



Quadcopters or Linguistic Corpora

Establishing RDM Services for Small-Scale Data Producers at Big Universities

Viola Voß

Universitäts- und Landesbibliothek (ULB), Münster, Germany
voss.viola@uni-muenster.de, orcid.org/0000-0003-3056-407X

Göran Hamrin

Kungliga Tekniska Högskolan (KTH) Library, Stockholm, Sweden
ghamrin@kth.se, orcid.org/0000-0003-4256-2960

Abstract

During an international library conference in 2017 the authors had many productive exchanges about similarities and differences in Swedish and German higher-education libraries. Since research data management (RDM) is an emerging topic on both sides of the Baltic Sea, we find it valuable to compare strategies, services, and workflows to learn from each other's practices.

Aim: In this paper, we aim to compare the practices and needs of small-scale data producers in engineering and the humanities. In particular, we try to answer the following research questions:

- What kind of data do the small-scale data producers produce?
- What do these producers need in terms of RDM support?
- What then can we librarians help them with?

Hypothesis: Our research hypothesis is that small-scale data producers have similar needs in engineering and the humanities. This hypothesis is based on

the similarities in demands from funding agencies on (open) research data and on the assumption that research in different subjects often creates results which are different in content but similar in structure.

Method: We study the current strategies, practices, and services of our respective universities (KTH Royal Institute of Technology Stockholm and Westfälische Wilhelms-Universität Münster). We also study the work and initiatives done on a more advanced level by universities, libraries, and other organisations in Sweden and Germany.

Results: The paper will give an overview of how we did the groundwork for the initial services provided by our libraries. We focus on what we are doing and why we are doing it. We find that we are following in the leading footsteps of other university libraries. The experiences shared by colleagues help us to adapt their best practices to our local demands, making them better practices for KTH and WWU researchers.

Limitation: We restrict ourselves to studying only researchers who create data on a small scale, since the large-scale data producers handle the RDM on their own.

Keywords: research data management; academic support services; university libraries; Sweden; Germany

1. Two Cultures, Two Countries, Two Universities

Let's start with a quiz: here are some questions by scientists of the authors' institutions. Can you guess the discipline of the researchers?

- a) "I am writing a funding grant application and I have to provide information about my research data. Can you help me with that?"
- b) "My colleagues don't want to use Dropbox. Where else can I save 25 gigabytes of data and share them with my colleagues?"
- c) "What do I have to do to follow data-protection regulations during my project?"
- d) "How do I set up a data-management plan?"
- e) "Who has the right to access or reuse my data?"
- f) "Why do I have to publish my data at all?"

Questions a, c, and e might have been questions asked by engineering faculty, b, d, and f by researchers of disciplines in the humanities – or the other way round.

So could it be that these disciplines that like to emphasise that they are completely different are not as different as they may think – at least regarding research data management?

The authors of this paper, librarians at a Swedish and a German university library, discussed this during a library conference in 2017. Since research data management (RDM) is an emerging topic in libraries on both sides of the Baltic Sea, comparing strategies, services, and workflows could help each of the authors to learn from the other.

In 1959 Charles Percy Snow, an English novelist and physical chemist, gave a lecture called *The Two Cultures* at the University of Cambridge (Snow, 1959). He painted a grim picture on the division between “scientists” (including applied scientists and engineers) and “intellectuals” (by which he meant humanists and to some extent social scientists).

This division is quite provocative – which was probably Snow’s intention, as he later (Snow, 1963, p. 53) said: “I hoped at most to act as a goad to action” –, but also quite timeless with regard to many current discussions in university teaching and research.

We will take Snow with us on a walk through the research data management situation in our respective subjects – engineering and the humanities –, countries – Sweden and Germany –, and institutions – The Royal Institute of Technology in Stockholm and the University of Münster –, highlighting important differences and similarities.

On our way we will try to answer questions like “What kind of data do small-scale data producers¹ in engineering and the humanities produce?”, “What do these producers need in terms of RDM support?”, or “What then can we librarians help them with?” – and whether the emerging polarisation between the scientific and intellectual communities that Snow postulated can still be found today.

2. Two Libraries in the Ocean of Research Data

To set the scene for our comparisons, we start with a short overview of the authors’ institutions.

2.1. KTH & KTHB

Kungliga Tekniska Högskolan (KTH),² the Royal Institute of Technology, has developed into Sweden's largest technical university since its beginnings in 1827, located in close proximity to research and innovation clusters in Stockholm. KTH currently hosts around 13,000 full-time equivalent (FTE) students, 3,700 FTE faculty, and around 2,000 PhD students. It fosters research and education with activities in all areas of the engineering sciences as well as in the natural sciences and architecture.

With a history as old as KTH, the KTH Library (KTHB)³ currently services KTH faculty and students on the different KTH campuses with the help of around 50 library workers.

KTHB provides a selection of research support services. For example, the library curates the KTH part of the Swedish publication database DiVA, which enables open access to works in its archives via parallel publishing. KTH-DiVA⁴ is also used by KTH for evaluation and bibliometrics, such as computing the publication output in the Annual Bibliometric Monitoring.⁵ This is possible because the KTH publication policy⁶ states that individual KTH departments are required by KTH management to include metadata for all their publications in DiVA. KTHB staff helps by continuously importing data concerning KTH publications from Web of Science and Scopus into DiVA.

The publication policy also encourages the KTH researchers to publish open access (OA), via parallel publishing in KTH-DiVA or via gold OA directly. To support this, some funding⁷ for processing charges (APCs) for journal articles is provided.

In recent times, several questions concerning the legal and practical implications of the Regulation (EU) 2016/679 (General Data Protection Regulation, GDPR) have been posed to KTHB. But KTHB currently has no formal mandate as a KTH RDM support function. Hence there is room for an expansion and improvement of the KTH(B) research support services.

2.2. WWU & ULB MS

With about 43,000 students, 675 professors, and academic staff of 5,050 the Westfälische Wilhelms-Universität (WWU) Münster, founded in 1780 and

again in 1902, is one of the biggest universities in Germany. Its 15 faculties cover the main scientific disciplines apart from engineering and veterinary medicine; over 120 subjects are taught in over 280 degree courses.

The Universitäts- und Landesbibliothek Münster (ULB MS), founded in 1588 as the library of a Jesuit seminary and transferred to the university in 1780, is the central literature- and information-supply institution for the WWU. A team of 248 colleagues for 182 FTE is in charge of the library system, encompassing one central library and about 100 departmental and institute libraries with a total collection of about 6.25 million printed and digital volumes.

Apart from the usual services of a big university and regional library, the ULB has a long history of services for WWU researchers regarding open access publishing: the MIAMI repository for publishing and archiving digital multimedia documents was set up in 2002. Since 2009, WWU authors have been able to publish their work as part of the hybrid online-and-print series “WWU Münster Academic Publications” that is edited by the ULB. Since 2011 WWU editors of open access journals have been using the ULB’s OJS (Open Journal System) server for their journals, and also since 2011, a fund managed by the ULB has been reimbursing publication fees for articles by WWU researchers published in open access journals that charge APCs.⁸

Expanding the WWU research support services provided by the ULB to the management of research data received “official status” in 2017 with the establishment of a Research Data Service Point and the publication of a WWU Research Data Policy. The steps leading there and the services established on the way and in the future will be presented in chapter 8.

3. RD(M) in Engineering

3.1. Disciplines

In *The Two Cultures* C. P. Snow talks about a distinguishing line between the natural sciences and engineering: “Pure scientists and engineers often totally misunderstand each other. Their behavior tends to be very different [...]”, where pure scientists are said to be left-leaning and engineers conservative

(cf. Snow, 1959). His description is very old-fashioned, and instead we would like to focus on the subjects rather than the researchers.

The engineering (or technological) sciences are similar to the natural sciences but also unlike them in some crucial aspects. The engineering sciences differ according to Hansson (2007) in that: “They (1) have human-made rather than natural objects as their (ultimate) study objects, (2) include the practice of engineering design, (3) define their study objects in functional terms, (4) evaluate these study objects with category-specified value statements, (5) employ less far-reaching idealizations than the natural sciences, and (6) do not need an exact mathematical solution when a sufficiently close approximation is available. In combination, the six characteristics are sufficient to show that the technological sciences are neither branches nor applications of the natural sciences, but form a different group of sciences with specific characteristics of their own.”

Thus we find among the engineering sciences subjects such as material sciences, mechanical engineering, applied physics, variants of chemistry (such as chemical technology, biochemistry, etc.), computer science (such as human-computer interaction (HCI), wireless communication, mechatronics, etc.), and many subjects which are multidisciplinary or blends of older subjects (such as medicine, technology, physics, chemistry, biology, mathematics) with at least some applied ingredient.

3.2. Data Characteristics and Usage

Research data collected in engineering are, like in the natural sciences, mostly *quantitative* and of the *ordinal* type, in that data entries can be compared with each other and ordered. Data sets can in principle be viewed as a large set of vectors, which you can analyse, compute upon, and visualise.

The engineering sciences can be defined as *the systematic study of reality via the construction of technological artefacts*. This differs from the natural sciences by the introduction of an extra, “artefactic”, level between the researcher and the reality studied.⁹ In this definition, research data collected will concern the *qualities* of the technological artefacts studied.

Three common examples of data are the following: fluid-mechanics data, computing data, and geopositioning data.

Fluid-mechanics data are collected when measuring the performance of gases or fluids around solid objects. These data can then be used in simulations and computational modelling for improving engineering design.

Computer programs which are resource-intensive have to be subject to practical performance estimates, before being used in industrial applications. This is a different process than estimating the theoretical complexity of algorithms, as done in mathematics. One example is High Frequency Trading, where it is vital to be fast enough. Many times optimisation for better user experience is important, for example effective ping in gaming applications. A more extreme example is cost reduction for large web applications, such as Twitter, which cut down 11 % on CPU use just by switching Java compilers.¹⁰

Any research object that is constructed to be moved around in 3D-space must be designed with an awareness of the physical forces that it is subject to, be it forces working in close proximity to the object (such as mechanical strains) or working from a distance (such as gravitational forces).

In, for example the construction of flying vehicles such as quadcopters (which normally refers to unmanned small aircrafts with four rotors, thus a particular kind of multicopter or drone, which can be autonomous or controlled by an operator), all the above kind of research data can or should be collected during the construction process. The aerodynamic properties of the aircraft are fluid-mechanics data, the ability of the steering-system software to adjust fast to input from sensors represents computing data, while this sensor input includes geopositioning data collected from gyros on the aircraft.

These examples indicate that most parts of the engineering sciences are empirical and highly data-driven research, which seemingly can be conveniently stored and subsequently accessed and computed on.

But this is not to say that all research data are “uncomplicated”, and that there are no challenges in engineering RDM. For example, some engineering science data are *sensitive*. With the fast-moving possibilities of data collection via GPS or other tracking technologies, some data are collected in engineering projects that involve human subjects. Any researcher working in traffic engineering will naturally have to consider that the collection of data on traffic flows via CCTV coverage will intrude on people’s privacy when linked with publicly accessible owner-vehicle registries. This can apply to the quadcopter

example, if you equip a quadcopter with a camera and visual recognition technology. As another example that has become apparent in recent years, any researcher in HCI studying the aggregated behaviour of human interaction with computers may have an integrity problem that can be exploited if that data is made openly available.

3.3. Initiatives Working on RDM

With the recent funding mandates, such as the ones from research funders such as the Swedish National Research Council (“Vetenskapsrådet”, VR), the EU Horizon 2020 requirements, etc., a need has emerged among all researchers to not only ensure long-term storage but also to archive and publish their data. Hence the still used “store the stuff on your hard drive” solution is no longer compliant with funders’ requirements.

R3data, the global registry of RD repositories, lists 22 results for Sweden. A few of these can be said to have some relevance in engineering, as they collect data in the atmospheric or environmental sciences or in life sciences.

The canonical example here is “The Human Protein Atlas”¹¹ project, which aims to investigate the human proteome. The project has its origin in research conducted at KTH with Mathias Uhlén as its principal investigator. The data volume(s) created in the different sub-projects are huge. For example, in the latest release of the atlas you can find more than 26,000 antibodies which can be used to target proteins from almost 17,000 human genes, which is said to correspond to ~87% of the human protein-coding genes.

Further international initiatives relevant for Swedish engineers can easily be found. One example is the Integrated Carbon Observation System (ICOS), which “is a European research infrastructure to quantify and understand the greenhouse gas balance of the European continent and of adjacent regions.”¹²

At CERN, Swedish high-energy physicists have been collecting particle-physics data. These data have for a long time been stored in the previously mentioned CERN Data Centre.

All these example repositories are primarily for large-scale data producers or domain-specific, hence they are perhaps not the natural choice for

data archiving on a small-scale national or single institutional level. One can then state the question whether the national publication database DiVA could be used for archiving the datasets? The answer is that the functionality of DiVA¹³ is not developed for archiving datasets at the moment.

That can be one argument in favour of the national initiative Swedish National Data Service (SND), which is a general research data access and support hub for researchers in the humanities, social sciences, and health sciences. It is structured as a consortium, with many Swedish higher education institutions taking part as members on different levels.

The future of the SND includes an expansion of its services for the natural sciences, and here lies one possible solution for RDM in engineering. But to realise this, a large and distributed storage solution has to be implemented. Such a solution has been proposed by the SUNET organisation, which has been providing IT services for research and higher education in Sweden since 1984. The distributed or federated solution would be constructed as approximately 2 to 5 storage facilities affiliated with different institutions, each with a storage capacity in the range of 10 petabytes of data.

