Improvement of Research Platforms to Raise Researchers Acceptance for Research Data Management - First Study on Results of Surveys among German Universities (2018)

Konstantin Wilms
Ania Lopez

1. Introduction

With the arise of digitalization in science, the amount of research data, i.e. data produced or used during research, has been growing very fast. With the technological development and raising of new scientific methods, also the relevance of research data has increased enormously. Reliable verification of scientific results is possible with the deployment of new technological methods. Furthermore, technology makes it feasible to make new and innovative research built upon existing data. Research data has moved to be an important resource for research, being even called the gold of science [1].

Type of research data differs within different scientific disciplines, given different formats (e.g. text, images, video, audio recordings, etc.). At the same time, the needs for handling data differ, including different levels of legal treatment of sensitive data in some disciplines, handling of very large data sets in computational intensive science, access restrictions for re-use, etc. Some scientific disciplines show a long tradition in handling their data in a standardized way, there exist large international well-accepted data centers, while other disciplines lack of standards and experience in management of research data [2], [3].

Platforms exist, where researchers are able to store, manage and spread their data [4]. As literature shows, these platforms struggle in terms of adoption and acceptance rate [5]. At the same time, researchers are starting to formulate their needs with respect to the management of their research data, addressing the necessity for technical platforms to support the whole research process (e.g. sharing and storage of data or preparation for publication).

To handle the ever-growing number of research data in academia, universities and research institutions started to set up own technical infrastructures to manage the increasing flood of information.

To avoid the development of inadequate information systems, universities started to carry out surveys asking their researchers what they want. The resulting information is highly relevant for current IS development and could help to design and improve current SaaS applications [6].

The latter resulted in the following research question:

RQ: How can technical research solutions be improved in order to raise researchers' acceptance for Research Data Management?

As stated 2015 in [7] there are still "only few empirical studies on the subject of research data management in scientific practice [and] these are often characterized by very small samples [and] restrictions to individual departments". Still, to find responses to the stated research question, it cannot be reasonable to wait until every single university or research institution carries out a survey asking researchers for their needs for research platforms. Therefore, this study compares results from ten surveys that have already been carried out at universities, with a focus on German universities. Being the first comparison of survey results regarding RDM-practice among scientists, this thesis has a strong exploratory character.

The structure of the thesis will be presented as follows:

- A Literature Review presents the state of the art regarding research in the topic of acceptance of research data management connected with the evolution of research platforms.
- The method for comparison of survey results is explained in detail, containing the description of the qualitative content analysis, and the description of quantitative facts and the distribution of frequencies among the defined categories.
- The results are presented, describing the categories that resulted from the qualitative content analysis.
- In a discussion, the results are evaluated in relation to the research question showing new outcomes for the research question and describing the limitations of the method.
- In a conclusion, we explain our interpretation of results an give an outlook for further research questions.

2. Literature Review

There exist several definitions of the term research data. In this thesis we refer to research data management as “the organization of data, from its entry to the research cycle through to the dissemination and archiving of valuable results” [8].

This includes the idea of coherent standards and policies on data exchange via different platforms, long-term storage of data, quality of data, security, and also free data access [9], [10]. Several research funding institutions, such as the National Science Foundation (NSF)[1] or the German Research Foundation (DFG)[2] have set up guidelines on adequate RDM. For example, the guidelines of the DFG recommend researchers to store their primary research data for at least ten years on suitable memories [11]. The use of data management plans is becoming a standard requirement for application for funding [12]–[15]. Various publishers are more and more demanding good data management practices from scientists, such as the publication of research data associated to classical paper publications [16]–[19]. To follow these guidelines means to conduct RDM in a specific way.

Besides direct requirements to the researchers coming from funding agencies and publishers, a lot of political statements reflect the importance of the topic. In Germany, the importance of RDM has been promoted by several institutions, such as the German Rectors' Conference[3] [20], the Alliance of Science Organisations in Germany[4] [21] and the Council for Scientific Information Infrastructures[5] [22]. In a next scale, the European Commission has described a vision of "a scientific e-infrastructure that supports seamless access, use, re-use and trust of data" (see [23], p.4).

