Humboldt University of Berlin

Institute for Library and Information Science

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

Ronny Gey

September 22, 2020

Zusammenfassung

Zusammenfassung

Abstract

Abstract

Contents

List of Figures

List of Tables

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.

2. Theory

2.1. Information Seeking

- Information Science briefly described

Seeking, Searching, Retrieval

- Information Seeking Research (Ingwersen2005) – Concepts: Strategies – Collaborative IS: Shah2013

- Distinction between behaviour and practices:

the concepts of information behavior and information practice emerge from different discourses that open alternative viewpoints on information seeking. Savolainen2007

Bates2010 - information behaviour, Case2007 - information behaviour

- behaviour: wilson, ellis, kuhltau, Niedzwiedzka2003 – different conceptualizations: intra/inter/extrapersonal (Feinman – transgender ib: pohjanen2016

- practices: McKenzie and Talja

2.2. Information Practices

- Introduction: 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)).

- Starting with McKenzie McKenzie2003, 2003a

- and Talja

- further examples of Information Practices: Savolainen 2007 - LitReview

2.3. Research Software

- definition - examples - importance for research, in the research process

2.3.1. Information Practices towards research software

- examples of studies, what has been studied yet

2.4. Domain Analysis: Humanities/Philology

- short: humanities, long: philology - definition - characteristics: subjects, work procedures, tools, ...

2.4.1. 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). With information practices we mean practices of seeking, managing, giving, and using information in context (Palmer and Cragin, 2009). 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 science[1]. Hence, I have decided that all data generated during the concept, survey, analysis, and writing phase is publicly available on GitHub[2], 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/
[2] https://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 arbitraryly 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 test persons 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 interview[3]. In the beginning of the interviews, I briefly introduced myself and the back-

[3] see 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 orthografically transcribed the interviews recordings using EasyTranscript, an open source transcription software[4]. 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 QualCoder[5], 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).

4. Findings

4.1. Introduction of the Empirical Field

Classical Philology is one of the oldest subjects at the university. It first flourished in the sixteenth century. In the nineteenth century, the institute was a center of classical philology. The dominant field of work in the nineteenth century 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 classical philology or philologists that they 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 interview[1]].

> *ein teilbereich usnerer arbeit nennt sich ja auch quellenkritik und das ist eine ganz zentrale frage, also ähm ja und die lässt sich auch in verschieden teilbereiche spalten. das eine ist ja die sicherheit oder nichtsicherheit, die man haben kann, was die gestalt des textes selber angeht, der ja wie gesagt, der ja überlieferungsprozesse und zum teil auch transformationsprozesse durchlaufen hat über zwei jahrtausende in aller regel.* [Peter, 1]

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

Table 2.: Overview of interviewees

Name	Position	No. Interviews	Duration of Interviews
Sandra	Post-doc	2	2h25min
Peter	PhD student	1	2h09min
Marie	Post-doc	2	2h44min
Marco	PhD student	2	2h56min
Christian	PhD student	2	2h07min

4.2. Information Practices in Research

4.3. Software Use

4.4. Information Practices Towards Software

4.5. Domain Factors

5. Discussion

- 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

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

Ahmadianyazdi, F. and Chandrashekara, M. (2018). Research challenges and delayed gratification in information seeking behavior: a case study of research scholars. *Library Philosophy & Practice*.

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.

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.

Bøyum, I. and Aabø, S. (2015). The information practices of business PhD students. *New Library World*, 116(3/4):187–200.

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., Jirotka, M., Eccles, K., and Madsen, C. M. (2011). Reinventing research? information practices in the humanities. *SSRN Electronic Journal*.

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.

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.

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.

Deutsche Forschungsgemeinschaft (2019). Guidelines for safeguarding good research practice. code of conduct. Technical report, Deutsche Forschungsgemeinschaft.

Edmond, J. (2005). The role of the professional intermediary in expanding the humanities computing base. *Digital Scholarship in the Humanities*, 20(3):367–380.

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.

Fry, J. (2006). Scholarly research and information practices: a domain analytic approach. *Information Processing & Management*, 42(1):299–316.

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.

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.

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.

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.

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.

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.

Kethers, S., Treloar, A., and Wu, M. (2017). Building tools to facilitate data reuse. *International Journal of Digital Curation*, 11(2):1–12.

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.

Lamnek, S. (2005). *Qualitative Sozialforschung*. Beltz Verlag, Weinheim, 4 edition.

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).

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.

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*.

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.

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.

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.

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.

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.

Scheliga, K., Pampel, H., Konrad, U., Fritzsch, 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.

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. 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.

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.

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.

Wilson, T. (1994). Information needs and uses: Fifty years of progress. *Fifty Years of Information Progress: A Journal of Documentation Review*.

Wolski, M., Howard, L., and Richardson, J. (2017). The importance of tools in the data lifecycle. *Digital Library Perspectives*, 33(3):235–252.

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

Humboldt-Universität zu Berlin
Philosophische Fakultät
Institut für Bibliotheks- und Informationswissenschaft

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 ……