For ongoing KTH projects, researchers have had the option to store data at the Swedish National Infrastructure for Computing (SNIC).¹⁴ Funded by the Swedish National Research Council and 10 Swedish universities, SNIC is a national research infrastructure which aims to provide resources and user support for data storage and computing for all scientific disciplines. This service is provided via six partner centres in Sweden, one of them being the KTH PDC Center for High Performance Computing.¹⁵ For active research projects, it is possible to get up to 200 TB for 4 to 5 years by applying for the SNIC resources through the SNIC User and Project Repository (SUPR).¹⁶

At KTHB, we are lucky to have staff in-house with subject qualifications relevant for the engineering sciences. Besides expertise in library and information science, in particular a concentration in bibliometrics, we have skills in chemistry, biochemistry, ecology/conservation, mathematics, and computer science via a number of staff with own research experience. This provides good conditions for establishing an RDM support unit at KTH, to which we will return in chapter 7.

4. RD(M) in the Humanities

4.1. Disciplines

Questions like “What are the humanities?”, “When and where do they begin to be digital humanities?”, “What is research data in the ‘traditional’ humanities?”, and “What is it in the digital humanities?” each fill dozens of papers and book chapters. For the purpose of this paper, we will concentrate on the “pragmatic side of things”: who are the researchers that are contacting the library with questions regarding their data, and what kind of data do they bring?

As a very pragmatic first step, we here define “humanities” in the tradition of the German term “Geisteswissenschaften”¹⁷ as “all non-natural-science disciplines” (i.e. “Naturwissenschaften”), thus covering not only philologies, but also arts, history, law, theology, and social sciences. As such “the humanities reveal probably the most diverse mix of data in the entire academic faculty” (Peukert, 2017, p. 235). Regarding some aspects of RD(M), it may however be useful to distinguish the humanities not only from the natural sciences and engineering, but also from social sciences, law, and economics.

The “lowest common denominator” of the definitions of *digital* humanities (or e-humanities or computational humanities) – be they a discipline on their own, an “auxiliary discipline”, a method, or a way of thinking – is the use of computer-assisted methods and digital(ized) resources and the reflection of these uses, thus taking the “traditional” or “mainstream” humanities closer to computer science.

4.2. Examples for DH Research

Three projects give a short glimpse in current German research in digital humanities and challenges they face with regard to their research data.¹⁸

A project at the University of Leipzig, “Music information retrieval for handwritten folksongs”,¹⁹ aims to analyse handwritten music scores from a large collection of German folk tunes. To be able to find e.g. similarities or regional patterns in the melody of different tunes, the scores had to be digitised and transcribed as machine-readable music. As tests of optical music

recognition software had shown that the precision of these tools is still quite low for handwritten music, the project decided on a crowdsourcing approach – and faced the next problem: there are frameworks for the “normal” transcription of music sheets but not for a crowdsourcing variant. Together with the University Library of Regensburg (the owner of the folk song collection to be analysed), the project developed a platform for collaborative music transcription.²⁰

Connecting archaeological findings to geographical information and presenting the combined data online was the scope of a joint project by the WWU’s Institutes of Egyptology and Coptology and Geoinformatics.²¹ The archaeologists had to note geospatial information during excavation projects to map the structure of the sites and their findings. The resulting platform²² brings together these data with additional information like aerial photography of the sites and photos of the excavated objects, making it possible to search and browse the findings according to different categories. The platform can be used by all WWU archaeologists for publishing their present and future findings and to enable exchange of data with scientists – also citizen scientists – anywhere else in the world.

Meanwhile WWU historians are working on medieval heraldry from the perspective of cultural history. They are asking how coats of arms became an omnipresent and strong means of communication, thus contributing to the understanding of the complex period of the Late Middle Ages.²³ For this purpose, different sources like images, artefacts, architectural information, and texts have to be made available and then reassessed for a combined analysis, bringing together history, art history, and medieval philology for interdisciplinary discussions of semiotics and visual culture. One of the project’s outcomes will be a machine-readable ontology of coats of arms that enables the description, documentation, retrieval, and processing of relevant data. With these methods the “auxiliary science” of heraldry can be enhanced on a large scale.

4.3. Data Characteristics and Usage

While the *research lifecycle* in the humanities does not differ too much from research in other disciplines, *research data* in the humanities often differs in several aspects regarding the type of data and its usage.²⁴

It starts with the fact that humanists who have “grown up” in the traditional era of humanities “would not easily speak of their objects of study as ‘data’”, as Schöch (2013, p. 2) remarks: “they would rather speak of books, paintings and movies; of drama and crime fiction, of still lives and action painting; of German expressionist movies and romantic comedy. They would mention Denis Diderot or Toni Morrison, Chardin or Jackson Pollock, Fritz Lang, or Diane Keaton. Maybe they would talk about what they are studying as texts, images, and sounds. But rarely would they consider their objects of study to be ‘data’.” It is only when they are asked about the materials they generate and collect for their research – “yeah, I guess that’s data” (Thoegersen, 2018, p. 499) – or when some form of digitisation sets in that these objects might be perceived as data – and then the potential they offer for humanist analyses may be recognised as not only entraining technical, but also theoretical, methodological, and social issues.

Research in natural or (quantitative) social sciences is mostly based on data from measurements or surveys, while the humanities work on (digital) representations of cultural artefacts like texts, images, sheets of music, audio or video recordings, or other physical objects. And while measurements and surveys mostly lead to structured data, the data derived from research objects in the humanities is often only modelled during the research itself – by digitising, describing, sorting, annotating, visualising, and interpreting the data. As the data can then reflect different perspectives on a topic and can be saved in different formats and in different levels of aggregation, they are mostly very heterogeneous regarding content and structure. Apart from this varied content, there are also external dimensions like “historicity” and “context” that are important for the interpretation.

Research in the humanities is often conducted in open, non-linear research cycles, thus the different steps of the research process are often not as separate as they are in other disciplines. The data can constantly be extended (“I have found another loan word for my collection”), refined (“I have discovered more information about these two painters in a new publication”), or enriched (“I decided to tag another feature for my statistical analysis”), the data can be reused in different contexts and connected to other data. This makes a differentiation between data “levels” like raw/primary data or processed/secondary data difficult, and this leads to an inherently dynamic character for which the usual storage types are not very well suited.

At the same time, many digital representations of objects have a continuous relevance for research, especially in the case of unique items like manuscripts or image/audio/video documentations of historic events. These resources have to be kept safe for longer than a certain period such as the ten years suggested by some funders.

While in the natural or the social sciences there are often complete datasets or data collections that could be considered for reuse, the database for research projects in the humanities can be based on single objects in one collection or on a set of objects distributed over several collections. The presentation or publication often happens via websites based on databases and scripts, and for a reproduction of this research, the data would have to be put together in the same way as used in a specific project, annotated in the same ways and analysed with the same set of tools with the same configuration. This is why RDM in the humanities not only has to consider the data as such, it also has to factor in the software environment for the processing, analysis, and presentation.

This leaves us with a complex situation: diverse types of data arranged in different layers during the research process, linked to other data types and sources and “corralled” in specific technical settings that have to be kept as “living systems” to make a “useful” reuse or reproduction possible.

As if this were not enough, there is also the dimension of legal questions: the different parts of the data collection for a project are often subject to several different copyrights and property, exploitation, or personal rights. Without the consent of all rights holders involved, publication of the data is not permissible – the more complex the data collection, the more difficult this aspect.

4.4. RDM Standards

When it comes to defining requirements for the handling of research data, national and international funding institutions often expect applicants to adhere to “standards and rules used in the specific disciplines or communities”. But when you start looking for them in the humanities, you will find that there are only very few disciplines that have issued standards or rules for their data (cf. e.g. Schirnbacher, 2017, pp. 398–399). One example is e.g. the “Austin Principles for Data Citation in Linguistics”.²⁵

This probably does not come as a surprise, as we have seen how varied the data landscape in the humanities can be. Apart from the different contents, there are also organisational questions about *who* should be responsible for developing and “running” these standards or about *how* it should be done (top-down or bottom-up processes?; open community or institutional platforms? etc.).²⁶

Given the fact that most researchers from the humanities cannot fall back on set standards for their data, establishing a Data Management Plan (DMP) or meeting funders’ requirements can be more complex than in other disciplines. Instead of using a “package solution” they have to combine recommendations e.g. for single data types or look out for data management recommendations published by researchers or research projects in the respective disciplines.²⁷

4.5. Initiatives Working on RDM

Several interest groups assemble people active in the digital humanities for promoting and supporting research, collaboration, publication, teaching, and training. They collect and disseminate news and information about DH, organise conferences, they run journals, book series, publication platforms, and discussion mailing lists, and they keep contact to other institutions.

There are e.g. the international Alliance of Digital Humanities Organizations (ADHO²⁸) (with the Special Interest Group for “Libraries and Digital Humanities”), the European Association for Digital Humanities (EADH), or the group “DHd” (“Digital Humanities im deutschsprachigen Raum”²⁹) covering German-speaking people active in DH.

One topic of their activities is of course RDM in DH. The ADHO has e.g. set up a liaison with the Research Data Alliance, and one of DHd’s recent publications is a paper on data centres for DH (DHd AG Datenzentren, 2018).

4.6. Data Tools and Infrastructures

There are at least as many software tools used in the humanities as there are data types or formats. As they can have a big influence on how data is

collected, annotated, or analyzed during a research project, collecting and disseminating knowledge about these tools is important.³⁰

To prevent that every project has to build its own infrastructures composed of different tools, initiatives like DARIAH (“Digital Research Infrastructure for the Arts and Humanities”) or CLARIN (“Common Language Resources and Technology Infrastructure”) offer sets of services with tools for collaboration or project management, virtual machines with ready-to-use software components, or repositories for storing research data safely.³¹

A good solution for storing research data are data centres that make sure that the data will still be available when the life cycles of research projects, data formats, or software have come to their ends.

As mentioned above, the DHd Working Group on data centres has published a comprehensive compilation of goals, tasks, and services these centres should offer and which types of organisational models are possible. Two aspects from the perspective of the humanities are the heterogeneity of the data and the blurred lines between input data, working data, and output data that entail many standards and rules which are relatively open and flexible. (The many variants of the use of the Text Encoding Initiative (TEI) standard can serve as an example.) Data centres for the humanities have thus to be differentiated from repositories built for homogeneous types of data or for systems concentrating on bitstream preservation: they are more like “living archives” that not only give access to single data objects, but that also keep generic and project-specific software systems running and that also provide active research data management once a project has ended.³²

Apart from repositories or data centres, there are also specialized data journals emerging in the humanities, like the Research Data Journal for the Humanities and Social Sciences (RDJ³³), the Journal of Open Humanities Data (JOHD³⁴), or the Journal of Open Archaeology Data (JOAD³⁵).

5. RDM in Sweden

Sweden is a small country with little or no federalist tendencies. It has around 50 universities or colleges of higher education, many of them are quite small

and distributed over different parts of Sweden, as a consequence of regional politics. This gives the individual institutions some freedom.

But since most of the universities are managed by the Swedish state, and all those are government agencies which have to follow Swedish law, in particular the “Lag (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning”³⁶ which was installed in April 2018, the differences between the institutions’ organisations tend to be minor.

There is also room for centralized initiatives on a national level.

The overview of RDM activities in Sweden will follow the three dimensions of politics of science, organisation, and technical aspects (cf. Schirnbacher, 2017).

5.1. Politics of Science: Papers and Recommendations

On the political level, RD and RDM have to be discussed regarding infrastructures, legal conditions, and financial frameworks (cf. Schirnbacher, 2017, p. 393).

In 2015, the Swedish government gave the Swedish National Research Council (“*Vetenskapsrådet*”, VR) the commission to investigate and formulate guidelines on the open access to scientific information (in accord with EU initiatives on Open Access). VR functions as a research funding agency and a research policy maker and provides general support for the Swedish government. This makes VR a central player in forming Swedish research policies.

The report “*Förslag till Nationella riktlinjer för öppen tillgång till vetenskaplig information*” (Vetenskapsrådet, 2015) contains a chapter on open access to research data. VR outlines the guidelines which are based both on what is stipulated in Swedish law (relevant laws are SFS 1949:105, SFS 2009:400, SFS 1998:204, SFS 2003:460 or SFS 1990:782, which regulate issues like the freedom of speech, the Swedish principle of open access to agency information, personal data, ethics and archive laws), as well as views from the EU commission and Horizon 2020 statements.

VR suggests that open data should be the norm, and it also highlights three points from the Swedish public office tradition. First, that RD are

public material if created by an agency and hence not owned the individual researchers. Second, that all RD are open, in that any individual may request them from the agency, and the agency has to comply (with the exception of RD involving human subjects, to which access is restricted). The third point VR makes is that it is clear that responsibility for the long-term storage and archiving of RD lies with the universities, not the researchers.

Philosophically, VR advocates for open public research data based on arguments of democracy, transparency, synergistic effects for research, innovation, and other uses outside of research, as well as for the future evaluation of research via citation analyses on data sets.

Economically, VR makes the case for open RD by implementing this in their funding requirements. Other funding providers have followed suit, such as the FORMAS policy of open data (Formas, 2016).

The guidelines are at the moment (2015–2020) restricted to only apply to some pilot projects, and then only the RD in publicly funded research *which have resulted in a publication*. VR suggests that in the future all data from publicly funded research shall be openly available.

A side note on this matter: in May 2018, the Bibsam consortium of Swedish universities, led by the National Library of Sweden (“Kungliga Biblioteket”), cancelled the subscriptions to Elsevier’s journal packages after negotiations had failed to reach an agreement. The press release from Kungliga Biblioteket states that the reason for cancellation is that Elsevier has failed to supply an offer which would fulfil the aim stated by the Swedish government that all publicly funded research should be published open access immediately from 2026. This decision by the Swedish universities can be viewed in the larger context of aiming to make all research openly available.