Beyond external requirements, appropriate research data management offers several benefits to the scientific community, and for researchers and scholars in general. For example, the idea of open data access is said to lead to higher citation rates [24], as well as perceived benefits regarding the own career, and altruism [25]. Due to openly shared data the research process becomes more transparent [26]. Even though researchers often fear negative criticism about their research as a result of data sharing, studies revealed that the granting of insights defend against the allegation of misconduct [26]–[28]. According to [29] there are four major goals for researchers to share data: "(1) to reproduce or to verify research, (2) to make results of publicly funded research available to the public, (3) to enable others to ask new questions of extant data, and (4) to advance the state of research and innovation" (Borgman 2012, pp. 1). These results go in line with findings of [30], [31], whereto the access to useful resources can motivate researchers to increase their collaboration with others. RDM does not only support data exchange and collaboration but also helps to avoid data get lost over time that can't be traced anymore, so called "dark data" [32]. As previous studies indicate, a majority of researchers have already experiences with data loss [33], [34].

Researchers need adequate IT solutions in order to perform adequate RDM. As stated by [5], institutions should offer platforms with standardized guidelines for researchers in order to support RDM and to help researchers with their publications by organizing their data. Additionally, already existing platforms need to be improved, for example in terms of data description, yet there already exist platforms that do partly meet the demanded requirements [5], [35]. As recent research on a status quo indicates, research platforms primarily focus on basic functionalities regarding publication management, while collaboration features have mostly been ignored [4].

According to [33], researchers will need at least 100GB for research data storage in the future, and even more for the medical and artistic institutions. As researchers seem to describe their data individually and by this inconsistent, [33] recommend a comprehensive research data management platform or infrastructure with regard to the disciplinary requirements, just as standardized guidelines.

1. www.nsf.gov (retrieved June 8, 2017)

2. The German Research Foundation (DFG) is the largest independent research funding organization in Germany, see also www.dfg.de/en/ (retrieved June 8, 2017)

3. The German Rectors' Conference (HRK) is the voluntary association of public and government- recognised universities and other higher education institutions in Germany, having currently 268 member institutions in which around 94 per cent of all students in Germany are enrolled, see also https://www.hrk.de/home/ (retrieved June 8, 2017).

4. The Alliance of Science Organisations in Germany (Allianz der Wissenschaftsorganisationen) is a union of the most important German research organisations. It issues statements relating to research policy and funding and the structural development of the German research system. See also https://www.leopoldina.org/en/about-us/cooperations/alliance-of-science-organisations/ (retrieved June 8, 2017).

5. The Council for Scientific Information Infrastructures (RfII) was initiated by the Joint Science Conference of the Federal States (GWK) and the Federal Government of Germany for a service period of four years in 2014. The 24 members represent a broad spectrum of scientific disciplines and institu- tions. The Council´s scope of work is the strategic development of a contemporary and sustainable infrastructure for access to scientific Information. See also http://www.rfii.de/en/ (retrieved June 8, 2017)

Yet, current guidelines on RDM do not restrict the usage of information technologies and specific IT solutions are not defined in any detail. Hence, there is no specification whether RDM is performed within a repository (e.g. as in [6]), a research platform or within a web-based research portal, where for example research activities are presented and discussed (e.g as in [4]). Thus, adequate technology descriptions could be located at the infrastructure level [6], [35], but also on the software level, where different software services like e-learning [36] or data management applications, support the academic workflow [37]. As described by [6] adequate technologies for RDM require cloud technology. In Germany, several universities and research institutions have started to see the need for new IT technologies to address the changes that are coming into the digitalization of science and scientific methods. Not only moved by national and international recommendations[6], but also by the results of surveys regarding the use of RDM-platforms among researchers, universities see the possibility to develop services and to find new roles within the context of digitalization of science.

This results in the building and development of technical (storage) infrastructures, training and support offers, implementation of research data policies, etc. While the majority of technical infrastructures are used as cloud-services, they have the potential to serve as research platforms [6].

[37] point out that cloud technologies are based on hardware and software components, which are provided as services via the internet. Three different service types can be distinguished: The infrastructure layer (Infrastructure-as-a-Service, IaaS), the platform layer (Platform- as-a-Service, PaaS), and the software layer (Software-as-a-Service, SaaS). IaaS describes a service which provides scalable IT-infrastructures to the customers. In this context [38] differentiate between the service of external storage capacity and the service of using external IT performance. Besides the infrastructure layer, the PaaS layer is built on the infrastructure level. This layer offers frameworks in the form of development or application environments to the user [39]. This type of platform service allows users to upload programmed applications to the cloud, so that they can be used as web-based applications [39]. The third layer is the software layer. At this level web-based application and various cloud services such as office, e-mail or calendar applications are offered to the user. Data management applications can be controlled as well [37].

According to the latter definitions, we define virtual research environments (VRE) as cloud based technologies including three service layers. VRE include infrastructure services (e.g. repositories which are run individually to store research data) [35], platform services as defined by [39], and software services, which e.g. support researchers in terms of collaboration, documentation or data management.

