edu>

⁵<https://about.brepolis.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/>

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. 49) 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 4.3.1. The very dominant mode, seeking, is then discussed in section 4.3.2. This is followed by explanations on scanning (section 4.3.3) and monitoring (section 4.3.4). 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 4.3.5) through colleagues or friends.

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*

⁹<https://scholar.google.com>

¹⁰<https://www.academia.edu>

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 speech 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.

4.3.4. Monitoring

Monitoring as an information practice seems to occur as rarely as scanning (section 4.3.3). 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:

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

"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 4.3.3) 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 4.3.2. 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 4.3:

"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 4.3.2, 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 4.3.2), 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. McKenzie (2003b, 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 Wittgenstein (1990) 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 4.3.2 and 4.3.5). 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 deficits 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 4.3.1). 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 4.4.5) and thus to a certain extent they have no choice but to trust them. In section 4.3.2 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 4.4.2, 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 4.3, 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 4.3.2). 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

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

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

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

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 and only segments and not the complete texts are available (Lang, 2018). 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.

5. Discussion

5.1. Theoretical Implications

- classical philologists do not reflect much on the research process - which is why they do not regarding software use either =, as a matter of fact, the analysis focus is on information practices in the widest sense
 - tool selection is very much trust based (recommendations/consulting from colleagues and friends/family)
 - negative experiences in digital humanities projects induce skepticism towards digital tools
- van Zundert (2012) large scale digi infrastructures as dead end Neuefeind et al. (2020)
- Sustainability Strategies for Digital Humanities Systems
 - difficulties in formulating the right search terms (what they describe as well as what I could listen to during the interviews) ==Savolainen (2015): cognitive barriers to information seeking - Inability to articulate one's information needs; Poor search skills
- Bulger et al. (2011) =, Viele Parallelen zu meinen Ergebnissen
- If the information has already reached an Interviewee via a proxy, he or she no longer needs to actively search for it.

5.2. Practical Implications

Constant et al. (1996) - kindness of strangers Edmond (2005) - role of prof intermediaries
Gunning (1978) - librarian should participate in research process Monroe-Gulick et al. (2013) - librarians as partners

6. Conclusion

7. Zusammenfassung (German conclusion)

Bibliography

- Ahenkorah-Marfo, M. (2016). Domain analytic approach to the use of academic databases by graduate students. *International Information & Library Review*, 49(1):1–10.
- Ahmadianyazdi, F. and Chandrashekara, M. (2018). Research challenges and delayed gratification in information seeking behavior: a case study of research scholars. *Library Philosophy & Practice*.
- Allen, A. (2019). Receiving Credit for Research Software. In Teuben, P. J., Pound, M. W., Thomas, B. A., and Warner, E. M., editors, *Astronomical Society of the Pacific Conference Series*, volume 523 of *Astronomical Society of the Pacific Conference Series*, page 593.
- Allen, A., Nemiroff, R., Ryan, P., Schmidt, J., and Teuben, P. (2020). Best ways to let others know how to cite your research software. In *American Astronomical Society Meeting Abstracts*, American Astronomical Society Meeting Abstracts, page 10912.
- Álvarez-Moreno, M., de Graaf, C., López, N., Maseras, F., Poblet, J. M., and Bo, C. (2014). Managing the computational chemistry big data problem: The ioChem-BD platform. *Journal of Chemical Information and Modeling*, 55(1):95–103.
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2):141–161.
- Anzt, H., Bach, F., Druskat, S., Löffler, F., Loewe, A., Renard, B. Y., Seemann, G., Struck, A., Achhammer, E., Aggarwal, P., Appel, F., Bader, M., Brusch, L., Busse, C., Chourdakis, G., Dabrowski, P. W., Ebert, P., Flemisch, B., Friedl, S., Fritzsche, B., Funk, M. D., Gast, V., Goth, F., Grad, J.-N., Hermann, S., Hohmann, F., Janosch, S., Kutra, D., Linxweiler, J., Muth, T., Peters-Kottig, W., Rack, F., Raters, F. H. C., Rave, S., Reina, G., Reißig, M., Ropinski, T., Schaarschmidt, J., Seibold, H., Thiele, J. P., Uekerman, B., Unger, S., and Weeber, R. (2020). An environment for sustainable research software in germany and beyond: Current state, open challenges, and call for action. *arXiv e-prints*.
- Araújo, C. A. Á. (2019). Information practices: the relevance of the concept to information user studies. *Annals of Library and Information Studies (ALIS)*, 66(3).
- Badreddin, O., Hamou-Lhadj, W., and Chauhan, S. (2019). Susereum: Towards a reward structure for sustainable scientific research software. In *Proceedings of the 14th*