5.2. Organisational Dimensions

The organisational dimension of RDM takes into account the cooperation of the different players and asks e.g. which structures are needed or how the different tasks can be efficiently distributed between the infrastructural institutions and the scientists (cf. Schirmbacher, 2017, p. 396).

RD and RDM support in Sweden are handled by different initiatives depending on university, subject, etc. Since on the national level the Swedish National Data Service (SND³⁷) is a national research infrastructure funded by VR and a consortium of seven universities, SND's main goals are to support the sharing, archiving and reuse of RD and related material as well as to serve the research community by receiving metadata descriptions and RD and organising their secure long-term storage. SND is also a communication platform, and it also helps research communities by providing tutorials and learning support.

SND claims it is an initiative with the intent of being a hub for assisting all Swedish researchers, although it currently services mostly researchers in medicine, natural sciences, or humanities. RD submitted to SND does not automatically become openly available via SND, as researchers have the right to decide who is going to be allowed to access their data.

As for Swedish individual institutions, the situation differs strongly. Some universities have developed their own repositories, such as the Tilda repository at the Swedish University of Agricultural Sciences (SLU). Tilda was built to archive and publish environmental and climate data from SLU researchers. It standardises routines for handling data, offering a manual interface for adding metadata, connecting publications with datasets, and it also supplies a communication interface for external sources. In the future, Tilda aims to provide automatised metadata enrichment and also connect the metadata with publications in the SLU publication database SLUpub and the underlying archived datasets.

The SLU DCU ("Data Curation Unit") claims that this will fulfil requirements made by Swedish law and by funders as well as increasing the visibility of research, both on the individual and institutional level. It is however unclear whether the Tilda archive currently fulfils these demands or principles such as FAIR.³⁸

Other universities have taken a different approach to RDM. They have first built a broad support organisation and combined it with a university RDM policy at a later time.

This has been done for instance at Stockholm University (SU) with its Research Data Policy.³⁹ As part of Sweden's largest university, the SU RDM

support organisation is naturally highly developed. It currently has a full-time staff of one coordinator, two analysts, and additional part-time staff who work in e-archiving, legal, and IT or as research secretaries. This support organisation started its development organically by re-allocating existing resources inside the library. It is an excellent example of how you can first develop a service on a small scale and only market and provide support on a larger scale to faculty after having gained support from the highest levels of university management.⁴⁰

Although this approach fails to give the SU researchers a “default” option for archiving and publishing their data, it still provides well-developed support for different questions. The SU RDM support web page claims that that SU aims to follow the FAIR principles and that the official SU position is that RD should not be archived with commercial vendors. But the SU researchers also have the option to store their data via Figshare services at su.figshare.com.

There are few collaborations between the Nordic countries and none important for small-scale data producers. There are however initiatives for large-scale producers, in particular when it comes to cooperation in the biomedical field. The joint Nordic Tryggve2⁴¹ initiative, for instance, will offer the ability to transmit sensitive data. The vision is to “develop secure services that enable large-scale biomedical research studies across countries.” (Pursula et al., 2018)

5.3. Technical Dimensions

In the technical dimension, issues like requirements for RD, RD metadata, or the storage and presentation of RD have to be considered (cf. Schirnbacher, 2017, p. 398).

As we covered earlier in this section, there have been multiple suggestions on how engineering data should be stored for the long term. We can compare the solutions proposed by SND via SUNET with the solutions given at SNIC. For scientists already storing their current projects at SNIC, there seems to be no reason to transfer their data to another server when the project has been completed. Instead one has to think of solving the financial issue of who is going to pay SNIC for long-term storage.

While long-term storage may seem like a simple problem of a more practical nature, the long-term archiving and, if needed, also publishing of RD is technically more challenging. Here SND can play an import role. SND has a number of metadata coordinators (“datasamordnare”) who can assist researchers when marking datasets with relevant metadata. So, a reasonable solution would be to describe your data via the SND metadata catalogue according to legal and funding requirements and then to only provide a link to the original data set. SND is committed to following the FAIR principles, which speaks for it being a repository/hub with high quality.

Unfortunately, the SND metadata coordinators have domain expertise mostly in the humanities and the health sciences. As such it is currently not particularly well suited for assisting engineering scientists.

5.4. Conclusions

We need more cooperation for RDM inside Swedish engineering sciences. With so many higher-education institutions doing research in technology, the situation almost demands that institutions work together when organising support functions. The development of common metadata standards in engineering and FAIR-complying repositories also further tasks.

Swedish research funders have long been advocating open access and we now also see them including statements on open data. This is a good driving force to open RDM.

Nevertheless, it is hard to see how good RDM support can be organised, if we do not involve the researchers in the fields, asking them what their needs are. For example, we need to have institutional policies for RDM that make clear the obligations that researchers have for (open) RDM and also take into account discipline-specific needs.

Reflecting on this and the good practices described earlier in this chapter, we relate these thoughts to two observations by Sesartic and Dieudé.

First, that “[i]n providing free expertise and consulting services to researchers, our team is building a solid basis for mutual respect, trust and exchange of best practices. The very fact of having dedicated persons willing to explore

with them the best options in their particular scenario while sharing their expertise, networks and knowhow proved to be very beneficial not only for the researchers and their lab, but also for the research institutions.” (Sesartic & Dieudé, 2017, p. 8)

And second, when speaking on the lesson learned during their work progress, they note that “RDM training leads to enhanced collaboration with scientists and better visibility for the libraries, as well as improving their image”, that “[c]onstructive collaborations are key (win-win approach): among colleagues from the library, among different sets of services within the institutions (including the library, IT, Research Office, and legal experts to name a few) and between institutions nationally and internationally”, and that “[c]entralized and harmonized communication, sensitizing actions and quality support service are key to build momentum and trust while changing the image of the library and of librarians” (Sesartic & Dieudé, 2017, p. 10).

Therefore, we urge increased RDM support by the KTH(B) – both for the value it brings to KTH and its individual researchers and for the increased visibility it brings to KTHB. We would also like to prioritise increased cooperation between KTH and other Swedish technological universities as a way to further advance our work in the engineering sciences. And KTH has to monitor the development inside the SND, which may mean that KTH has to become a full SND consortium member in the future.

6. RDM in Germany

Germany is a big federalist country with 16 states, 17 Ministries of Education and Research, over 425 universities or colleges of higher education, many big and small scientific organisations, 6 regional library consortia, and several nationwide or regional alliances and initiatives – it comes as no surprise that RDM in Germany, like other comprehensive topics, is dealt with in many a place, constellation, or context. So this chapter can only give a rough overview.

6.1. Politics of Science: Papers and Recommendations

In 2010, the Alliance of Science Organisations in Germany⁴² (Allianz der deutschen Wissenschaftsorganisationen, 2010), a union of the most important

German research organisations issuing statements relating to research policy and funding and the structural development of the German research system, put out principles of the management of research data, defining it as a strategic task for science, politics, and society and asking for a coordinated approach.

The German Council of Science and Humanities (Wissenschaftsrat⁴³), composed of scientists from different scientific institutions, provides advice to the federal and the state governments on the structure and development of higher education and research, considering societal, cultural, and economic contexts in Germany, Europe, and on an international level. In 2012, based on former papers by other committees that had worked on infrastructural questions, the Wissenschaftsrat published recommendations about the development of scientific information infrastructures, with RD as one important aspect (cf. Wissenschaftsrat, 2012).

As one result of these recommendations by the Council of Science and Humanities, a Council for Scientific Information Infrastructures (Rat für Informationsstrukturen, RfII⁴⁴) was established in 2014. It focuses on the strategic development of a contemporary and sustainable infrastructure for access to scientific information as part of the federal government's "Digital Agenda for Germany". In 2016 the RfII proposed the establishment of a national RD infrastructure (Nationale Forschungsdateninfrastruktur, NFDI), encompassing all scientific disciplines and communities and coordinating and governing RDM activities (cf. RfII, 2016, 2017, 2018).

The German [Universities'] Rectors' Conference (Hochschulrektorenkonferenz, HRK⁴⁵), recognising the importance of the topic, installed a committee for digital infrastructures that published recommendations especially for universities (cf. HRK, 2014, 2016).

Meanwhile, in 2013 one of the biggest research funders, the Deutsche Forschungsgemeinschaft (DFG), published a memorandum for "good scientific practice", proposing e.g. that primary data should be "securely stored for ten years in a durable form in the institution of their origin" (DFG, 2013, p. 74). This set a standard e.g. for grant applicants regarding RDM planning. DFG recommendations about RD date back to 2009, giving advice on the secure storage and provision of primary RD (cf. DFG, 2009), while a first version of the recommendations for good scientific practice had been published in 1998.

In February 2018, the Alliance of Science Organisations in Germany found that the discussion about RD, RDM, and research infrastructure had been very lively in the last years: there have been developments regarding research data infrastructures, and, with the NFDI, there are efforts to build a national data infrastructure. The Alliance gives five recommendations to enable a “working digital science and research landscape” in Germany, aiming at developing digital strategies for all research institutions, building a national RD infrastructure system, and establishing RD and RDM in general and specific courses and training (cf. Allianz der deutschen Wissenschaftsorganisationen, 2018b, p. 4).

Several institutions in the state of North Rhine-Westphalia formed an initiative working on the national infrastructure NFDI from the state institutions’ perspectives. This “Landesinitiative NFDI”⁴⁶ collaborated with an RDM expert group of the initiative “Digital University North Rhine-Westphalia” for a policy paper on the universities’ roles regarding the national infrastructure issued in April 2018 that encouraged institutions to think about ways of participating in the process of building the NFDI (cf. Curdt et al., 2018).

RDM is, of course, also a topic for the section “Academic Universal Libraries” of the German Library Association. In February 2018, it published a strategic paper “forecasting” the situation of academic libraries in the year 2025. RDM is the third of eight fields of activities. The paper assumes that libraries, having adapted their organisational structure to the demands of RDM services, will e.g. work on the further development of national and international standards while providing technical infrastructures and services in collaboration with computer and data centres (cf. DBV Sektion 4, 2018, p. 13–15).

The three latest additions to the collection of statements and recommendations date from May 2018. The German Initiative for Network Information (Deutsche Initiative für Netzwerkinformation, DINI⁴⁷), a network of academic libraries, media centres, IT centres, and scientific societies working in the field of infrastructure, published “Theses about the future information and communication infrastructure”. The 11 2018 theses are an update to those published in 2008, based on former considerations about infrastructure questions. They cover the four main areas “digital transformation”, “openness”, “research”, and “learning and teaching”; RD(M) is dealt with in “openness” and “research”, asking for open access to scientific data and publications and postulating that RDM has to be dealt with on every level of academia, while the retrievability and reusability of data require a sustained metadata

concept (cf. DINI, 2018). In a paper by the DFG discussing how to strengthen the system of academic libraries in Germany, RD and RDM are identified as an important field for libraries (cf. DFG, 2018). Furthermore, “nestor”,⁴⁸ a network of expertise in long-term storage of digital resources in Germany with members from different institutions dealing with digital preservation, published a statement on the organisational structure of the national infrastructure NFDI (cf. nestor, 2018).

6.2. Organisational Dimensions

While all these papers agree on the importance of RDM and the necessity of a structural approach, there are various realisation approaches and different implementation efforts.

6.2.1. National, regional, or discipline-specific initiatives and networks

Some initiatives are working on a nationwide level, aiming at services for all disciplines.

By organising conferences and training, RDA Deutschland,⁴⁹ a German association affiliated with the Research Data Alliance (RDA), wants to promote the exchange and reuse of data.

The initiative DINI, mentioned above, started the “AG Forschungsdaten”,⁵⁰ a working group on RD, together with the nestor network, which is dedicated to long-term storage and has members from different institutions dealing with digital preservation.⁵¹ The working group wants to encourage the interdisciplinary and inter-institutional exchange of experiences and to coordinate RD activities in German-speaking countries.

One joint project is Forschungsdaten.org,⁵² a central wiki collecting information about RD and RDM activities in Germany and beyond, grouped in categories like Education and Qualification, Data Publishing, Metadata, Policies, Projects, Software and Technology, or Networking.

Another information platform is Forschungsdaten.info,⁵³ initiated by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg

and maintained by several universities in this state. It provides general introductory information about RDM (Planning, Organisation and Work, Rights and Responsibilities, Preparation and Publication, Maintenance and Lasting Use) and lists ongoing e-science projects in Baden-Württemberg funded by the Ministry.

Information about the field of copyright and licensing of RD can be found at Forschungslizenzen.de.⁵⁴

One of the organisations behind this site is DARIAH-DE,⁵⁵ part of the European network DARIAH (“Digital Research Infrastructure for the Arts and Humanities”). This project aims to build an infrastructure for the humanities and cultural sciences that work with digital resources and methods. It provides various tools, services, and network activities for teaching, research, and research data. Coordinated at the State and University Library in Göttingen, DARIAH-DE is a collaboration of 19 institutions from the humanities and cultural sciences and from information technology, including universities, libraries, computer centres, non-university research institutions, academies of sciences.

Another German “branch” of a European initiative – and cooperating with DARIAH-DE – is CLARIN-D.⁵⁶ CLARIN (“Common Language Resources and Technology Infrastructure”), launched in 2012, is a network for archiving and processing language-related resources in the humanities and social sciences. CLARIN-D is a national network of nine certified centres at German universities, the Institute of German Language in Mannheim, and the Max Planck Institute for Psycholinguistics in Nijmegen (Netherlands). Each centre specialises on a certain type of data or service, e.g. corpora for different types of German language data, language statistics, multimodal data, tools for phonetics, or software for computer linguistics.