In order to improve the acceptance of those infrastructures, universities started to set up surveys and ask their researcher what they want.

3. Method

To answer the research question "*How can Virtual Research Environments be improved in order to raise researchers' acceptance for Research Data Management?*", a qualitative content analysis was set up. The analysis included currently available studies on RDM among universities in Germany. The aim of the analysis is to find out more about the current insights on technological requirements for VRE. While the studies which were consulted for the analysis were of both types, quantitative and qualitative, only content results were used for the examination. Since the status quo of RDM guidelines differs the most between countries [6], only RDM studies from Germany were selected in the first step. Hence, all survey results are associated with a German university, or in one case a network of universities in one federal state of Germany (Baden-Württemberg).

To make sure that all relevant research results on current RDM investigations among Germany were related in our analysis, the German research network for RDM "forschungsdaten.org" was consulted. This resulted in overall ten surveys from the universities of Berlin [40], [41], Münster [7] , Hannover [42], Göttingen [43]–[45], Kiel [46], Marburg [47], Hamburg [48], Trier [49], Siegen[7] , and a commulated study of all universities among the German state of Baden-Württemberg (bwFDMCommunities). Table 1 gives an overview of the sample size of all qualitative and quantitative survey settings.

Institution	Qualitative	Quantitative
	n	n

6 In Germany, for instance the German Rectors' Conference stated the importance of Research Data Management by calling it a "central strategic challenge for university management" [58].

7. The survey data and results from the IT center of the University of Siegen are just internal documents and not published.

University of Berlin	17	499
University of Hannover	20	294
University of Kiel	-	218
Universities of Baden-Württemberg (bwFDMCommunites)	779	-
University of Marburg	-	427
University of Hamburg	-	96
University of Trier	-	193
University of Münster	-	667
University of Siegen	-	200
University of Göttingen	-	877

Table 1. *Sample sizes of qualitative and quantitative survey results.*

The analysis of all ten survey results was based on Mayring's qualitative content analysis [50]. Due to this method, the material is generalized in the first place and then reduced during an abstraction process [50]. The central aspect of qualitative content analysis is the development of a categorical system which helps to identify the aspects that seem necessary for answering the research question from the wealth of the interview material [50]. The first coding process resulted in overall number of eight topics. All categories should be distinct. The validity of the first coding process was tested with a second and a third coder. As a decent accordance of the coders was proved, all results were finally coded.

By use of the inductive category development as suggested by [50], six main categories and eight sub categories were evolved. The categories "Features" and "Protection mechanisms and Security" did not include sub categories. Figure 1 shows the distribution of frequencies.

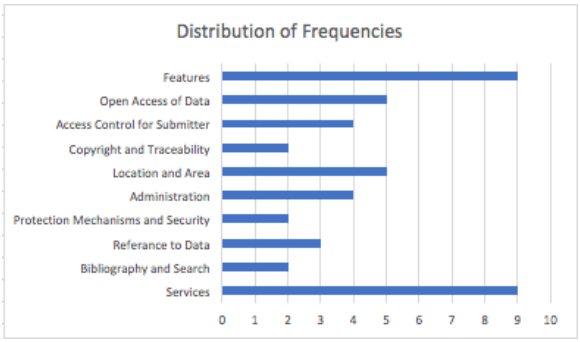

Figure 1. Frequency distribution over sub- and main categories.

4. Results

Features. The main category "Features" contains a summarization of features required by the researchers of the included studies. As stated by the Universities of Baden-Württemberg the VRE should include the possibility of "open reviews or awards for researchers, e.g. for best data publications of a doctoral thesis / research discipline" to increase researchers awareness for RDM and research transparency. According to findings of [3], [33], whereto incentives indeed have the potential to increase researchers' utilization of VRE, gamification elements like badges or leaderboards offer attractive cost-efficient solutions within VRE [6].

As another feature for a VRE respondents demand for an assistance feature which facilitates the documentation of the data over the duration of the project. Related to this demand, a researcher of the University of Trier suggests that (semi-)automated approaches from the area of "machine learning" could help the scientists to supplement research data with meta information. The participant rely on the results of the research project "DataWiz" [51] of the Leibniz Institute in Germany which is designed as a semi-automated assistance system that supports researchers in the implementation of research data management procedures (eg. planning, data processing, documentation, versioning and archiving) during the research process and provides the necessary functionalities.