- International Workshop on Software Engineering for Science*, SE4Science '19, pages 51–54, Piscataway, NJ, USA. IEEE Press.
- Barrett, A. (2005). The information-seeking habits of graduate student researchers in the humanities. *The Journal of Academic Librarianship*, 31(4):324–331.
- Bauer, V., Eckhardt, J., Hauptmann, B., and Klimek, M. (2014). An exploratory study on reuse at google. In *Proceedings of the 1st International Workshop on Software Engineering Research and Industrial Practices - SER&IPs 2014*. ACM Press.
- Baxter, R., Chue Hong, N., Gorissen, D., Hetherington, J., and Todorov, I. (2012). The Research Software Engineer. In *Proceedings of the Digital Research 2012*. Digital Research 2012 ; Conference date: 10-09-2012 Through 12-09-2012.
- Benardou, A., Constantopoulos, P., and Dallas, C. (2013). An approach to analyzing working practices of research communities in the humanities. *International Journal of Humanities and Arts Computing*, 7(1-2):105–127.
- Biernacki, P. and Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research*, 10(2):141–163.
- Bobra, M. G., Mumford, S. J., Hewett, R. J., Christe, S. D., Reardon, K., Savage, S., Ireland, J., Pereira, T. M. D., Chen, B., and Pérez-Suárez, D. (2020). A survey of computational tools in solar physics. *Solar Physics*, 295(4):57.
- Bonner, A. and Lloyd, A. (2011). What information counts at the moment of practice? information practices of renal nurses. *Journal of Advanced Nursing*, 67(6):1213–1221.
- Borne, K., Accomazzi, A., Bloom, J., Brunner, R., Burke, D., Butler, N., Chernoff, D. F., Connolly, B., Connolly, A., Connors, A., Cutler, C., Desai, S., Djorgovski, G., Feigelson, E., Finn, L. S., Freeman, P., Graham, M., Gray, N., Graziani, C., Guinan, E. F., Hakkila, J., Jacoby, S., Jefferys, W., Kashyap, Kelly, B. o., Knuth, K., Lamb, D. Q., Lee, H., Lored, T., Mahabal, A., Mateo, M., McCollum, B., Muench, A., Pesenson, M., Petrosian, V., Primini, F., Protopapas, P., Ptak, A., Quashnock, J., Raddick, M. J., Rocha, G., Ross, N., Rottler, L., Scargle, J., Siemiginowska, A., Song, I., Szalay, A., Tyson, J. A., Vestrand, T., Wallin, J., Wandelt, B., Wasserman, I. M., Way, M., Weinberg, M., Zezas, A., Anderes, E., Babu, J., Becla, J., Berger, J., Bickel, P. J., Clyde, M., Davidson, I., van Dyk, D., Eastman, T., Efron, B., Genovese, C., Gray, A., Jang, W., Kolaczyk, E. D., Kubica, J., Loh, J. M., Meng, X.-L., Moore, A., Morris, R., Park, T., Pike, R., Rice, J., Richards, J., Ruppert, D., Saito, N., Schafer, C., Stark, P. B., Stein, M., Sun, J., Wang, D., Wang, Z., Wasserman, L., Wegman, E. J., Willett, R., Wolpert, R., and Woodroffe, M. (2009). Astrominformatics: A 21st Century Approach to Astronomy. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010.
- Bourdieu, P. and Pratique, R. (1994). Sur la théorie de l’action. *Editions du Seuil*.