The aforementioned Landesinitiative NFDI (LNFDI) wants to coordinate and steer RDM activities in the state of North Rhine-Westphalia, acting as a hinge between the government initiating the national RD infrastructure NFDI and the universities and other research institutions in the state. The LNFDI has e.g. developed sample policies and guidelines for these universities, and it offers consultancy and regular meetings or talks and workshops for the infrastructure institutions in the state. A similar initiative for the state of Bavaria is “Forschungsdatenmanagement Bayern”.⁵⁷

6.2.2. National, regional, or discipline-specific recommendations and handouts

To raise scientists' awareness of the topic, the Alliance of Science Organisations in Germany published an introductory brochure (cf. Allianz der deutschen Wissenschaftsorganisationen, 2018a) – as did the Landesinitiative NFDI (cf. Curdt et al., 2016), the virtual research infrastructure project WissGrid (cf. Ludwig & Enke, 2013), and several other initiatives. The Alliance has e.g. also published recommendations on the development, use, and provision of research software (cf. Katerbow & Feulner, 2018).

While the information shared by DARIAH-DE and CLARIN-D concentrates on research and RD in the humanities and social or cultural studies, there are several other societies or institutions working on information material for a specific “clientele”. To name just a few: the Rat für Sozial und Wirtschafts-Daten, an advisory council to the federal government for developing a RD infrastructure for empirical social and behavioural sciences and economics, has published working papers⁵⁸ on RD topics, the German Psychological Society has specified the general DFG guidelines for RD in psychological science (cf. Schönbrodt, Gollwitzer & Abele-Brehm, 2016), and the DFG has given recommendations for RD for studies in biodiversity.⁵⁹

6.2.3. Local initiatives and activities

According to López (2015, p. 7) there have been two trends among German universities: demands from specific disciplines or projects leading to the development of central infrastructures serving all disciplines, or centralised formal guidelines leading to the development of central services. In many cases it may have been a mixture of both: “bottom-up” activities starting from projects plus “top-down” activities from conceptual/strategic approaches.

Several universities have issued policies⁶⁰ and information websites and are offering consultation and training for their scientists and/or students.

The following list focuses on the universities that are a kind of “role model” for the services developed or planned for the WWU, presenting some aspects of their activities.

- The university library of Bielefeld is very active in the field of open access;⁶¹ it has developed a broad range of RDM services following a pilot project that determined the local needs and demands.⁶²
- At the University of Göttingen, RDM and Digital Humanities have been on the agenda for several years, and the university library is a strong partner in the Göttingen eResearch Alliance⁶³ and in the Göttingen Centre for Digital Humanities (GCDH⁶⁴), both working on RDM, among other services.
- In April 2017, Hamburg University opened a “Center for sustainable RDM”,⁶⁵ offering services and workshops for students and faculty.
- The RDM Service Team of the University of Hannover⁶⁶ has developed a concept based on four types of activities (cf. Meyer, Neumann & Soßna, 2017) and compiled an extensive website about RDM.
- The team for the Cologne Center for eHumanities (CCeH⁶⁷) and the Data Center for the Humanities (DCH⁶⁸) at the University of Cologne is over 30 members strong (most of the positions funded by projects). A key area of their activities is to keep RD “alive”: not just store away the data, but keep the applications used to search, present, or visualise the data up and running. The DCH will carry out a special project for this together with the Institute of Architecture of Application Systems at the University of Stuttgart.⁶⁹ For the university’s activities regarding RDM in general see e.g. Dierkes and Curdt (2018).
- The University of Trier has a Service Centre eSciences (SeS⁷⁰) offering services for RDM and working on the development of a research infrastructure for the university. For this, the centre cooperates e.g. with the University Library, central IT services, and the Trier Center for Digital Humanities.⁷¹ They have e.g. developed “FuD”,⁷² software for a virtual research environment for the humanities and social sciences that wants to help collect, analyse, edit, publish, and preserve RD.
- The corresponding centre at Tübingen University is also called “eScience Center”,⁷³ and the Digital Humanities are also a key area here. While other centres that have emerged from a DH context are “text-oriented”, offering e.g. annotation tools for digital text editions, space- and time-oriented subjects like art history or archaeology were added to the list of relevant disciplines in Tübingen. Tübingen is part of the CLARIN-D network with special expertise in annotated corpora (treebanks), lexical data, data from experiments, or web services.

6.3. Technical Dimensions

When it comes to the question of the standards that RD are expected to meet, political and strategic papers often refer to the standards and regulations given by the communities or societies of the respective disciplines. But there are only few disciplines that already have established norms that can e.g. be used for developing metadata schemes for repositories – a problem that is not limited to Germany (cf. Schirmbacher, 2017, pp. 398–399). An example for a discussion of a standard for the social sciences can be found in Jensen (2012).

The aforementioned CLARIN-D centres built an organisational, but also a technical network for offering an infrastructure for services for language-related resources.

A joint project of the German National Library and the library of Humboldt University in Berlin, “eDissplus”,⁷⁴ is working on conceptual and technical questions regarding archiving and publishing RD from doctoral research. The project aims to develop a prototypical integrated system for archiving and publishing the research data generated or used by doctoral students as part of their dissertation project. For this the existing workflow for the legal deposit of dissertations with the German National Library has to be enhanced and the URN service has to be expanded to allow RD to be identified, addressed, and linked to the corresponding thesis on a persistent basis.

r3data, the global registry of RD repositories, lists 321 entries for Germany – so there is a huge variety available, be they disciplinary or institutional repositories.⁷⁵ There are big ones developed jointly by several institutions, like “PsychData” and “Qualiservice” (for data from psychology and social sciences), RADAR (“Research Data Repository”, natural and information sciences), or the Humanities Data Centres (HDC), and there are more specialised ones like “plankton*net” run by the Alfred Wegener Institute for Polar and Marine Research or the animal sound archive called “Tierstimmenarchiv” by the Museum for Natural History in Berlin.

The German Research Foundation (DFG) also hosts a list of infrastructure services such as data repositories called “RIsources” (with RI as in “Research Infrastructure”).⁷⁶

Several German repositories have been awarded with the “Data Seal of Approval”.⁷⁷

There probably is a repository for nearly every need somewhere in Germany – but for scientists and for libraries it is not easy to keep up-to-date with the different services. And there are reasons why many institutions add their own repository to the list; see chapter 8 for the one to be developed for the WWU.

6.4. Conclusions

There are many papers to read, many institutions to know, many abbreviations to learn, many initiatives to follow – it is not easy to keep track of all that is going on in RD(M) in Germany.

As a relatively new topic, RDM emerged in a “scientific landscape” characterised by a federal structure and parallel developments in several places. While “many people working on a problem” can lead to many good ideas, it can also lead to the duplication of structures and developments and thus to higher costs and to a confusing overall picture. Hopefully the different projects will keep contact for “cross-fertilisation” and more central ideas like the NFDI will be introduced to develop sustainable solutions, to avoid more fraying, and to arrive at a more consolidated state.

Apart from organisational and technical issues, there are also still many open legal questions regarding RDM, e.g. in copyright or contract law: what kind of copyright should apply, which rights for the data are given to the operator of a repository, what kind of data protection is needed, or what are the regulations when two institutions from different states work together? And while many current German RDM activities are funded as projects, there have to be decisions on sustainable long-term financing: costs paid by federal government and states, by the universities or research institutions, by the researchers, or by share of cost? (Cf. e.g. Schirnbacher, 2017, pp. 394–396.)

The WWU will try to reuse as many ideas and solutions developed at other institutions as possible and to cooperate with several initiatives to create synergetic effects and to be able to offer good services in a short time and with a relatively small RDM team.

7. RDM at KTH & KTHB

7.1. The Current Setting

Considering that KTH aims to cover all subjects in the vast field of the engineering sciences, it is vital that a broad RDM support is implemented at KTH. There are already some support features in place. A special KTH Research Office unit has long been active in helping researchers during grant applications, which includes assisting them when they write Data Management Plans (DMPs).

As of today (August 2018), the President of KTH has not given any formal mandate to establish a support function. But we still believe that we have to develop some support services even without formal mandate. Our informal working group with people from KTHB, Archive, IT, Research Office, SNIC-PDC has now been active for less than a year. During that time, we have documented the current state and the future plans in a report for our Chief Librarian and we have started to attend selected networking or informational meetings in order to meet and engage with researchers and other parties at KTH that are vital for RDM.

We have built a support website with Q&A. We have also started to improve staff knowledge on GDPR and RDM. At the moment, frequent questions concern elements of the earlier parts of the research process, such as setting up a DMP and questions concerning GDPR. But there will inevitably be more questions concerning the long-term preservation of data in the future.

7.2. The Future Setting?

We are currently awaiting a formal mandate from KTH. After receiving that, we can continue our work, probably with the recruitment of special competencies necessary to expand and scale up our support services.

As of now, no RDM policy has been implemented at KTH. Hence, there is room for improvement! KTHB can work closely with researchers to develop this policy. It is probably important that this policy has some obligatory clauses, rather than only providing non-mandatory guidelines.

The question is also: what kind of repository solution(s) shall we recommend for small-scale data producers at KTH? With an unlimited amount of resources available, the solution is easy: let us at KTH develop a FAIR-compliant repository for KTH researchers and let us recruit metadata coordinators to help researchers when they are depositing their data. This would make the KTH RDM support unit's work much easier in that it would allow staff to point researchers to that repository.

However, although the transaction costs for making the RD publicly available may be marginal compared to the total cost for any research project, an official KTH data repository is more of an ideal solution than a feasible one. First, it seems highly improbable that "one solution fits all" would work in this case. We do have to accept disciplinary (and cultural) differences. Second, it is a labour-intensive and expensive solution, an observation from other data-curation activities. (KTH also has no tradition of publishing activities, we have no KTH university press, etc.). Hence, active data-curation activities by a KTH RDM support function are not feasible.

8. RDM at WWU & ULB MS

8.1. RDM Sneaking in: First Steps

After the papers mentioned in chapter 6, especially the 2014 HRK recommendations, German universities that had not yet done so also started thinking on how to tackle the vast field of RDM – i.e. vast in Germany where it was still emerging; in other countries like Great Britain or the Netherlands, it was already quite developed.

By that time, the WWU could look back on 15 years of close cooperation between the university library, central IT services, and the university's administration. This alliance, called "IKM" ("Information, Kommunikation und Medien" = "information, communication, media"), coordinates the strategic planning, development, and maintenance of digital infrastructures and services for students and faculty like computer-assisted teaching, e-learning, production of multimedia material, or file storage facilities.⁷⁸

The WWU rectorate commissioned the development of an RDM strategy for the university. This put the topic on the official agenda – an important step

for the establishment of such an extensive matter. A working group composed of IKM members and professors set about the task.

The IKM group then carried out a survey among faculty regarding their view on and their experiences with research data and RDM. The questionnaire asked for information about the kind of data the scientists were dealing with in their research, how they stored it, which subject-specific or funding-related guidelines were relevant for them, and which kind of information and services they expected from their university. The results corresponded with those of similar surveys at other universities: a wide range of types of data, often only saved on local computers and backed up irregularly and mostly not publicly available, due to legal concerns and/or a lack of time for the preparation of the publication. About 45% of the scientists answering the survey admitted to have only little or very little knowledge about RDM, wishing for technical and legal advice to be given by the university, while there were only few guidelines known (or relevant) for their research (cf. Meyer-Doeringhaus & Tröger, 2015).

These results served as starting points for the WWU journey into the world of RDM.

To anchor the topic on a university-wide level, an intensively discussed WWU Research Data Policy⁷⁹ passed the senate in summer 2017. With its publication and the establishment of a Research Data Service Point⁸⁰ structured WWU RDM activities officially started.⁸¹

Of course, RDM questions had been raised before 2017 – but only “here and there” and “now and then”, originating from single projects from e.g. the Institutes for New Testament Textual Research, German Studies, or Egyptology and Coptic Studies. From 2005 to 2010, the IKM group carried out a project about the management of scientific and administrative information,⁸² and the ULB joined forces with the Institute of Geoinformatics for research projects on linked spatio-temporal data⁸³ and Opening Reproducible Research.⁸⁴

8.2. Picking up Speed: RDM in the Wake of DH

While these smaller and bigger projects were very valuable for gaining experience with the handling of data in different contexts, the main focus of the

ULB's department for Science and Innovation in the last three years has shifted from RDM to Digital Humanities.

As the science disciplines like physics, chemistry, or informatics had already found solutions for their current RDM needs, the requests that the ULB got from WWU researchers centred on different aspects of DH, with RDM as one facet.⁸⁵ It became clear that most WWU DH projects were working independently from each other, and thus were running the risk of reinventing several wheels for each project. The ULB started acting as a coordinator, bringing together the researchers and establishing contact between the projects and their infrastructural needs.

An intensification of these first activities was triggered by the so-called WWU Cluster of Excellence⁸⁶ "Religion and Politics in Pre-Modern and Modern Cultures".⁸⁷ Established in 2007, this research group had to apply for follow-up funding in 2017 – and the funding regulations asked for detailed information about how DH and RDM would be handled in the next funding period.

Intense discussions between professors and faculty from DH projects and the Cluster of Excellence as well as representatives of the ULB and the central IT services led to the foundation of a Center for Digital Humanities (CDH) in July 2017. It brings together, coordinates, and accompanies DH projects at the WWU, and it will advance the teaching e.g. of the use of DH tools for students.