Support functionalities are also required in terms of quality management. Researchers of the Universities of Hannover, Kiel, Hamburg and Münster indicated, that research data are used to be stored in outworn data formats. To avoid incompatible data versioning, VRE should support researchers with feedback functionalities.

Overall, a broad majority of researchers called for adequate usability standards as primary demand for a VRE. These include interoperability and easy access, which is not restricted by user's operation system. Very important is also the compatibility with different devices.

Access Control. The second main category composes "Access Control" with its two sub categories "Open Access of Data", and "Access Control for Submitter".

"Open Access of Data" implies that research data is accessible for everyone, regardless if they are researchers or not. As expected and described in previous literature [6], [29], [52], the opinion of open access of research data within a VRE is still divided. While the results of different studies implicate that a broad majority of researcher see benefits in open data access in VRE (Kiel, Berlin), other survey results show, that researchers are still suspicious about public data access. The results of the University of Berlin for

example imply a high value in open data repositories, since these allows to reach a broad range of public. As stated by a participant of the University of Berlin “It also points to the positive effect that such a repository could have on the notoriety of less popular research areas”. At the University of Marburg, the results indicate a different perspective, where the acceptance of research repositories is negatively correlating with the “openness” of a system. At the same time, however, half of the respondents downloaded or planned to download research data from a data archive.

Likewise “Access Control for Submitter” is much needed, since multiple universities call for this feature. According to the results of the University of Hamburg in case of open access the submitter needs “full control over permissions and adjustable access for different user groups (especially external)”. As stated by a participant of Marburg, own control over research data is of central importance for researchers.

Copyright and Traceability. The main category “Copyright and Traceability” includes two sub-categories, namely “Copyright and User Agreement”, and “Usage, Documentation and Traceability”.

“Copyright and User Agreement” describes researchers’ demand for clear data policies and regulations. As stated by researchers of the Universities of Baden-Würtemberg, "this requires incentives and clarity about the legal status of the data transfer, guidelines, appropriate structures, and possibly new forms of publication”. As researchers work mostly with self-generated data, there is an urgent support for protection and marketing rights in order to be able to control data access on the research data. According to the findings of the University of Marburg, VRE should support copyright protection mechanisms. One idea is to set up digital utilization agreements in exchange for research data. According to those agreements, researchers could set up own conditions which need to be fulfilled in order to re-use research data.

“Usage, Documentation and Traceability” was called by the University of Hamburg where participants recommended that access attempts should be documented. In this context data could be traced over the border of the VRE.

Administration and Responsibility. “Administration and Responsibility” includes two sub-categories, namely “Location and Area”, and “Administration”.

“Location an Area” describes researchers need for a local or a central VRE infrastructure. As the results of the surveys show, researchers refuse to use intercontinental infrastructure, like infrastructure which is located in the US for example. These findings go in line with the findings of [35], [53], whereto researchers prefer local IT infrastructures as research repositories. In addition to unresolved copyright issues, intercontinental repositories are also criticized for massive data protection problems. Known data protection violations, as in the case of the NSA scandal, have had a lasting impact on user confidence in the cloud [54], [55]. As multiple universities stated out, local infrastructures located at the institutions are preferred.

"Administration" includes researchers’ requirements in terms of regulation and competence. Whether VRE are under central administration or could be run on a local server is a question that also shows up within the surveys. Administration does not only mean to maintain VRE’s hardware and software services, but also to serve as a source of responsibility, where researchers could belong to. Since there is no major institution yet, which is responsible for general and discipline-across RDM guidelines, and since institutional guidelines might differ from guidelines of third party funding institutions or from national recommendations, researchers ask for central administration. According to the answers of researchers from the Universities of Baden-Württemberg, Berlin, Kiel, Hannover and Hamburg it is important to establish a directory of specialised competences, skills and expertise.

Protection Mechanisms and Security. The category “Protection Mechanisms and Security” contains researchers' requirements for adequate data protection standards. As University of Hamburg surveys point out, there is a huge demand for data security. As pointed out by a researcher of the University of Kiel, an adequate VRE needs to 1) offer a secure and reliable storage space for research data, and 2) must guarantee data security. Specific security standards or technological demands were not named overall. The reasons for such a high security demand are different. As the results of the University of Hamburg indicate, researchers are dealing with a lot of research data out of (industrial) cooperations, where mostly qualitative raw data can not be completely anonymized. Hence the protection of the research data is of tremendous importance, since personal data of participants could be leaked.

Literature-management. Belonging to the main category "Literature-management", the two sub categories "Reference to Data" and "Bibliography and Search" were determined.