- Bøyum, I. and Aabø, S. (2015). The information practices of business PhD students. *New Library World*, 116(3/4):187–200.
- Brett, A., Croucher, M., Haines, R., Hettrick, S., Hetherington, J., Stillwell, M., and Wyatt, C. (2017). Research software engineers: State of the nation report 2017.
- Brian, M., Ciula, A., Geoffroy, N., Ginestra, F., Miguel, V., and Paul, C. (2020). The place of models and modelling in digital humanities: Intersections with a research software engineering perspective.
- Bronstein, J. (2007). The role of the research phase in information seeking behaviour of jewish studies scholars: A modification of ellis’s behavioural characteristics. *Information Research: An International Electronic Journal*, 12(3):n3.
- Bronstein, J. and Baruchson-Arbib, S. (2007). The application of cost—benefit and least effort theories in studies of information seeking behavior of humanities scholars: the case of jewish studies scholars in israel. *Journal of Information Science*, 34(2):131–144.
- Bryman, A. (2004). *Social research methods*. Oxford Univ. Press, 2. ed. edition. Includes index. - Bibliography. - Previous ed.: 2001.
- Bulger, M. E., Meyer, E. T., la Flor, G. D., Terras, M., Wyatt, S., Jirotko, M., Eccles, K., and Madsen, C. M. (2011). Reinventing research? information practices in the humanities. *SSRN Electronic Journal*.
- Caidi, N., Allard, D., and Quirke, L. (2010). Information practices of immigrants. *Annual Review of Information Science and Technology*, 44(1):491–531.
- Campbell, L. (2017). The information-seeking habits of architecture faculty. *College & Research Libraries*, 78(6).
- Case, M. M. (2008). Partners in knowledge creation: An expanded role for research libraries in the digital future. *Journal of Library Administration*, 48(2):141–156.
- Catalano, A. (2013). Patterns of graduate students' information seeking behavior: a meta-synthesis of the literature. *Journal of Documentation*, 69(2):243–274.
- Cohen, J., Katz, D. S., Barker, M., Haines, R., and Hong, N. C. (2018). Building a Sustainable Structure for Research Software Engineering Activities. In *2018 IEEE 14th International Conference on e-Science (e-Science)*. IEEE.
- Constant, D., Sproull, L., and Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization science*, 7(2):119–135.
- Corbin, J. and Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift für Soziologie*, 19(6):418–427.

- Cosmo, R. D., Gruenpeter, M., and Zacchioli, S. (2020). Referencing source code artifacts: A separate concern in software citation. *Computing in Science & Engineering*, 22(2):33–43.
- Cunningham, L. (2010). The librarian as digital humanist: The collaborative role of the research library in digital humanities projects. *Faculty of Information Quarterly*, 2(1):1–11.
- Dallmeier-Tiessen, S. (2016). Reproduzierbarkeit Und Open Science: Bestandteile Und Erste Erfahrungswerte - Mit Besonderem Augenmerk Auf Software. Online. Presented at Helmholtz Open Science Workshop 2016.
- de Souza, M. R., Haines, R., Vigo, M., and Jay, C. (2019). What makes research software sustainable? an interview study with research software engineers. *CoRR*, abs/1903.06039.
- Deutsche Forschungsgemeinschaft (2019). Guidelines for safeguarding good research practice. code of conduct. Technical report, Deutsche Forschungsgemeinschaft.
- Druskat, S. (2020). Software and Dependencies in Research Citation Graphs. *Computing in Science and Engineering*, 22(2):8–21.
- Dryden, M. D. M., Fobel, R., Fobel, C., and Wheeler, A. R. (2017). Upon the shoulders of giants: Open-source hardware and software in analytical chemistry. *Analytical Chemistry*, 89(8):4330–4338.
- Edmond, J. (2005). The role of the professional intermediary in expanding the humanities computing base. *Digital Scholarship in the Humanities*, 20(3):367–380.
- Edmond, J., Fischer, F., Mertens, M., and Romary, L. (2017). The DARIAH ERIC: Redefining Research Infrastructure for the Arts and Humanities in the Digital Age. *ERCIM News*, -(111).
- Edwards, J. A. (2003). 17 the transcription of discourse. *The handbook of discourse analysis*, 18:321.
- Ellis, D. (1993). Modeling the information-seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly*, 63(4):469–486.
- Fehr, J., Himpe, C., Rave, S., and Saak, J. (2019). Sustainable Research Software Hand-Over. *arXiv e-prints*, page arXiv:1909.09469.
- Fisher, K. E. and Julien, H. (2009). Information behavior. *Annual Review of Information Science and Technology*, 43(1):1–73.
- Fry, J. (2006). Scholarly research and information practices: a domain analytic approach. *Information Processing & Management*, 42(1):299–316.

- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Polity Press.
- Given, L. M. and Willson, R. (2018). Information technology and the humanities scholar: Documenting digital research practices. *Journal of the Association for Information Science and Technology*, 69(6):807–819.
- Goh, G. B., Hodas, N. O., and Vishnu, A. (2017). Deep learning for computational chemistry. *Journal of Computational Chemistry*, 38(16):1291–1307.
- Gray, N. and Mann, R. G. (2011). A journal for the astronomical computing community? *arXiv e-prints*.
- Greyson, D. (2017). Health information practices of young parents. *Journal of Documentation*, 73(5):778–802.
- Greyson, D. (2018). Information triangulation: A complex and agentic everyday information practice. *Journal of the Association for Information Science and Technology*, 69(7):869–878.
- Grimme, S. and Schreiner, P. R. (2017). Computational chemistry: The fate of current methods and future challenges. *Angewandte Chemie International Edition*, 57(16):4170–4176.
- Grosbol, P. and Tody, D. (2010). Making access to astronomical software more efficient. *arXiv e-prints*.
- Grossman, T., Fitzmaurice, G., and Attar, R. (2009). A survey of software learnability. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. ACM Press.
- Gunning, K. (1978). Increasing the reference librarian’s participation in the research process. *Journal of Academic Librarianship*, 4(4):216–217.
- Harlan, M. A. (2012). *Information Practices of Teen Content Creators: The Intersection of Action and Experiences*. PhD thesis, School of Information Systems. Science and Engineering Faculty. Queensland University of Technology.
- Hasselbring, W. (2019). Microservices as architectural style for research software. In *2nd Vienna Software Seminar (VSS) on DevOps and Microservice APIs*, Vienna.
- Hasselbring, W. and van Hoorn, A. (2020). Kieker: A monitoring framework for software engineering research. *Software Impacts*, 5:100019.
- Haupt, C., Meinel, M., and Schlauch, T. (2018). The software engineering initiative of dlr: overcome the obstacles and develop sustainable software. In *2018 IEEE/ACM 13th International Workshop on Software Engineering for Science (SE4Science)*, pages 16–19. IEEE.

- Hemminger, B. M., Lu, D., Vaughan, K., and Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 58(14):2205–2225.
- Hepworth, K., Mensing, D., and Yun, G. W. (2017). Journalism professors’ information-seeking behaviors: Finding online tools for teaching. *Journalism & Mass Communication Educator*, 73(3):255–270.
- Herraiz-Tabernero, I., Amor-Iglesias, J.-J., and del Castillo-San Felix, A. (2007). Libre software for research. *The European Journal for the Informatics Professional*, 9(6).
- Heyer, G., Kahmann, C., and Kantner, C. (2019). Generic tools and individual research needs in the Digital Humanities – Can agile development help? In Draude, C., Lange, M., and Sick, B., editors, *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft (Workshop-Beiträge)*, pages 175–180, Bonn. Gesellschaft für Informatik e.V.
- Hjørland, B. and Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science*, 46(6):400–425.
- Hong, N. C. (2019). Making open source research software visible: A path to better sustainability? (13.09.2019).
- Howison, J. and Bullard, J. (2015). Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature. *Journal of the Association for Information Science and Technology*, 67(9):2137–2155.
- Hucka, M. and Graham, M. J. (2018). Software search is not a science, even among scientists: A survey of how scientists and engineers find software. *Journal of Systems and Software*, 141:171–191.
- Imran, A. and Kosar, T. (2019). Software sustainability: A systematic literature review and comprehensive analysis. *arXiv e-prints*.
- Ince, D. C., Hatton, L., and Graham-Cumming, J. (2012). The case for open computer programs. *Nature*, 482(7386):485–488.
- Jay, C., Sanyour, R., and Haines, R. (2016). “not everyone can use git”: Research software engineers’ recommendations for scientist-centred software support (and what researchers really think of them). *Journal of open research software*, pages null–null.
- Johannisson, J. and Sundin, O. (2007). Putting discourse to work: Information practices and the professional project of nurses. *The Library Quarterly*, 77(2):199–218.
- Johanson, A. and Hasselbring, W. (2018). Software engineering for computational science: Past, present, future. *Computing in Science & Engineering*, pages 1–1.