Meanwhile, the WWU – under a new rectorship since autumn 2016 – had adopted a "digitalisation strategy"⁸⁸ regarding every aspect of the university life, be it teaching, research, administration, or infrastructure. RDM and DH features as important factors in the strategic WWU development plan published in spring 2018.⁸⁹

8.3. The Current Setting

The framing of all DH, RDM, and other digital scholarship services is set by the WWU's eScience strategy:

- The so-called *eScience Center – e* as in "enhanced" – will be the competence and services centre for digital methods and resources for all WWU departments.

- A *Service Point Digital Humanities* as part of the eScience Center will take care of specific DH tasks like consulting projects and developing central services. It will start working in autumn 2018, staffed with four positions (a DH coordinator, a development coordinator, two software developers).⁹⁰ Until these new positions are filled, the ULB will continue its coordinating tasks, collecting needs and demands and preparing central services.
- The aforementioned *Service Point Research Data Management* will provide information and services for RDM in general.
- More Service Points e.g. for digitisation of research objects (text and audiovisual media, 3D objects) will follow.

Organizationally, the eScience Center and the Service Points are affiliated with the ULB department “Science and Innovation”.

Responsibilities and competencies for RDM are shared among the IKM partners: while the administration is e.g. responsible for the research information system (CRIS) and the management of WWU personnel data, the IT services deal with the technical aspects of file storage or information security. The library, on the other hand, runs the RDM Service Point as the central unit, coordinating services, workflows, and processes as well as contributing to topics like metadata handling, publishing, or RD information literacy. As there are several overlapping points between RDM and DH, the two Service Points will work closely together and develop their services in cooperation.

Alongside working on local services, the eScience Center also monitors the ongoing work of initiatives in Germany and abroad to stay up-to-date with current activities, to identify tools and services that can be integrated into the RDM at the WWU, and to get ideas for future developments, be it specifically for the WWU or as joint projects with other universities.

8.4. Next Steps

Regarding RDM the next main tasks for the eScience Center and the Service Points will be:

- giving advice on RD and RDM in general and on how to “bring the WWU RD policy to life”

- developing a repository for WWU research data
- interlinking this repository, the MIAMI document repository, the CRIS research information system, and ORCID, aiming at a better data exchange and a reduced administrative workload for the scientists
- developing a DMP tool on the basis of RDMO⁹¹
- developing “sciebo Research Data Services” (sciebo.RDS⁹²) as an RDM toolbox
- developing an “eScience Cloud” with OpenStack cloud computing as Infrastructure as a Service (IaaS)
- combining all these RDM and DH tools and services for use inside the eScience Cloud, thus building a kind of “scholarly makerspace” (cf. Kaden, 2018b) with emphasis on integrating and connecting existing services and enabling continuous workflows for research projects modelled on different use cases
- developing an aggregation of WWU research data that can be used for testing or “playing around with data” for pilot studies
- developing a business model for extensive data curation and data storage
- developing training and workshops for students and faculty to raise awareness of and competencies in RD and RDM
- coordinating WWU projects and consolidating RDM activities to build a strong network inside the university

While there is a lot to talk about for each single step, let us pick out three for some additional remarks.

The publication of the *WWU RDM policy* has been a big step towards establishing the topic at the university. But the principles laid down in the document are quite general and abstract, and the detailed specification is left to the faculties – who have not yet published corresponding papers. Thus the researchers need advice on how to “read” the policy and how to put the principles to practice. Therefore one of the main tasks for the library is consulting about the policy and, in many cases, going further back to the beginnings and discussing what kind of research data is expected in a planned project or has been produced in a completed project.

As e.g. surveys have shown that WWU scientists would prefer to put their data in a WWU repository rather than one “somewhere else” and as other

repositories do not always offer features WWU scientists are looking for, the development of a *RD repository* is one of the first tasks currently being worked on. The repository will have to be able to handle all kinds of data types and formats. The consultations have shown that sometimes scientists contacting the library cannot yet answer the DMP question “which kind of data types and which data formats will be produced in your project?”, as they are still thinking about how to tackle their research questions. One of the questions to decide is e.g. whether the data should be kept “interpretable” or executable, or whether it just has to be stored to keep it available on request. So the repository has to be able to “ingest everything” from PDF to XML or TIF files to “dark archives”, i.e. encrypted data packages that will be stored on behalf of the respective researchers but that will not be openly published or processed for long-term preservation. Several repository software systems have been tested in the last months, but so far the ideal solution has not been found. The WWU repository will combine a basic system with specific additions concerning use case workflows.

The *services* that the Service Points for RDM and for DH will develop will concentrate on generic services that can be reused by several projects all over the WWU. Highly specific services for single projects will have to be developed within the projects or bought in from other suppliers,⁹³ with the WWU Service Points being available for discussion, but not for programming. Using the (slightly simplified yet handy) four-field-strategy matrix by Peukert (2017) set up by the variables “data” and “usage” with the two properties “same” and “different”, the Service Points will develop services on the “standard level” (same data, same usage) and on the “modularize data workflow” level (different data, same usage) and “modularize usage workflow” level (same data, different usage), but not on the “individualize” level (different data, different usage).

8.5. What the ULB will not do (yet)

Roughly speaking, the first set of RDM tools and services offered for the WWU will concentrate on those services asked for by the scientists and/or needed with regard to funding requirements. Other services that would be welcome additions but are not urgently needed might be added later, if demand rises – and the number of RDM staff increases.

As mentioned above, services and tools that are very specific and will probably be used only in one or very few projects will not be developed by the WWU eScience Center. Instead they will have to be taken care of in the respective projects.

For the long-term preservation of stored data, the WWU will use regional and national infrastructures like e.g. the one provided by the North Rhine-Westphalian Library Service Centre hbz based on the Exlibris software “Rosetta”.⁹⁴

And for the time being, bibliometric analyses will not be a part of the ULBs research support services.

8.6. Some Lessons Learned

As we have seen in chapter 6, some German universities have already established big competence centres for RDM and some offer training for RDM or even Bachelor or Master courses dedicated to DH. While there have been several small WWU research projects using DH methods and/or needing RDM support, their number and with them the awareness of RDM and DH topics has grown in the last two years – the WWU is catching up: The ULB and the IT centre are now dealing with more and higher demands of professors, faculty, and projects.

Experiences so far are consistent with e.g. those reported for the library of the Technical University (ETH) in Zurich⁹⁵ (cf. Töwe, 2017): RDM is a topic that affects not only nearly all units of the Digital Services department dealing with metadata management, repository services, or publishing, copyright, and open access, but also the information department or the subject or liaison librarians. It intensifies the cooperation with the university’s central IT services, and it brings new people to the library who often are not trained librarians but “practitioners” coming from research and bringing a “refreshing lack of understanding” (cf. Töwe, 2017, p. 367) of the structures and working methods of a library. While they should become a part of and identify with “the library spirit”, they should also keep “the researcher’s perspective” as long as possible – a challenge for both sides.

Consultations with researchers often exceed “simple” questions like “Where do we store our files?” and can lead to discussions about “What kind of data

are you dealing with?” not only in terms of general definitions, but also in terms of the various workflows in the specific projects, about criteria for RDM in the respective disciplines, or even about the fundamental aims and goals of a given research project. Ideally these discussions take place “ab ovo”, i.e. before a project is started, e.g. while writing a grant application. Sometimes researchers contact the RDM Service Point “in vivo”, during an ongoing project, while some only start asking questions “post mortem”, when a project has come to its end.⁹⁶ Each stage demands for different aspects to be dealt with during the consultations, and each researcher and each project can be different: from perfectly prepared documents with only some specific questions left to answer to “loose leaf notes” with nothing more than a first rough draft of a project, from researchers interested in and open to the ideas of sharing data to sceptics who only want to fulfil the minimum grant requirements as quickly and as simply as possible – we have to be prepared for everything.

While there are limits to what the library can do, it can always “help the researchers help themselves”: it cannot e.g. provide detailed RDM implementation regulations for each discipline – but it can help researchers to find out whether there are existing standards or best practices used in the respective disciplines or what considerations and decisions a project has to take to fulfil RDM demands. And while the “cultural change” towards Open Science can only emerge from the scientific community itself, the library can at least try to encourage scientists to discuss these topics in their disciplines – or at least with the colleagues from across the hall who are working on comparable questions ...

It has become clear that the WWU eScience Center will have to continuously work on building awareness not only for RDM in general, but also for the competencies and services that are already available for the researchers at WWU. In such a big university, this information often takes long to spread, and while monitoring activities at other universities and in other countries are always interesting, we also have to keep track of new (or not yet known) projects at the WWU to get in contact with them and thus to bring together the scientists and their demands with our RDM specialists – but also bring together the scientists with other scientists working on similar topics or using similar tools.

On the technical side, it has become clear that while there are many software programs and tools made available by other projects, they often lack detailed documentation. This makes reusing them in other projects or expanding

them with additional features very difficult or impossible: It would take longer to try to understand the source code and get it running than to develop a new program. Furthermore the technical infrastructure and the workflows differ in every institution, making individual local adjustment necessary. This is why the WWU eScience team will e.g. develop parts of the RD repository on its own. The focus will be on engines for workflows that are tailored to meet the demands of certain use cases. Because of the use of the Business Process Model and Notation (BPMN) standard, the definition (and with it the documentation) of the workflows has to be done before the actual programming. This way the software that is developed locally will hopefully be reusable for other institutions – and this way of programming also mirrors the learnings from many a consultation with researchers: most of their questions and demands are about RDM workflows.

With regard to the time that is needed to build knowledge and competences in RDM and cooperation between the many institutes of a big university, it is good that the WWU eScience center and Service Points are run with permanent positions and not on a project basis. This will hopefully help to develop sustainable solutions for the WWU.

9. “Same same but different”?

What have we learned from our short run-through?

9.1. Engineering vs Humanities

One prejudice one might have when thinking about “the people with screwdrivers” vs “the people with dictionaries/Old French dramas/oil paintings from the Baroque” could be true: while engineers are fully aware of the fact that they are handling data, this is not necessarily true at least for “traditional” humanists. But the closer we come to digital humanities, the smaller this gap becomes.

An interesting difference has been observed at the WWU: while scientists having questions regarding RDM tend to contact the IT/computing centre instead of the library, humanists tend to contact the library instead of the

IT centre. The two institutions have to combine their “established trusts” to make sure all researchers get the information they need.

We have seen that both data from engineering and the humanities are different from data from the natural sciences in that the former are often human-made or based on artefacts. But while humanist data can come in many different “flavours” engineering data is mostly of the ordinal type, making it relatively easy to proceed, compute, or compare.

This entrains the biggest difference between the “two cultures”: the demands on the technical infrastructure for storing the data. If it were only for “keep it safe”, relatively simple storage solutions would be enough for both sides. But if it is about “keep it alive and running”, things are more complex on the humanist side: the data for just a single research project can be a wickerwork of data types or formats and layers of annotations, visualizations, and interpretations. Data centres have to be able to rebuild this as a “living system” to make the results comprehensible and reusable as close as possible to the situation given during the research project.

Both disciplines share the fact they may encounter sensitive data like geopositions or personal information about survey participants, but this is a universal property that can be found in every discipline, with health data as a prime example.

And at the latest when it comes to “getting money”, they all meet again: “While the data produced in scientific research differ from the digital humanities in many ways, both disciplines have similar themes when it comes to research data management as exhibited by the similarities in the data management plan requirements of the funding agencies NSF [National Science Foundation, USA; V.V.] and NEH [National Endowment for the Humanities, USA; V.V].” (Dressel, 2017, p. 6)

Other aspects regarding RDM are more dependent on the personality of a researcher and not that much on her or his discipline: for some it is a matter of course that research data should be published – and ideally published open access –, others are difficult to persuade and will only do so when e.g. funders make them to. Some know that their data might be useful for other researchers, others think that publishing them would not be worthwhile – or they do not want their data and ideas be “stolen” by others. Some think

about RDM questions right from the beginning of drafting a project, others just start thinking about it two days before they have to hand in their grant application or their final project report.

Kaden (2018a) has collected a long list of reasons why researchers do not publish research data; most reasons will probably be found evenly distributed in all disciplines.

Most researchers e.g. want to spend their time researching, but not organizing the research process. In this regard RDM is comparable to knowledge and information management: “the type of work you need, but nobody wants to see: Underwear work” (cf. Bohle Carbonell, 2018; Goble, 2018).

The researchers’ personalities are an important factor that institutions working in the field of RDM should keep in mind, e.g. when developing an RDM policy, as the study by Linek et al. has shown: “[i]f policy makers want to foster data sharing, it is not sufficient to concentrate only on global interventions. Rather they have also to consider the individual needs and apprehensions in relation to the researchers’ personality” (Johnson & Steeves, 2018; Linek, Fecher, Friesike & Hebing, 2017, p. 20; Zenk-Möltgen, Akdeniz, Katsanidou, Naßhoven & Balaba, 2018).

So libraries as – ideally – institutions “in the middle of the campus” have to know about how data types and ways of research differ in the different disciplines, but they also have to see the similarities of doubts and demands of researchers. With this they help avoid the “silozation of the academy”, as Moritz et al. (2017, pp. 4–5) put it: “In order to promote greater flexibility, deeper collaboration, and increased innovation across campus, digital scholarship questions library service models that restrict communication between disciplinary scholars and the library to a single departmental liaison.”

9.2. Sweden vs Germany

Both countries started intensely working on RDM at nearly the same time, following discussions about good scientific work and open access.