The sub category "Reference to Data" was created due to different statements of researchers calling for a repository, were publications are interlinked with their research data. Researchers at the University of Hamburg recommend a solution, where research data was interlinked with a specific digital object identifier (DOI) in order to be able to cite on specific research data. Participants of the University of Berlin specifically asked for "web-based solutions that enable the linking of research data with the author's online publications".

Within the sub category "Bibliography and Search", the demand for search and filter functions within a research platform are discussed. As the results from the universities of Baden-Wurttemberg and the University of Hamburg imply, there is a need for search and filter functionalities in order to locate publications and research data within the repository structure. As a researcher of Baden-Württemberg stated, "If repositories are better standardized, repositories, links and a consistent repository culture are also conceivable in the storage of research data". As the survey results of the University of Hamburg indicate, there is a direct demand for filter and search functionalities among research repositories.

Services. The last main category "Services" consists of researchers demands for service requirements in order to use research platforms. Expect the University of Trier all universities call for service support in term of legal issues. As several researchers respond, there are huge uncertainties among universities employees in terms of legal restrictions. Researchers do not know which data they have to offer or which data they are not allowed to offer. Furthermore, there are questions about legal restrictions in term of institutional cooperations or data which is collected under the waiver of disclosure. Seven out of ten surveys implied a urgent need for technical support. According to [56], technical support is of tremendous importance for IT projects, since insufficient services might lower IT projects success. As third service requirement researchers ask for a general support in terms of questions related to RDM. Here, the researchers indicate that the current knowledge about RDM restrictions is quite low, and universities were in demand to brief their employees adequately.

5. Discussion

The results of the qualitative content analysis serve IT consultants at universities and research departments to increase their awareness of what researchers expect of research platforms or research repositories.

While most of the researchers demand for research repositories which could be used to store data for at least 10 years, others demand for extensive virtual research environments with access control, supported metadata or collaboration functions. The categories resulting from the analysis show, that besides several functionalities, there are consistent demands among all universities.

The analysis included 10 survey with a total participation of 816 researchers within qualitative studies and 3471 researchers within quantitative studies. The results show that researchers' needs differ to a great extent.

Within the categories that resulted from the qualitative content analysis, the category *Services* shows the highest demand. Although *Services* is not directly connected to technical development, the prominence of the appearance goes in line with the findings, that technological success is strongly connected to the technical competences of users and technical support [56].

Not surprising is the fact, that within the category *Features* a high demand of different technical requirements was represented. This shows the heterogeneity between the different scientific disciplines but also the singular working methods of researchers. Hence, from IT consultant's perspective there must be weighted to what extend this singular needs can be fulfilled in a realistic way.

A returning issue within the category *Administration and Responsibility* and *Protection Mechanisms and Security* was the demand for infrastructure to be local at institutional level. Here, universities gain advantage from the fact to be considered trustworthy

and reliable. This fact builds a barrier for external developer to get into the field. Especially in academia, security issues seem to be a major factor that keeps researchers within institutional IT borders, and lowers the acceptance of external IT solutions [57].

The category *Access Control* shows the different perception of the topic Open Science. In terms of open data access, access permission should be at least in control of the submitter, since specific data is not meant to be freely available. This type of data is mostly related to the private sector where qualitative interviews are collected on the waiver of disclosure of third parties. Also, some data can't be modified to an extend where the originator will be completely anonym. These personal data need to be protected and can't be shared with public. Rather than hold back this

sensitive data, researchers call for adequate data protection standards which allow them to store the data in central repositories or platforms. These platforms are deemed to run on local infrastructures and to be protected by national protection law.

The within literature [29], [52] divided opinion towards the meaning and importance of sharing research data is clearly reflected. In a conjunction to the expressed needs within the category *Literature-management*, especially the sub category *Reference to Data,* researchers fears and prejudices regarding sharing of data could be picked up by bringing up the benefits in terms of reputation that brings the citation of shared data. This of course assumes the existence of persistent reference to data and semantic connection between classical text publications and data.

The presented research shows several limitations. First of all, the qualitative nature of the content analysis provides an in-depth insight into researchers demand of digital platforms at the university to conduct RDM. However, the insights are limited due to the small number of surveys which were consulted in the analysis. There is no guarantee that the included surveys are representative of the population of researchers. Furthermore, the included surveys only represent universities among Germany. Germany might differ from other countries in terms of RDM regulations and requirements. There is a need for further investigation, which can be supported by quantitative research. Also, a future investigation on an international level is required.