- Kanewala, U. and Bieman, J. M. (2018). Testing scientific software: A systematic literature review. *arXiv e-prints*.
- Katerbow, M. and Feulner, G. (2018). Recommendations on the development, use and provision of Research Software. Technical report, Deutsche Forschungsgemeinschaft (DFG).
- Katz, D. S. (2017). Software in Research: Underappreciated and Underrewarded. Presented at eResearch Australia.
- Katz, D. S., Druskat, S., Haines, R., Jay, C., and Struck, A. (2019). The state of sustainable research software: Learning from the workshop on sustainable software for science: Practice and experiences (WSSSPE5.1). *Journal of Open Research Software*, 7.
- Katz, D. S., Hong, N. P. C., Clark, T., Fenner, M., and Martone, M. E. (2020). Software and data citation. *Computing in Science & Engineering*, 22(2):4–7.
- Katz, D. S., McInnes, L. C., Bernholdt, D. E., Mayes, A. C., Hong, N. P. C., Duckles, J., Gesing, S., Heroux, M. A., Hettrick, S., Jimenez, R. C., Pierce, M., Weaver, B., and Wilkins-Diehr, N. (2018). Community Organizations: Changing the Culture in Which Research Software Is Developed and Sustained. *Computing in Science & Engineering*, pages 1–1.
- Kethers, S., Treloar, A., and Wu, M. (2017). Building tools to facilitate data reuse. *International Journal of Digital Curation*, 11(2):1–12.
- Kislinskiy, S. and Goch, C. (2016). Qualitätsmanagement von und infrastruktur für open source software in der wissenschaft am beispiel des medical imaging interaction toolkits (mitk).
- Kitzie, V. L., Wagner, T. L., and Vera, A. N. (2020). “in the beginning, it was little whispers... now, we’re almost a roar”: Conceptualizing a model for community and self in LGBTQ health information practices. In *Sustainable Digital Communities*, pages 15–31. Springer International Publishing.
- Kohlbacher, F. (2006). The Use of Qualitative Content Analysis in Case Study Research. *Forum: Qualitative Social Research*, 7(1):23.
- Korobili, S., Malliari, A., and Zapounidou, S. (2011). Factors that influence information-seeking behavior: The case of greek graduate students. *The Journal of Academic Librarianship*, 37(2):155–165.
- Krippendorff, K. (2012). *Content Analysis: An Introduction to its Methodology*. Sage Publications, 3 edition.
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5):361–371.

- Lamnek, S. (2005). *Qualitative Sozialforschung*. Beltz Verlag, Weinheim, 4 edition.
- Lang, S. A. (2018). Perseus digital library. *RIDE*.
- Lee, B. D. (2018). Ten simple rules for documenting scientific software. *PLOS Computational Biology*, 14(12):e1006561.
- Liew, C. L. and Ng, S. N. (2006). Beyond the notes: A qualitative study of the information-seeking behavior of ethnomusicologists. *The Journal of Academic Librarianship*, 32(1):60–68.
- Liyana, S. and Noorhidawati, A. (2017). How graduate students seek for information: Convenience or guaranteed result? *Malaysian Journal of Library & Information Science*, 19(2).
- Lloyd, A. (2007). Learning to put out the red stuff: Becoming information literate through discursive practice. *The Library Quarterly*, 77(2):181–198.
- Lloyd, A., Kennan, M. A., Thompson, K. M., and Qayyum, A. (2013). Connecting with new information landscapes: information literacy practices of refugees. *Journal of Documentation*, 69(1):121–144.
- Löffler, F., Hammitzsch, M., Schieferdecker, I., Nüst, D., and Druskat, S. (2019). Rse4nfdi - safeguarding software sustainability in the nfdi.
- Lubbock, A. L. R. (2019). Accredited scientific software for sustainability. *Nature*, 572(7771):586–586.
- Mangul, S., Martin, L. S., Eskin, E., and Blekhman, R. (2019). Improving the usability and archival stability of bioinformatics software. *Genome Biology*, 20(1).
- Markowetz, F. (2017). All biology is computational biology. *PLOS Biology*, 15(3):e2002050.
- Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(2):1–10.
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Beltz.
- McKenzie, P. J. (2002). Communication barriers and information-seeking counterstrategies in accounts of practitioner-patient encounters. *Library & Information Science Research*, 24(1):31–47.
- McKenzie, P. J. (2003a). Justifying cognitive authority decisions: Discursive strategies of information seekers. *The Library Quarterly*, 73(3):261–288.
- McKenzie, P. J. (2003b). A model of information practices in accounts of everyday-life information seeking. *Journal of Documentation*, 59(1):19–40.