In a small country, administrative things can be easier than in a big country. With the Vetenskapsrådet as a major player acting as an advisor for the

government, a policy maker, and a funding agency, and with the Swedish National Data Service as a big service and communication platform provider, some important elements are much more centralized than they are in Germany – thus it is easier to keep track of services and developments.

On the other hand a small country does not have the financial and staffing capacities of a bigger one. This may be the opportunity for close cooperation between Swedish universities and other scientific institutions within Sweden, but also with institutions in other countries, while in Germany the risk of duplicate structures and developments is immanent in RDM.

As an example of a similarity, we have noticed that many RDM policies have little or no connection to the actual needs of the affiliated researchers in both countries. They will have to be complemented with more specific recommendations, which should take into account the respective disciplines. Here we may have a kind of divide between the disciplinary cultures, but we can easily transcend country borders: Swedish engineers could e.g. find inspiration at German universities, while German humanists could have a look at Swedish recommendations – or the other way round. We could also add other countries to the mix.

9.3. KTHB & KTHB vs WWU & ULB MS

Being “tired of using the word infrastructure all the time, with its attendant technophilic tendencies” and inspired by “conversations about the ecosystem of academic publication”, Vandegrift (2018) introduces the concept of “ecologies” defined as ‘interactions between organisms and their environment’ for a holistic view of all the activities regarding digital scholarship. In Vandegrift’s vision, there are no “digital centers” in libraries, but instead there will be “an ecosystem wherein new ideas are generated, incubated, and enveloped into the university, the locality, and the global community” (Moritz et al., 2017; Vandegrift, 2018). This will be achieved by four “near-term revolutions that will lead to long-term evolution”, triggered by digital scholarship: “Librarianship is/will be 1) omnipresent in the research lifecycle, 2) data-focused, 3) infrastructure aware, and 4) essential to the community, within the university and beyond” (Vandegrift, 2018).

The (r)evolution at KTHB and ULB may not yet have come that far, but their services, infrastructures, and expertise are already leading to more interaction between the libraries and other parts the universities. As both libraries are just starting to develop and establish services, we cannot yet present best-practice examples, but we will continue to look out for practices at other libraries and universities.

Apart from the differences in size and number of staff, KTHB can concentrate on subjects taught at KTH, while the ULB MS – as the library of a big university where nearly every subject can come along – has to be “prepared for everything”.

Nevertheless WWU and ULB can learn from KTH(B), e.g. in the area of engineering data, which might also be of interest for researchers in physics or other technically oriented subjects at the WWU. Of course there are also overlaps in biotechnology, chemistry, or computer science. Regarding engineering and also architecture, the other academic library in Münster, the one from the University of Applied Sciences Münster,⁹⁷ could be another interesting partner for joint considerations on RDM.

A problem all libraries have to tackle is the fact that keeping track of all the developments in RDM is already a sizeable task. For this reason and for deciding which services to implement for the respective library, it is important to train continuously staff and to establish close cooperation between library and faculty – and between other libraries, as e.g. Faniel and Connaway (2018, p. 115) note: “Finding ways to share support for RDM efforts, particularly technical and human resources, reduces burdens on individual libraries and their institutions. By growing the infrastructure together, no library or institution should create and sustain RDM programs alone.”

And the discussions between the authors of this paper, originating in an coincidental conference meeting, have shown that “cross discipline/country/library type” exchanges can also be fruitful when discussing RDM and other library topics: you learn by explaining the context you are working in and you learn by trying to understand other contexts. Library conferences can help avoid “siloization” in this regard as well.

10. Conclusions

10.1. From Local Services ...

There is different background knowledge about “data” in general and RDM in particular in the different disciplines. There are differences regarding the technical aspects of preserving data. However, many aspects of RDM are the same: scientists have similar questions, similar fears, or a similar lack of time, no matter where they are from – from the sciences or the humanities, from a small country or a big one, from a small university or a big one.

Many papers and presentations in the last years have asked whether RDM could or should be a new task for libraries. This could be classified as a rhetorical question today,⁹⁸ but, as Töwe (2017, p. 370) notes, the external conditions are different to other topics: regarding RDM, libraries do not decide on the agenda or on the tempo. Instead the demands of funders, editorial boards, or university administrations steer the needs and respective measures that have to be taken. To keep at least some influence over how to implement the measures, libraries have to be well connected to the important players at their universities. This leads us back to the mission statement that Vandegrift (2018) postulated. The measures that KTHB and ULB MS have undertaken so far aim in that direction.

For framing future RDM developments at KTHB and ULB – and at other libraries –, the model developed by Pinfield, Cox and Smith (2014) could be useful. It was developed on the basis of interviews with library professionals in the UK and “is intended to address the ‘who?’, ‘what?’, ‘why?’, and ‘how?’ of RDM, particularly in relation to the library’s involvement” (Pinfield et al., 2014, p. 22).

Six stakeholder groups (“who?” – library, information technology (IT) services, academic departments, senior university managers, research support services, other support services) represent the main actors involved in institutional RDM that is “fuelled” by seven drivers (“why?” – storage, security, preservation, compliance, quality, sharing, jurisdiction). The RDM programme itself is composed of six main components (“what?” – strategies, policies, guidelines, processes, technologies, services) that are shaped in different ways by the drivers, the stakeholders, and 12 influencing factors

(“how”? – acceptance, cultures, demand, incentives, roles, governance, politics, resources, projects, skills, communications, context).

Defining these four elements and analyzing their roles for and relations to RDM activities at KTHB and ULB will hopefully help build a lively digital scholarship ecology at both KTH and WWU.

10.2. ... to the Bigger Picture

Finally, taking a perspective on research data that is larger than developing local RDM support functions, we now come full circle and return to C. P. Snow’s 1959 lecture.

While Snow’s positions on the divide between “the scientists” and “the literary intellectuals” have been and still are disputable, there is one aspect that has not lost much of its immediacy.

In a comment on the *The two cultures* published four years later, Snow (1963, p. 79) mentioned: “Before I wrote this lecture I thought of calling it ‘The Rich and the Poor’, and I rather wish that I hadn’t changed my mind.” It was only the last part of his lecture that got this title, and this chapter was the one Snow “intended to be the centre of the whole argument”.

In his view the scientific revolution – i.e. “the application of real science to industry” (Snow, 1959) – is the only way to increase the well-being of society, e.g. with regard to public health. For this revolution to be successful a change in the educational system is needed (from an early specialisation in either the sciences or the non-sciences to a broader education) to minimize the division and to enable communication between the two sides. In this, the richer countries have to help the poorer countries to enable them to take part in the scientific revolution as well.

In the nearly 60 years that have passed since Snow’s lecture, much has happened in education, in technology, or in the communication between the different scientific disciplines, politics, and society. But the idea that supporting scientific progress in all countries is the only reasonable way to a better life on and for this planet is not the least outdated.

Apart from general education, industrial development, or public health, it can also be found in higher education and science: in the current framework for scientific publishing, most publications and data are behind a paywall or locked-in in closed repositories or on local computers or USB sticks. This reduces the possibility for less-developed countries to access the current state of research. In this instance, the approach to *open* research data management is vital.

The outbreaks of Ebola in Africa or Zika in South America may hold as an example. In the first case, no structured collection of data from the spread of the disease or patient history was made; instead the data is located on individual hospital teams' computers (cf. e.g. Hodson, 2018). In the second case, data from the Zika cases was collected both in more detail and in a more structured manner, helping to handle this outbreak more efficiently.

These circumstances have e.g. also been noted by the President of the Karolinska Institutet (KI), one of the biggest medical universities, on the occasion of the 2018 cancellation of Elsevier journal packages by the Swedish Bibsam consortium: "This was one of several issues that we discussed in our commission on global governance for health. We concluded – in no uncertain terms – that restrictions on access to knowledge serve to aggravate extant knowledge disparities and health inequities. Equal access to information – irrespective of geography and economy – is central to improvement of health, the very mission of KI. In my mind, it is in society's interest – and also in our own interest as scientists – that what we publish actually reaches all those who need the knowledge and who stand to benefit from it." (Ottersen, 2018)

By encouraging "our" faculty and students to open up their research and with developing RDM services that facilitate open science, we can add some small tesserae to bridge the gap between the rich and the poorer countries, at least with regards to the barriers of access to scientific data. This is not "rainbows, unicorns, & puppies"⁹⁹ but a fundamental concern.

Acknowledgements

Göran Hamrin would like to thank his KTHB colleague Rosa Lönneborg for interesting discussions and for her inspiring work with RDM at KTH.

Viola Voß would like to thank her ULB colleagues Jörg Lorenz and Marc Schutzeichel for insights in and discussions about the ULB's eScience strategy and RDM services, Peter te Boekhorst for bringing C. P. Snow into play, and Else Gellinek for brushing up our English.

References

- Allianz der deutschen Wissenschaftsorganisationen. (2010). *Grundsätze zum Umgang mit Forschungsdaten*. <https://doi.org/10.2312/ALLIANZOA.019>.
- Allianz der deutschen Wissenschaftsorganisationen. (2018a). *Forschungsdatenmanagement. Eine Handreichung*. <http://doi.org/10.2312/allianzoa.029>.
- Allianz der deutschen Wissenschaftsorganisationen. (2018b). *Research Data Vision 2025 – ein Schritt näher*. <http://doi.org/10.2312/allianzoa.024>.
- Andorfer, P. (2015). *Forschungsdaten in den (digitalen) Geisteswissenschaften: Versuch einer Konkretisierung*. DARIAH-DE working papers, nr. 14. Göttingen: Georg-August-Universität. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.
- Bohle Carbonell, K. (2018). *Enigmas, networks, and people at work — The underwear of data science* [Blog post]. <https://katerinabc.com/make-open-science-successful/>.
- Bollen, J. (2016). Data models for theatre research: people, places, and performance. *Theatre Journal*, 68(4), 615–632. <https://doi.org/10.1353/tj.2016.0109>.
- Burghardt, M. (2018a). Digital Humanities in der Musikwissenschaft – Computergestützte Erschließungsstrategien und Analyseansätze für handschriftliche Liedblätter. *Bibliothek – Forschung und Praxis*, 42(2), 324–332. <https://doi.org/10.1515/bfp-2018-0032> [preprint available at <http://dx.doi.org/10.18452/18952>].
- Burghardt, M. (2018b). *Music information retrieval meets digital humanities: Computergestützte Erschließung und Analyse von handschriftlichen Volksliedblättern* [Slideshow]. https://ch.uni-leipzig.de/images/2018-Weimar-MIR/2018_Weimar-MIR.pdf.
- Cremer, F., Klaffki, L., & Steyer, T. (2018). Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften. *o-bib*, 5(2), 142–162. <https://doi.org/10.5282/o-bib/2018H2S142-162>.
- Curdt, C., Krämer, F., Hess, V., López, A., Magrean, B., Rudolph, D., & Vompras, J. (2016). *Einführung in Forschungsdatenmanagement*. <https://doi.org/10.5281/zenodo.165126>.
- Curdt, C., Hess, V., López, A., Magrean, B., Rudolph, D., & Vompras, J. (2017). Herausforderung Forschungsdatenmanagement – Unterstützung der Hochschulen

- durch eine einrichtungsübergreifende Kooperation in NRW. In: J. Kratzke & V. Heuveline (Eds.), *E-Science-Tage 2017: Forschungsdaten managen* (pp. 95–103). Heidelberg: heiBOOKS. <http://dx.doi.org/10.11588/heibooks.285.377>.
- Curdt, C., Grasse, M., Hess, V., Kasties, N., López, A., Magrean, B., ..., & Winter, N. (2018). *Zur Rolle der Hochschulen: Positionspapier der Landesinitiative NFDI und Expertengruppe FDM der Digitalen Hochschule NRW zum Aufbau einer Nationalen Forschungsdateninfrastruktur*. <https://doi.org/10.5281/zenodo.1217527>.
- DBV Sektion 4. (2018). *Wissenschaftliche Bibliotheken 2025*. https://www.bibliotheksverband.de/fileadmin/user_upload/Sektionen/sektion4/Publikationen/2018_02_27_WB2025_Endfassung_endg.pdf.
- DFG (Deutsche Forschungsgemeinschaft). (2009). *Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten*. http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf.
- DFG (Deutsche Forschungsgemeinschaft). (2013). *Sicherung guter wissenschaftlicher Praxis, Safeguarding good scientific practice*. Amended edition. http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf.
- DFG (Deutsche Forschungsgemeinschaft). (2018). *Stärkung des Systems wissenschaftlicher Bibliotheken in Deutschland. Ein Impulspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft*. http://www.dfg.de/download/pdf/foerderung/programme/lis/180522_awbi_impulspapier.pdf.
- DHd AG Datenzentren. (2018). *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten*. <http://doi.org/10.5281/zenodo.1134760>.
- Dierkes, J., & Curdt, C. (2018). Von der Idee zum Konzept – Forschungsdatenmanagement an der Universität zu Köln. *o-bib*, 5(2), 28–46. <https://doi.org/10.5282/o-bib/2018H2S28-46>.
- DINI (Deutsche Initiative für Netzwerkinformation). (2018). *Thesen zur Informations- und Kommunikationsinfrastruktur der Zukunft*. <http://dx.doi.org/10.18452/19126>.
- Dressel, W.F. (2017). Research data management instruction for digital humanities. *Journal of eScience Librarianship*, 6(2), e1115, 1–7. <https://doi.org/10.7191/jeslib.2017.1115>.
- Faniel, I.M., & Connaway, L.S. (2018). Librarians' perspectives on the factors influencing research data management programs. *College & Research Libraries*, 79(1), 100–119. <https://doi.org/10.5860/crl.79.1.100>.
- Formas. (2016). *Formas policy för öppen tillgång till forskningsdata*. (Policy decision made at The Swedish Research Council for Environment, Agricultural Sciences and

Spatial Planning.) 2016-06-30. Diarienummer 1.4-2016-01894. http://www.formas.se/PageFiles/15559/Beslut_policy_oppna_data.pdf.