To validate the findings in this work, future research is needed to investigate in a broader range of sample size. Future investigations need to research requirements for VREs on an international level. Especially requirements like open access might be different in countries which are more collectivistic and countries where data protection laws are not as strict as in Germany might have another understanding of data protection and security. Finally, future research should also include approaches of design science, in order to design and conceptualize appropriate virtual research environments. Instead of continuous investigations on status quo of RDM among universities, future investigations should focus on current barriers that seem to diminish the acceptance of RDM.

6. Conclusion

From a practical viewpoint, there are several implications for IT consultants to design and improve VREs or extend research repositories.

The results show that researchers often have a different idea of the concept of research platform or VRE. Depending on the discipline or individual experiences with digital research methods, this can range from just needing infrastructure (i.e. IaaS) to a complete VRE, where complex software applications are required (i.e. intelligent support systems). Currently, there are strongly diverging perspectives of what VRE should offer and how RDM should be technically supported in the future. While data preservation and reuse practices are already a topic in some field of science, other disciplines still have have a number of reservations against digital research practices.

Therefore, the development of VREs will have to take place on all service types (IaaS, PaaS an SaaS) and will have to address very heterogeneous and specific needs, focussing on interoperability of single tools shifting away from monolithically constructions or "one-fits-all"-solutions. The presented results from the analysed surveys underline the need for locally or institutional solutions for handling of data. This result may be special for German universities because law restrictions. Further research and analyse will be needed to understand if this result is transferable. As the results show, future technological investigations need to be pushed on an institutional or governmental level, since researchers in most cases are restricted to internal infrastructures and policies.

Since there is yet only a lit number of research investigations within the field of RDM, the development of VREs is still being in a trial-and-error-phase. Future will show in which way the actual heterogeneous demands will remain different and to what point local conditions or political stakeholders can and will influence the development of merely institutional or governmental VREs. It remains an important task to address the different levels of understanding, usage and competencies dealing with digital technologies among researchers. Researchers who still use USB-sticks as only working equipment for the management of valuable research data should be picked up and get enrolled in the process, this is only possible with an adequate offer or services.

Overall, the results show researchers' raising demand for digital solutions in scientific work processes. Future IT solutions like VRE need to respect the interdisciplinary character of scientific work processes and be agile in terms of integration. Yet, the IS discipline, has widely missed the chance to investigate in this ever growing research field and capitalize on the shift so far [2]. The IS discipline is in charge to conceptualize valuable IS investigations which could offer valuable research on this novel research field of RDM.

7. References

[1] D. Siegfried, “Open Science: Das neue Gold von Wissenschaft sind Forschungsdaten,” *Bibliotheksdienst*, vol. 51, no. 5, pp. 454–460, Jan. 2017.

[2] K. Lyytinen, “Data matters in IS theory building,” *J. Assoc. Inf. Syst.*, vol. 10, no. 10, pp. 715–720, 2009.

[3] M. Feijen, “What researchers want,” Utrecht, 2011.

[4] J. Becker, R. Knackstedt, L. Lis, A. Stein, and M. Steinhorst, “Research Portals: Status Quo and Improvement Perspectives,” *Int. J. Knowl. Manag.*, vol. 8, no. 3, pp. 27–46, 2012.

[5] R. C. Amorim, J. A. Castro, J. Rocha da Silva, and C. Ribeiro, “A comparison of research data management platforms: architecture, flexible metadata and interoperability,” *Univers. Access Inf. Soc.*, pp. 1–12, Jun. 2016.

[6] K. Wilms, C. Meske, S. Stieglitz, D. Rudolph, and R. Vogl, “How to Improve Research Data Management,” in *Human Interface and the Management of Information: Applications and Services. HIMI 2016. Lecture Notes in Computer Science*, 9735th ed., Y. S., Ed. Springer, Cham, 2016, pp. 434–442.

[7] D. Rudolph, A. Thoring, and R. Vogl, “Research Data Management: Wishful Thinking or Reality?,” *PIK - Prax. der Informationsverarbeitung und Kommun.*, vol. 38, no. 3–4, pp. 113–120, Jan. 2015.

[8] A. Whyte and J. Tedds, “Making the case for research data management,” *Digital Curation Centre*, 2011.

[9] R. Devarakonda, G. Palanisamy, J. M. Green, and B. E. Wilson, “Data sharing and retrieval using OAI-PMH,” *Earth Sci. Informatics*, vol. 4, no. 1, pp. 1–5, 2011.

[10] R. Higman and S. Pinfield, “Research data management and openness: The role of data sharing in developing institutional policies and practices,” *Progr. Electron. Libr. Inf. Syst.*, vol. 49, no. 4, pp. 364–381, 2015.