- Mohammad, A. (2020). “she’s dead!” – nursing simulation practices: a discourse analysis approach. *Journal of Public Health Research*, 9(1).
- Monroe-Gulick, A., O’Brien, M. S., and White, G. W. (2013). Librarians as partners: Moving from research supporters to research partners. In *Association of College and Research Libraries Conference*.
- Muna, D., Alexander, M., Allen, A., Ashley, R., Asmus, D., Azzollini, R., Bannister, M., Beaton, R., Benson, A., Berriman, G. B., Bilicki, M., Boyce, P., Bridge, J., Cami, J., Cangi, E., Chen, X., Christiny, N., Clark, C., Collins, M., Comparat, J., Cook, N., Croton, D., Davids, I. D., Éric Depagne, Donor, J., dos Santos, L. A., Douglas, S., Du, A., Durbin, M., Erb, D., Faes, D., Fernández-Trincado, J. G., Foley, A., Fotopoulou, S., Frimann, S., Frinchaboy, P., Garcia-Dias, R., Gawryszczak, A., George, E., Gonzalez, S., Gordon, K., Gorgone, N., Gosmeyer, C., Grasha, K., Greenfield, P., Grellmann, R., Guillochon, J., Gurwell, M., Haas, M., Hagen, A., Haggard, D., Haines, T., Hall, P., Hellwing, W., Herenz, E. C., Hinton, S., Hlozek, R., Hoffman, J., Holman, D., Holwerda, B. W., Horton, A., Hummels, C., Jacobs, D., Jensen, J. J., Jones, D., Karick, A., Kelley, L., Kenworthy, M., Kitchener, B., Klaes, D., Kohn, S., Konorski, P., Krawczyk, C., Kuehn, K., Kuutma, T., Lam, M. T., Lane, R., Liske, J., Lopez-Camara, D., Mack, K., Mangham, S., Mao, Q., Marsh, D. J. E., Mateu, C., Maurin, L., McCormac, J., Momcheva, I., Monteiro, H., Mueller, M., Munoz, R., Naidu, R., Nelson, N., Nitschelm, C., North, C., Nunez-Iglesias, J., Ogaz, S., Owen, R., Parejko, J., Patrício, V., Pepper, J., Perrin, M., Pickering, T., Piscionere, J., Pogge, R., Poleski, R., Pourtsidou, A., Price-Whelan, A. M., Rawls, M. L., Read, S., Rees, G., Rein, H., Rice, T., Riemer-Sørensen, S., Rusomarov, N., Sanchez, S. F., Santander-García, M., Sarid, G., Schoenell, W., Scholz, A., Schuhmann, R. L., Schuster, W., Scicluna, P., Seidel, M., Shao, L., Sharma, P., Shulevski, A., Shupe, D., Sifón, C., Simmons, B., Sinha, M., Skillen, I., Soergel, B., Spriggs, T., Srinivasan, S., Stevens, A., Streicher, O., Suchyta, E., Tan, J., Telford, O. G., Thomas, R., Tonini, C., Tremblay, G., Tuttle, S., Urrutia, T., Vaughan, S., Verdugo, M., Wagner, A., Walawender, J., Wetzel, A., Willett, K., Williams, P. K. G., Yang, G., Zhu, G., and Zonca, A. (2016). The astropy problem. *arXiv e-prints*.
- Murphy-Hill, E., Lee, D. Y., Murphy, G. C., and McGrenere, J. (2015). How do users discover new tools in software development and beyond? *Computer Supported Cooperative Work (CSCW)*, 24(5):389–422.
- Neuefeind, C., Schildkamp, P., Mathiak, B., Karadkar, U., Stigler, J., Steiner, E., Vasold, G., Tosques, F., Ciula, A., Maher, B., Newton, G., Arneil, S., and Holmes, M. (2020). Sustainability strategies for digital humanities systems. In *Proceedings of the DH2020 Humanities Commons*.
- Niedźwiedzka, B. (2003). A proposed general model of information behaviour. *Information research*, 9(1):9–1.