Gaillard, M. (2017). *CERN Data Centre passes the 200-petabyte milestone*. <https://home.cern/about/updates/2017/07/cern-data-centre-passes-200-petabyte-milestone>. Retrieved September 20, 2018, from <https://cds.cern.ch/record/2276551>.

Goble, C. (2018). *Building the FAIR Research Commons: A Data Driven Society of Scientists*. Talk given at the Symposium "The Future of a Data-Driven Society" at Maastricht University on the 25th of January 2018. Video available at <https://www.maastrichtuniversity.nl/events/review-symposium-future-data-driven-society> (talk starts at 09:24), slides available at <https://www.slideshare.net/carolegoble/building-the-fair-research-commons-a-data-driven-society-of-scientists>.

Gradl, T., Henrich, A., & Plutte, C. (2015). Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen. *Zeitschrift für digitale Geisteswissenschaften, Sonderband 1* (n.p.): C. Baum & T. Stäcker (Eds.), Grenzen und Möglichkeiten der Digital Humanities. http://doi.org/10.17175/sb001_020.

Hahn, H., Kalman, T., Kolbmann, W., Kollatz, T., Neuschäfer, M., Pielström, S., ..., & Tonne, D. (2015). *DARIAH-DE -Handbuch Digital Humanities: Anwendungen, Forschungsdaten und Projekte*. Berlin: Open Knowledge Foundation. <http://handbuch.io/w/DH-Handbuch>.

Hansson, S.O. (2007). What is technological science? *Studies in History and Philosophy of Science Part A*, 38, 523–527. <https://doi.org/10.1016/j.shpsa.2007.06.003>.

Hodson, S. (2018). *African open science platform*. Talk given at the 11th Research Data Alliance meeting Berlin, 21.-23.3.2018.

HRK (Hochschulrektorenkonferenz). (2014). *Management von Forschungsdaten – eine zentrale strategische Herausforderung für Hochschulleitungen*. http://www.hrk.de/uploads/tx_szconvention/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf.

HRK (Hochschulrektorenkonferenz). (2016). *Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können: Orientierungspfade, Handlungsoptionen, Szenarien*. https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-10-Publikationsdatenbank/Beitr-2016-01_Forschungsdatenmanagement.pdf [English version: *How university management can guide the development of research data management. Orientation paths, options for action and scenarios*. https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-10-Publikationsdatenbank/Beitr-2016-01_Research_Data_Management.pdf].

IANUS. (2017). *IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften*. <http://dx.doi.org/10.13149/000.111000-a>.

Ingram, C. (2017). *Audiovisual research data: A guide to managing digital media for anyone using it as an integral part of carrying out research*. <https://www.jisc.ac.uk/guides/>

audiovisual-research-data. [German version: Ingram, C. (2018). Audiovisuelle Forschungsdaten: Ein Leitfaden zur Verwaltung multimedialer Forschungsdaten für alle, die diese als integralen Bestandteil der Forschung nutzen. *B.I.T.online*, 21(1), 36–41. <http://www.b-i-t-online.de/heft/2018-01-nachrichtenbeitrag-ingram.pdf>].

Jensen, U. (2012). *Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten*. TECHNICAL Reports. 2012 | 07. Köln: GESIS – Leibniz-Institut für Sozialwissenschaften. https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf.

Johnson, K., & Steeves, V. (2018). *Research data management among life sciences faculty: Implications for library service* [Preprint]. LIS Scholarship Archive. <https://dx.doi.org/10.17605/OSF.IO/Q36UV>.

Kaden, B. (2018a). Warum Forschungsdaten nicht publiziert werden. *LIBREAS. Library Ideas*, 33. <http://libreas.eu/ausgabe33/kaden-daten/>.

Kaden, B. (2018b). Was sind Scholarly Makerspaces? – eine erste Überlegung [Blog post]. *Future e-Research Support in the Humanities*, 14.5.2018. <https://blogs.hu-berlin.de/furesh/2018/05/14/scholarly-makerspaces/>.

Kaden, B., & Kleineberg, M. (2018). *Anforderungen und Trends werkzeuggestützter Forschung in den Geisteswissenschaften* [Blog post]. <https://blogs.hu-berlin.de/furesh/2018/05/15/anforderungen-und-trends-werkzeuggestuetzter-forschung-in-den-geisteswissenschaften/>.

Katerbow, M., & Feulner, G. (2018). *Recommendations on the development, Use and Provision of Research Software*. <https://doi.org/10.5281/zenodo.1172988>.

Kronenwett, S. (2017). *Dienstleistungsangebote für digital arbeitende GeisteswissenschaftlerInnen. Vorstellung des Cologne Center for eHumanities (CCeH) & Data Center for the Humanities (DCH)* [slideshow]. http://dch.phil-fak.uni-koeln.de/sites/dch/Materialien_Aktivitaeten/2017/Vortrag_Kronenwett-ZBIW-20170328.pdf.

Linek, S.B., Fecher, B., Friesike, S., & Hebing, M. (2017). Data sharing as social dilemma: Influence of the researcher's personality. *PLoS One*, 12(8), e0183216, 1–24. <https://doi.org/10.1371/journal.pone.0183216>.

López, A. (2015). *Forschungsdatenmanagement an deutschen Universitäten—ein neues Arbeitsfeld für Bibliotheken?* [slideshow] <https://nbn-resolving.org/urn:nbn:de:0290-opus4-17614>.

Ludwig, J., & Enke, H. (eds.) (2013). *Leitfaden zum Forschungsdaten-Management, Handreichungen aus dem WissGrid-Projekt*. Göttingen: Universitätsverlag. <http://resolver.sub.uni-goettingen.de/purl?isbn-978-3-86488-032-2>.

Meyer, A., Neumann, J., & Soßna, V. (2017). Service durch Kompetenzbündelung – Das institutionelle Konzept zum Forschungsdatenmanagement der Leibniz

- Universität Hannover. In: J. Kratzke & V. Heuveline (Eds.), *E-Science-Tage 2017: Forschungsdaten managen* (pp. 13–21). Heidelberg: heiBOOKS. <http://dx.doi.org/10.11588/heibooks.285.377>.
- Meyer-Doeringhaus, U., & Tröger, B. (2015). Forschungsdatenmanagement als Herausforderung für Hochschulen und Hochschulbibliotheken. *o-bib*, 2(4), 65–72. <http://dx.doi.org/10.5282/o-bib/2015H4S65-72>.
- Moritz, C., Smart, R., Retteen, A., Hunter, M., Stanley, S., Soper, D., & Vandegrift, M. (2017). *De-centering and recentering digital scholarship: A manifesto* [Preprint]. LIS Scholarship Archive. <https://dx.doi.org/10.17605/OSF.IO/T7HFU>.
- nestor. (2018). *nestor als Kern einer Langzeitarchivierungs- und Langzeitverfügbarkeits-Infrastruktur für digitale Daten und Dokumente in Deutschland*. [http://www.langzeitarchivierung.de/Subsites/nestor/SharedDocs/Downloads/berichte/ nestorPositionspapier2018.pdf](http://www.langzeitarchivierung.de/Subsites/nestor/SharedDocs/Downloads/berichte/nestorPositionspapier2018.pdf).
- Neuroth, H. (2018). *Forschungsdaten(management) in den Digital Humanities – Eine Annäherung*. Talk given at the lecture series “Digital Humanities: Die digitale Transformation der Geisteswissenschaften” at Humboldt University Berlin, 30.1.2018 [Video]. <http://www.bbaw.de/mediathek/archiv-2017/24-10-2017-digital-humanities/https://rs.cms.hu-berlin.de/rv-dh2017/pages/view.php?ref=29&k=b234f4cf40>.
- Ottersen, O.P. (2018). The circuitous road towards open access: Swedish universities to pull the plug on Elsevier [Blog post]. *Karolinska Institutet President's blog*. <http://blog.ki.se/rektor/the-circuitous-road-towards-open-access-swedish-universities-to-pull-the-plug-on-elsevier/>.
- Owens, T. (2011). Defining data for humanists: Text, artifact, information or evidence? *Journal of Digital Humanities*, 1(1), 6–8. http://journalofdigitalhumanities.org/files/jdh_1_1.pdf/http://journalofdigitalhumanities.org/1-1/defining-data-for-humanists-by-trevor-owens/.
- Peukert, H. (2017). Curating humanities research data: managing workflows for adjusting a repository framework. *International Journal of Digital Curation*, 12(2), 234–245. <https://doi.org/10.2218/ijdc.v12i2.571>.
- Pinfield, S., Cox, A.M., & Smith, J. (2014). Research data management and libraries: relationships, activities, drivers and influences. *PLoS One*, 9(12), e114734, 1–28. <https://doi.org/10.1371/journal.pone.0114734>.
- Pursula, A., Hedlund, J., Jareborg, N., Kaunisvaara, J., Azab, A., & Syed, A. (2018). *Tryggve – Nordic collaboration for sensitive data*. Poster presented at ELIXIR All Hands meeting 4.-7.6.2018. <http://dx.doi.org/10.7490/f1000research.1115559.1>.
- RfII (Rat für Informationsinfrastrukturen). (2016). *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. <http://www.rfii.de/download/rfii-empfehlungen-2016/> [English

version: *Enhancing research data management: performance through diversity. Recommendations regarding structures, processes, and financing for research data management in Germany.* <http://www.rfii.de/?p=2075>].

RfII (Rat für Informationsinfrastrukturen). (2017). *Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI).* <http://www.rfii.de/?p=2269>.

RfII (Rat für Informationsinfrastrukturen). (2018). *Zusammenarbeit als Chance. Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland.* <http://www.rfii.de/?p=2529>.

Schirmbacher, P. (2017). Dimensionen des Forschungsdatenmanagements im digitalen Zeitalter. In P. Hauke, A. Kaufmann, & V. Petras (Eds.), *Bibliothek – Forschung für die Praxis. Festschrift für Konrad Umlauf zum 65. Geburtstag* (pp. 389–405). Berlin: De Gruyter. <https://doi.org/10.1515/9783110522334-034>.

Schöch, C. (2013). Big? Smart? Clean? Messy? Data in the humanities. *Journal of Digital Humanities*, 2(3), 2–13. http://journalofdigitalhumanities.org/files/jdh_2_3.pdf / <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.

Schönbrodt, F., Gollwitzer, M., & Abele-Brehm, A. (2016). Der Umgang mit Forschungsdaten im Fach Psychologie: Konkretisierung der DFG-Leitlinien im Auftrag des DGPs Vorstands. *Psychologische Rundschau*, 68, 20–35. <https://doi.org/10.1026/0033-3042/a000341> [English version available via https://www.dgps.de/fileadmin/documents/Empfehlungen/Data_Management_eng_9.11.16.pdf].

Sesartic, A., & Dieudé, A. (2017). Research data management training and support services at both ETH Zurich and EPF Lausanne. Lessons learned, best practices and the way forward. In J. Kratzke & V. Heuveline (Eds.), *E-Science-Tage 2017: Forschungsdaten managen* (pp. 5–12). Heidelberg: heiBOOKS. <https://doi.org/10.11588/heibooks.285.377>.

Snow, C.P. (1959). The two cultures. In C.P. Snow & S. Collini (Eds.), *The two Cultures* (pp. 1–51). Cambridge: Cambridge University Press (2012).

Snow, C.P. (1963). The two cultures: a second look. In C.P. Snow & S. Collini (Eds.), *The Two Cultures* (pp. 53–107). Cambridge: Cambridge University Press (2012).

Stäcker, T., & Steenweg, H. (2018). Forschungsdaten – Aufgabe und Herausforderung für Bibliotheken. Editorial. *o-bib*, 5(2), IV–V. <https://doi.org/10.5282/o-bib/2018H2SIV-V>.

Thalinger, C. (2017). *Twitter's quest for a wholly Graal runtime* [Video]. Talk given at the conference JavaOne 2017. <https://youtu.be/G-vlQaPMAxg>.

Thoegersen, J.L. (2018). “Yeah, I guess that’s data”: Data practices and conceptions among humanities faculty. *Portal: Libraries and the Academy*, 18(3), 491–504. <https://doi.org/10.1353/pla.2018.0030>.

- Töwe, M. (2017). Wie Forschungsdaten die Bibliothek verändern. Erfahrungen aus der ETH-Bibliothek. *B.I.T.online*, 20(5), 361–370. <http://www.b-i-t-online.de/heft/2017-05/fachbeitrag-toewe.pdf>.
- Vandegrift, M. (2018). Designing digital scholarship ecologies [Preprint & Slideshow]. *LIS Scholarship Archive*. <http://doi.org/10.17605/OSF.IO/93ZVB>.
- Vetenskapsrådet. (2015). *Förslag till nationella riktlinjer för öppen tillgång till vetenskaplig information*. Report number 1.1.2-2015-8. <https://www.vr.se/analys-och-uppdrag/vi-analyserar-och-utvarderar/alla-publikationer/publikationer/2015-01-07-forslag-till-nationella-riktlinjer-for-oppen-tillgang-till-vetenskaplig-information.html>.
- Weisbrod, D., Kaden, B., & Kleineberg, M. (2017). eDissPlus – Optionen für die Langzeitarchivierung dissertationsbezogener Forschungsdaten aus Sicht von Bibliotheken und Forschenden. In J. Kratzke & V. Heuveline (Eds.), *E-Science-Tage 2017: Forschungsdaten managen* (pp. 189–198). Heidelberg: heiBOOKS. <http://dx.doi.org/10.11588/heibooks.285.37>.
- Wissenschaftsrat. (2012). *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*. <https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>.
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behavior of researchers in sociology and political science. *Journal of Documentation*, 74(5), 1053–1073. <https://doi.org/10.1108/JD-09-2017-0126>.