[11] DFG Deutsche Forschungsgemeinschaft, “Memorandum Safeguarding Good Scientific Practice,” 2013.

[12] DFG Deutsche Forschungsgemeinschaft, “Leitlinien zum Umgang mit Forschungsdaten,” 2015.

[13] European Commission, “H2020 Programme - Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020,” 2017.

[14] Wellcome Trust, “Policy on data management and sharing.” [Online]. Available: https://wellcome.ac.uk/funding/managing-grant/policy-data-management-and-sharing. [Accessed: 15-Jun-2017].

[15] NSF National Science Foundation, “Data Management & Sharing FAQs | NSF - National Science Foundation.” [Online]. Available: https://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp. [Accessed: 15-Jun-2017].

[16] AGU American Geophysical Union, “AGU Publications Data Policy - Publications.” [Online]. Available: http://publications.agu.org/author-resource-center/publication-policies/data-policy/. [Accessed: 15-Jun-2017].

[17] PLOS ONE - Public Library of Science - accelerating the publication of peer-reviewed science, “Materials and Software Sharing.” [Online]. Available: http://journals.plos.org/plosone/s/materials-and-software-sharing. [Accessed: 15-Jun-2017].

[18] American Economic Association, “Data Availability Policy.” [Online]. Available: https://www.aeaweb.org/journals/policies/data-availability-policy. [Accessed: 15-Jun-2017].

[19] Cell, “Information for Authors.” .

[20] HRK Hochschulrektorenkonferenz, “Wie Hochschulleitungen die Entwicklung des Forschungs- datenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien, Empfeh- lung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel,” 2015.

[21] Alliance of German Science Organisations, “Principles for the Handling of Research Data,” 2010.

[22] RfII German Council for Scientific Information Infrastructures, “Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes and financing for research data management in Germany,” 2016.

[23] European Union, “Riding the wave. How Europe can gain from the rising tide of scientific data,” 2010.

[24] H. A. Piwowar, R. S. Day, and D. B. Fridsma, “Sharing Detailed Research Data Is