- Nielsen, L. H. and Van De Sandt, S. (2019). Tracking citations to research software via pids.
- Niemeyer, K. E., Smith, A. M., and Katz, D. S. (2016). The challenge and promise of software citation for credit, identification, discovery, and reuse. *Journal of Data and Information Quality*, 7(4):1–5.
- Oishi, J. S., Brown, B. P., Burns, K. J., Lecoanet, D., and Vasil, G. M. (2018). Perspectives on reproducibility and sustainability of open-source scientific software from seven years of the dedalus project. *arXiv e-prints*.
- Olsson, M. (2007). Power/knowledge: The discursive construction of an author. *The Library Quarterly*, 77(2):219–240.
- Olsson, M. (2016). Making sense of the past: The embodied information practices of field archaeologists. *Journal of Information Science*, 42(3):410–419.
- Olsson, M. R. (2010). All the world's a stage – the information practices and sense-making of theatre professionals. *Libri*, 60(3).
- Palmer, C. L. and Cragin, M. H. (2009). Scholarship and disciplinary practices. *Annual Review of Information Science and Technology*, 42(1):163–212.
- Pan, X., Yan, E., and Hua, W. (2016). Disciplinary differences of software use and impact in scientific literature. *Scientometrics*, 109(3):1593–1610.
- Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*. SAGE, 3 edition.
- Pianosi, F., Sarrazin, F., and Wagener, T. (2019). How successfully is open-source research software adopted? results and implications of surveying the users of a sensitivity analysis toolbox. *Environmental Modelling & Software*, page 104579.
- Richardson, C. (2017). The research software engineer- an emerging new role in academia in the uk.pdf.
- Rimmer, J., Warwick, C., Blandford, A., Gow, J., and Buchanan, G. (2006). Humanities scholars' information-seeking behaviour and use of digital resources. *Digital Libraries in the Context of Users' Broader Activities*, page 19.
- Rinsdorf, L. (2013). Qualitative methoden. In Umlauf, K., Fühles-Ubach, S., and Seadle, M., editors, *Handbuch Methoden der Bibliotheks-und Informationswissenschaft.*, pages 64–79. DeGruyter Saur, Berlin/Boston.
- Roos, A. (2015). Medical scientists' information practices in the research work context. *Health Information & Libraries Journal*, 32(1):23–36.
- Rupp-Serrano, K. and Robbins, S. (2013). Information-seeking habits of education faculty. *College & Research Libraries*, 74(2):131–142.

- Saputri, T. R. D. and Lee, S.-W. (2020). Software analysis method for assessing software sustainability. *International Journal of Software Engineering and Knowledge Engineering*, 30(01):67–95.
- Savolainen, R. (2007). Information behavior and information practice: Reviewing the “umbrella concepts” of information-seeking studies. *The Library Quarterly*, 77(2):109–132.
- Savolainen, R. (2015). Cognitive barriers to information seeking: A conceptual analysis. *Journal of Information Science*, 41(5):613–623.
- Schelig, K., Pampel, H., Konrad, U., Fritzsche, B., Schlauch, T., Nolden, M., Zu Castell, W., Finke, A., Hammitzsch, M., Bertuch, O., and Denker, M. (2019). Dealing with research software: Recommendations for best practices. Technical report, Helmholtz Association.
- Shamir, L., Wallin, J. F., Allen, A., Berriman, B., Teuben, P., Nemiroff, R. J., Mink, J., Hanisch, R. J., and DuPrie, K. (2013). Practices in source code sharing in astrophysics. *Astronomy and Computing*, 1:54–58.
- Shortridge, K. (2009). Talking amongst ourselves - communication in the astronomical software community. *arXiv e-prints*.
- Smith, A. M., Katz, D. S., and and, K. E. N. (2016). Software citation principles. *PeerJ Computer Science*, 2:e86.
- Soito, L. and Hwang, L. J. (2017). Citations for software: Providing identification, access and recognition for research software. *International Journal of Digital Curation*, 11(2):48–63.
- Struck, A. (2018). Research Software Discovery: Challenges, Risks And Opportunities. In *IEEE eScience 2018*. Zenodo.
- Talja, S. (2005). The domain analytic approach to scholars’ information practices. In *Theories of Information Behavior*. Information Today, Inc.
- Talja, S. and Maula, H. (2003). Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. *Journal of Documentation*, 59(6):673–691.
- Talja, S. and McKenzie, P. J. (2007). Editors’ Introduction: Special Issue on Discursive Approaches to Information Seeking in Context. *The Library Quarterly*, 77(2):97–108.
- Talja, S., Tuominen, K., and Savolainen, R. (2005). “Isms” in information science: constructivism, collectivism and constructionism. *Journal of Documentation*, 61(1):79–101.