Notes

¹ We restrict ourselves to studying only researchers who create data on a small scale, since the large-scale data producers handle the RDM on their own. We here define small-scale – rather arbitrarily – to be up to 200 terabytes of data. This is a volume which few research projects exceed. As an example for comparison, there are over 200 petabytes of data stored at the CERN Data Centre (cf. e.g. Gaillard, 2017), so three orders of magnitude greater.

² Cf. <https://www.kth.se>.

³ Cf. <https://www.kth.se/kthb>.

⁴ Cf. <http://kth.diva-portal.org>.

⁵ Cf. <https://intra.kth.se/bibliometri/public/start>.

⁶ Cf. <https://www.kth.se/en/kthb/publicering/vagledning/kth-s-policy-for-publicering-1.623564>.

⁷ Cf. <https://www.kth.se/en/kthb/publicering/open-access/oa-erbjudanden-for-k/oa-erbjudanden-for-kth-forskare>.

⁸ Cf. <https://miami.uni-muenster.de>, <https://www.ulb.uni-muenster.de/wissenschaftliche-schriften/>, <https://www.uni-muenster.de/Ejournals/>, and <https://www.ulb.uni-muenster.de/publikationsfonds>.

⁹ The distinction is not entire crystal clear, naturally.

¹⁰ Cf. e.g. the talk by Thalinger (2017). This performance-optimisation can seem like a minor detail for most of us, but for Twitter it means significant savings in electricity and hence money. Hypothetical example given was that renting 10,000 cloud servers would result in saving 353,280 dollar/year.

¹¹ Cf. <https://www.proteinatlas.org>.

¹² Cf. <http://www.icos-sweden.se>.

¹³ Or the related database SwePub, cf. <http://www.swepub.se>.

¹⁴ Cf. <https://www.snic.se>.

¹⁵ Cf. <https://www.snic.se/resources/compute-resources/beskow/>.

¹⁶ Cf. <https://supr.snic.se>. Bigger allocations are possible, but they require a separate process, usually getting in touch with the SNIC Storage Coordinator. (Dejan Vitlacic, personal email communication, May 2018).

¹⁷ Cf. e.g. <https://de.wikipedia.org/wiki/Geisteswissenschaft>.

¹⁸ For some illustrated examples see e.g. Neuroth (2018). More interesting projects can be found e.g. in journals like “Frontiers in Digital Humanities” (<https://www.frontiersin.org/journals/digital-humanities>), the “Journal of Digital Humanities” (<http://journalofdigitalhumanities.org>), the “Journal of Open Humanities Data” (<https://openhumanitiesdata.metajnl.com>), in the blog of the DHd group (<https://dhd-blog.org>), in Hahn et al. (2015) (chapter “Digital Humanities in der Praxis”) or on the projects site of the WWU Center for Digital Humanities (<https://www.uni-muenster.de/CDH/projekte/>).

¹⁹ Cf. <https://ch.uni-leipzig.de/research/>. For an overview see e.g. Burghardt (2018a,b).

²⁰ The platform “Allegro” can be tested under the address <http://138.68.106.29/>.

²¹ “Wadi Abu Dom Itinerary (W.A.D.I.), cf. <http://en.wadi-abu-dom.de>. See <https://www.uni-muenster.de/news/view.php?cmdid=9565> for a short overview of the project.

²² WebGIS, cf. <http://wadi-abu-dom.de/web-gis-2/>.

²³ Project “Die Performanz der Wappen. Zur Entwicklung von Funktion und Bedeutung heraldischer Kommunikation in der mittelalterlichen Kultur (12.–15. Jahrhundert)” (“Coats of arms in practice. The history of heraldic communication

in late medieval culture and society (12th–15th centuries)"); cf. <https://www.uni-muenster.de/Geschichte/histsem/JPGHuSMA/forschen/PerformanzderWappen.html>.

²⁴ Cf. for this chapter e.g. Owens (2011), Andorfer (2015), Cremer, Klaffki, and Steyer (2018), and especially DHd AG Datenzentren (2018), p. 5–11.

²⁵ Cf. <http://www.linguisticsdatacitation.org>.

²⁶ Cf. e.g. the discussion about data centres at the DHd 2018 conference, <https://dhd-blog.org/?p=9740>.

²⁷ Cf. e.g. IANUS (2017) for different file formats used in Ancient Studies, Ingram (2017) for audio files, or Bollen (2016) for data in theatre studies.

²⁸ Cf. <https://adho.org>.

²⁹ Cf. <https://dig-hum.de>.

³⁰ Cf. e.g. Kaden and Kleineberg (2018).

³¹ Cf. e.g. Hahn et al. (2015) (chapter "Forschungsinfrastrukturen nutzen").

³² For more on data centres or data federations for the humanities cf. e.g. DHd AG Datenzentren (2018) or Gradl, Henrich, and Plutte (2015).

³³ Cf. <https://www.brill.com/rdj>.

³⁴ Cf. <https://openhumanitiesdata.metajnl.com>.

³⁵ Cf. <https://openarchaeologydata.metajnl.com>.

³⁶ Cf. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2018218-med-kompletterande-bestamnelser_sfs-2018-218.

³⁷ Cf. <https://snd.gu.se/en>.

³⁸ Cf. <https://internt.slu.se/stod-service/admin-stod/it/tjanster/systemutveckling/verktyg-och-system-under-utveckling/tilda/>.

³⁹ Cf. <https://www.su.se/english/library/publish/research-data/policies-and-guidelines-1.359054>.

⁴⁰ This was e.g. an approach advocated by several participants at the Research Data Alliance plenary meeting in Berlin, March 2018.

⁴¹ Cf. <https://www.csc.fi/-/tryggve2-increases-nordic-efforts-within-e-infrastructure-for-sensitive-data>.

⁴² Cf. http://www.dfg.de/en/dfg_profile/alliance/.

⁴³ Cf. <https://www.wissenschaftsrat.de>.

⁴⁴ Cf. <http://www.rfii.de>.

⁴⁵ Cf. <https://www.hrk.de>.

⁴⁶ Cf. <http://fdm-nrw.de>. For the activities leading to the Landesinitiative cf. e.g. Curdt et al., 2017.

⁴⁷ Cf. <https://www.dini.de>.

⁴⁸ Cf. <http://www.langzeitarchivierung.de>.

⁴⁹ Cf. <http://rda-deutschland.de>.

⁵⁰ Cf. http://www.forschungsdaten.org/index.php/AG_Forschungsdaten.

⁵¹ Cf. <http://www.langzeitarchivierung.de>.

⁵² Cf. <http://www.forschungsdaten.org>.

⁵³ Cf. <https://www.forschungsdaten.info>.

⁵⁴ Cf. <http://forschungslizenzen.de>.

⁵⁵ Cf. <https://de.dariah.eu>.

⁵⁶ Cf. <https://www.clarin-d.net>.

⁵⁷ Cf. <https://www.fdm-bayern.org>.

⁵⁸ Cf. <https://www.ratswd.de/en/publications/working-papers>.

⁵⁹ Cf. http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten_biodiversitaetsforschung.pdf.

⁶⁰ Cf. e.g. http://www.forschungsdaten.org/index.php/Data_Policies for a list.

⁶¹ Two examples: the BASE search engine for open access documents was developed and hosted in Bielefeld (as its full name reveals: “Bielefeld Academic Search Engine”; cf. <https://www.base-search.net>), and the Bielefeld library is one of the partners in the INTACT project monitoring APCs paid for open access journal articles (cf. <https://www.intact-project.org>).

⁶² Cf. <https://data.uni-bielefeld.de>.

⁶³ Cf. <http://www.eresearch.uni-goettingen.de>.

⁶⁴ Cf. <http://www.gcdh.de>.

⁶⁵ Cf. <https://www.fdm.uni-hamburg.de>.

⁶⁶ Cf. <https://www.fdm.uni-hannover.de>.

⁶⁷ Cf. <http://cceh.uni-koeln.de>.

⁶⁸ Cf. <http://dch.phil-fak.uni-koeln.de>.

⁶⁹ Cf. <http://gepris.dfg.de/gepris/projekt/379522012?language=en>.

⁷⁰ Cf. <http://www.esciences.uni-trier.de>.

- ⁷¹ Cf. <http://kompetenzzentrum.uni-trier.de>.
- ⁷² Cf. <http://www.fud.uni-trier.de>. “FuD” stands for “Forschungsnetzwerk und Datenbanksystem”, “research network and database system”.
- ⁷³ Cf. <http://www.escience.uni-tuebingen.de>.
- ⁷⁴ Cf. <https://www2.hu-berlin.de/edissplus/>; cf. e.g. also Weisbrod, Kaden, & Kleineberg, 2017.
- ⁷⁵ Cf. [https://www.re3data.org/search?query=&countries\[\]=DEU](https://www.re3data.org/search?query=&countries[]=DEU). Germany has the highest number among the European countries. But as the r3data platform has been developed and is run by a consortium of mostly German institutions the numbers could be influenced by the fact that German repositories are perhaps more likely to be listed than those based in other countries.
- ⁷⁶ Cf. <http://risources.dfg.de>.
- ⁷⁷ Cf. <https://www.datasealofapproval.org>.
- ⁷⁸ Cf. https://www.uni-muenster.de/ZIV/Kooperationen/IKM_kooperation.html and <https://www.uni-muenster.de/Rektorat/ikm/>.
- ⁷⁹ “Grundsätze zum Umgang mit Forschungsdaten an der Westfälischen Wilhelms-Universität Münster”, cf. <https://www.uni-muenster.de/Forschungsdaten/information/richtlinien/>; up until now only available in German, an English translation is on its way.
- ⁸⁰ Cf. <https://www.uni-muenster.de/Forschungsdaten/>.
- ⁸¹ Cf. <https://www.uni-muenster.de/news/view.php?cmdid=9529>.
- ⁸² Project “MIRO”; cf. https://www.ulb.uni-muenster.de/bibliothek/aktivitaeten/projekte/projekt_miro.html.
- ⁸³ Project “LIFE” 2013–2015; cf. <https://www.uni-muenster.de/Geoinformatics/en/research/projects/LIFE.html>.
- ⁸⁴ Project “o2r”, 2015–2017, cf. <http://o2r.info>. An application for the follow-up project “o2r2” 2018–2021 has been submitted.
- ⁸⁵ The same development can e.g. be seen at Trier University; cf. chapter 6.
- ⁸⁶ For information about the German “Universities Excellence Initiative” cf. e.g. <https://www.bmbf.de/en/excellence-strategy-5425.html> or https://en.wikipedia.org/wiki/German_Universities_Excellence_Initiative; for information about the WWU projects see <https://www.uni-muenster.de/forschung/en/profil/exzellenzstrategie/>.
- ⁸⁷ Cf. <https://www.uni-muenster.de/Religion-und-Politik/en/>.
- ⁸⁸ Cf. <https://www.uni-muenster.de/news/view.php?cmdid=9330>.
- ⁸⁹ Cf. https://www.uni-muenster.de/imperia/md/content/wwu/rektorat/dokumente/180417_hochschulentwicklungsplan.pdf.

⁹⁰ Cf. <https://www.ulb.uni-muenster.de/digital-humanities>, <https://www.uni-muenster.de/news/view.php?cmdid=8993>, <https://www.uni-muenster.de/news/view.php?cmdid=9040>.

⁹¹ “Research Data Management Organiser”, a web application project for planning, implementing and administering data in a scientific project; cf. <https://rdmo.aip.de>.

⁹² Cf. <https://www.uni-muenster.de/ZIV/Aktuell/2018/news25042018.html>. sciebo. RDS expands sciebo, a non-commercial cloud storage and sharing service developed for 28 universities in the state of North Rhine-Westphalia, led by the WWU IT Services; cf. <https://www.sciebo.de/en/project/>. Elements that will be added by sciebo.RDS are e.g. the encoding of and checksums for data or a refined roles and rights management.

⁹³ The WWU Institute for New Testament Textual Research (INTF) has e.g. commissioned the Cologne Center for eHumanities for a database project; cf. <http://ntg.cceh.uni-koeln.de>.

⁹⁴ Cf. <https://www.hbz-nrw.de/produkte/langzeitverfuegbarkeit/langzeitverfuegbarkeit-fuer-hochschulen> and <http://www.exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation/>.

⁹⁵ Cf. <http://www.library.ethz.ch/en/ms/Digitaler-Datenerhalt-an-der-ETH-Zuerich>.

⁹⁶ The three labels for the different stages in project consulting are taken from the Cologne Center for eHumanities; cf. e.g. Kronenwett, 2017, p. 33.

⁹⁷ Cf. <https://en.fh-muenster.de/bibliothek/>.

⁹⁸ Cf. e.g. the fact that the German librarianship journal “o-bib” got so many responses for a Call for Papers regarding RDM services in libraries that the editors decided to do two special issues on the topic instead of only one (cf. Stäcker & Steenweg, 2018).

⁹⁹ Cf. the answer of an Elsevier official replying to a statement arguing that access to information should not be limited to privileged people (<https://twitter.com/mrgunn/status/1028812448063664129>). For a discussion cf. e.g. <https://twitter.com/Protohedgehog/status/1028819653982736389>.