Associated with Increased Citation Rate," *PLoS One*, vol. 2, no. 3, p. e308, Mar. 2007.
[25] Y. Kim and J. M. Stanton, "Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis," *J. Assoc. Inf. Sci. Technol- ogy*, vol. 67, no. 4, pp. 776–799, 2012.
[26] M. Joshi and S. S. Krag, "Issues in Data Management," *Sci. Eng. Ethics*, vol. 16, no. 4, pp. 743–748, 2010.
[27] S. H. Koslow, "Should the neuroscience community make a paradigm shift to sharing primary data?," *Nat. Neurosci.*, vol. 3, no. 9, pp. 863–866, 2000.
[28] B. Nelson, "Data Sharing: Empty archives," *Nature*, vol. 461, no. 7261, pp. 160–163, 2009.
[29] C. L. Borgman, "The Conundrum of Sharing Research Data," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 6, pp. 1059–1078, 2012.
[30] H. Bukvova, "Research Proposal : Supporting the Initiation of Research Collaborations," *Jena Res. Pap. Bus. Econ.*, vol. 10, pp. 64–70, 2009.
[31] D. Beaver and R. Rosen, "Studies in scientific collaboration - part I. the professional origins of scientific co-authorship," *Scientometrics*, vol. 1, no. 1, pp. 65–84, 1978.
[32] P. Bryan Heidorn, "Shedding Light on the Dark Data in the Long Tail of Science," *Libr. Trends*, vol. 57, no. 2, pp. 280–299, 2008.
[33] B. Bauer *et al.*, "Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung - Report 2015," 2015.
[34] C. J. Savage and A. J. Vickers, "Empirical Study of Data Sharing by Authors Publishing in PLoS Journals," *PLoS One*, vol. 4, no. 9, pp. 70–78, 2009.
[35] C. Meske, S. Stieglitz, R. Vogl, and D. Rudolph, "Demand for Cloud Services as an Infrastructure in Higher Education Demand for Cloud Services as an Infrastructure in Higher Education," in *Proceedings of the 35th International Conference on Information Systems (ICIS)*, 2014, no. 10, p. 7.
[36] B. Dong, Q. Zheng, M. Qiao, J. Shu, and J. Yang, "BlueSky Cloud Framework: An E-Learning Framework Embracing Cloud Computing," Springer, Berlin, Heidelberg, 2009, pp. 577–582.
[37] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition."
[38] C. Weinhardt *et al.*, "Cloud Computing - A Classification, Business Models, and Research Directions," *Bus. Inf. Syst. Eng.*, vol. 1, no. 5, pp. 391–399, Oct. 2009.
[39] G. Lawton, "Developing software online with platform-as-a-service technology," *Computer (Long. Beach. Calif).*, vol. 41, no. 6, pp. 13–15, 2008.
[40] E. Simukovic, M. Kindling, and P. Schirmbacher, "Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin," 2013.
[41] E. Simukovic, R. Thiele, A. Struck, M. Kindling, and P. Schirmbacher, "Was sind Ihre Forschungsdaten? Interviews mit Wissenschaftlern an der Humboldt-Universit{ä}t zu Berlin." 2014.
[42] R. Hauck, R. Kaps, H. G. Krojanski, A. Meyer, J. Neumann, and V. Soßna, "Der Umgang mit Forschungsdaten an der Leibniz Universit{ä}t Hannover: Auswertung einer Umfrage und erg{ä}nzender Interviews 2015/16." 2016.
[43] "Nutzerbefragung 2014 der SUB Göttingen: Online-Fragebogen (deutsch)."
[44] S. Glitsch and K. Helmkamp, "Nutzerbefragung 2014 der SUB Göttingen Abschlussbericht," 2015.
[45] Niedersächsische Staats- und Universitätsbibliothek, "Nutzerbefragung 2014 der SUB Göttingen Kurzbericht," 2014.
[46] T. Paul-Stüve, G. Rasch, S. Lorenz, and Christian-Albrechts-Universität Kiel, "Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Christian-Albrechts-Universit{ä}t zu Kiel (2014)." 2014.
[47] Philipps-Universität Marburg, "Forschungsdatenmanagement an der Philipps-Universit{ä}t Marburg: Die Ergebnisse der Umfrage zum Forschungsdatenmanagement im November 2014." 2015.
[48] Technische Universität Hamburg and Universitätsbibliothek Hamburg, "Digitale Forschungsdaten f{ü}r die Zukunft sichern: Umfrage zum Umgang mit Forschungsdaten an der TU Hamburg: Auswertung." 2016.
[49] M. Lemaire, Y. Rommelfanger, J. Ludwig, A. Lürken-Uhl, B. Merkler, and P. Sturm, "Umgang mit Forschungsdaten und deren Archivierung: Bericht zur Online-Bedarfserhebung an der Universit{ä}t Trier," no. 2. Trier, 2016.

[50] P. Mayring and T. Fenzl, “Qualitative Inhaltsanalyse,” in *Handbuch Methoden der empirischen Sozialforschung*, N. Baur and J. Blasius, Eds. Wiesbaden, 2014, pp. 543–556.

[51] DataWiz, “ZPID DataWiz - Management psychologischer Forschungsdaten.” [Online]. Available: https://www.zpid.de/index.php?wahl=forschung&uwahl=informationstechnologie&uuwahl=DataWiz. [Accessed: 11-Jun-2017].

[52] H. A. Piwowar, “Who shares? Who doesn’t? Factors associated with openly archiving raw research data,” *PLoS One*, vol. 6, no. 7, 2011.

[53] C. Meske, S. Stieglitz, R. Vogl, D. Rudolph, and A. Öksüz, “Cloud Storage Services in Higher Education — Results of a Preliminary Study in the Context of the Sync & Share-Project in ...,” in *Proceedings of the 16th International Conference on Human Computer Interaction (HCI International)*, 2014, pp. 161–171.

[54] S. Srinivasan, “Risk Management in the Cloud and Cloud Outages,” in *Security, trust and regulatory aspects of cloud computing in business environments*, S. Srinivasan, Ed. Hershey, Pennsylvania: Information Science Reference, 2014.

[55] D. Puthal, B. P. S. Sahoo, S. Mishra, and S. Swain, “Cloud Computing Features, Issues, and Challenges: A Big Picture,” in *2015 International Conference on Computational Intelligence and Networks*, 2015, pp. 116–123.

[56] M. Benaroch and A. Chernobai, “Linking Operational IT failures to IT control weaknesses.” Americas Conference on Information Systems, 2015.

[57] A. Bhattacherjee and S. C. Park, “Why end-users move to the cloud: a migration-theoretic analysis,” *Eur. J. Inf. Syst.*, vol. 23, no. 3, pp. 357–372, 2014.

[58] HRK Hochschulrektorenkonferenz, “Management von Forschungsdaten – eine zentrale strategische Herausforderung für Hochschulleitungen. Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main,” 2014.