- Tennant, J., Agarwal, R., Baždarić, K., Brassard, D., Crick, T., Dunleavy, D. J., Evans, T. R., Gardner, N., Gonzalez-Marquez, M., Graziotin, D., Tzovaras, B. G., Gunnarsson, D., Havemann, J., Hosseini, M., Katz, D. S., Knöchelmann, M., Madan, C. R., Manghi, P., Marocchino, A., Masuzzo, P., Murray-Rust, P., Narayanaswamy, S., Nilsson, G., Pacheco-Mendoza, J., Penders, B., Pourret, O., Rera, M., Samuel, J., Steiner, T., Stojanovski, J., Uribe-Tirado, A., Vos, R., Worthington, S., and Yarkoni, T. (2020). A tale of two 'opens': intersections between free and open source software and open scholarship.
- Tuominen, K., Talja, S., and Savolainen, R. (2005). The social constructionist viewpoint on information practices. In *Theories of information behavior*. Information Today, Inc.
- van Zundert, J. (2012). If you build it, will we come? large scale digital infrastructures as a dead end for digital humanities. *Historical Social Research / Historische Sozialforschung*, 37(3 (141)):165–186.
- Veinot, T. C. (2007). “the eyes of the power company”: Workplace information practices of a vault inspector. *The Library Quarterly*, 77(2):157–179.
- Waltemath, D. and Wolkenhauer, O. (2016). How Modeling Standards, Software, and Initiatives Support Reproducibility in Systems Biology and Systems Medicine. *IEEE Transactions on Biomedical Engineering*, 63(10):1999–2006.
- Wang, P., Dervos, D. A., Zhang, Y., and Wu, L. (2008). Information-seeking behaviors of academic researchers in the internet age: A user study in the united states, china and greece. *Proceedings of the American Society for Information Science and Technology*, 44(1):1–29.
- Weiner, B. J., Blanton, M. R., Coil, A. L., Cooper, M. C., Davé, R., Hogg, D. W., Holden, B. P., Jonsson, P., Kassin, S. A., Lotz, J. M., Moustakas, J., Newman, J. A., Prochaska, J. X., Teuben, P. J., Tremonti, C. A., and Willmer, C. N. A. (2009). Astronomical software wants to be free: A manifesto. *arXiv e-prints*.
- White, S. R., Amarante, L. M., Kravitz, A. V., and Laubach, M. (2019). The Future Is Open: Open-Source Tools for Behavioral Neuroscience Research. *eNeuro*, 6(4):ENEURO.0223–19.2019.
- Wilson, T. (1994). Information needs and uses: Fifty years of progress. *Fifty Years of Information Progress: A Journal of Documentation Review*.
- Wittgenstein, L. (1990). *Tractatus logico-philosophicus*. Reclam-Bibliothek Philosophie, Geschichte, Kulturgeschichte. Reclam, Leipzig, 1. Aufl. edition. Lizenzausg. des Suhrkamp-Verl., Frankfurt am Main. - Bibliogr. und Literaturverz. S. 478 - [485].
- Wolski, M., Howard, L., and Richardson, J. (2017). The importance of tools in the data lifecycle. *Digital Library Perspectives*, 33(3):235–252.

- Ye, Y., Boyce, R. D., Davis, M. K., Elliston, K., Davatzikos, C., Fedorov, A., Fillion-Robin, J. C., Foster, I., Gilbertson, J., Heiskanen, M., Klemm, J., Lasso, A., Miller, J. V., Morgan, M., Pieper, S., Raumann, B., Sarachan, B., Savova, G., Silverstein, J. C., Taylor, D., Zelnis, J., Zhang, G. Q., and Becich, M. J. (2019). Open source software sustainability models: Initial white paper from the informatics technology for cancer research sustainability and industry partnership work group. *arXiv e-prints*.
- Zirkelbach, C., Krause, A., and Hasselbring, W. (2020). The collaborative modularization and reengineering approach coral for open source research software. *International Journal On Advances in Software*, 13(1&2):34–49.
- Zwart, S. P. (2018). Computational astrophysics for the future. *Science*, 361(6406):979–980.

Glossary of Mentioned Software and Services

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

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

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 .

Dropbox .

Excel (Microsoft) .

Germanistik.net .

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

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

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

JSTOR .

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

Little Scott Jones .

Mozilla Thunderbird .

Perseus .

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=>.

Word (Microsoft) .

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

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

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: Vorname:

Matr.Nr.:

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

.....
.....
.....

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

Unterschrift