

Bulletin

Research Data Management

Mit Beiträgen von – avec des contributions de

John R Helliwell

**Silke Bellanger, Gérard Bagnoud, Eliane Blumer,
Carmen Jambé, Simone Rosenkranz,
René Schneider, Florian Steurer, Stefanie Strebel,
Clemens Trautwein, Matthias Töwe**

**Caterina Barillari, Henry Lütcke,
Ana Sesartic Petrus, Matthias Töwe,
Thomas Wüst, Bernd Rinn**

Professor of Animal Nutrition

The Department of Environmental Systems Science (www.usys.ethz.ch) at ETH Zurich invites applications for the above-mentioned position.

The new professor will be responsible for developing a leading research and teaching program that focuses on the sustainability and resilience of livestock nutrition systems. The professorship's main research topics may include: (i) maximizing nutrient utilization efficiency to meet nutritional needs of a growing human population while minimizing the environmental impacts of animal-based food production, (ii) resilient and diverse feeding systems, including the use of novel, human-inedible feedstuff, (iii) the impacts of feedstuff on the composition of rumen/gut microbiomes and on long-term animal health, or (iv) experimental approaches to precision animal nutrition to enable tailored individual feeding.

The professorship will be part of the Institute of Agricultural Sciences. It has full access to the well-equipped facilities of the research and teaching center AgroVet-Strickhof, which includes metabolic chambers suitable for animal phenotyping. Furthermore, the Functional Genomics Center Zurich offers state-of-the-art molecular user labs as well as bioinformatic support and service.

The successful candidate is expected to teach undergraduate level courses (German or English) and graduate level courses (English) in the Agricultural Sciences study program, covering both basic and advanced animal nutrition science

Please apply online: www.facultyaffairs.ethz.ch

Applications should include a curriculum vitae, a list of publications, a statement of future research and teaching interests, the names and contact details of three referees, and a description of the three most important achievements*. The letter of application should be addressed to the **President of ETH Zurich, Prof. Dr. Joël Mesot. The closing date for applications is 30 September 2020.** ETH Zurich is an equal opportunity and family friendly employer, strives to increase the number of women professors, and is responsive to the needs of dual career couples.

**ETH Zurich emphasises qualitative assessment of academic accomplishments. This is why you are kindly asked to submit a short description of your three most important achievements (maximum a half page each). Besides research findings, these could also be extraordinary achievements in teaching and its further development, services in the benefit of the academic community or society, software development, patents, knowledge transfer and its practical application, spin-off companies or similar.*

Editorial	2
Gernot Kostorz	

Research Data Management

Open science and reproducibility of the scientific record	3
John R Helliwell	
Forschungsdatenmanagement und Bibliotheken	11
Silke Bellanger, Gérard Bagnoud, Eliane Blumer, Carmen Jambé, Simone Rosenkranz, René Schneider, Florian Steurer, Stefanie Strebel, Clemens Trautwein, Matthias Töwe	
Supporting Scientists with FAIR Research Data Management – Initiatives at ETH Zurich and beyond	22
Caterina Barillari, Henry Lütcke, Ana Sesartic Petrus, Matthias Töwe, Thomas Wüst, Bernd Rinn	

Stellenausschreibung – Poste à pourvoir	ii
--	----

Inserat / Annonce	10
--------------------------	----



Editorial

Gernot Kostorz

Liebe Leserin, lieber Leser

Man kann die derzeit alles beherrschende Virus-Pandemie für vieles verantwortlich machen, aber nicht für das späte Erscheinen dieser Ausgabe oder ihren geringen Umfang. Der Hauptgrund ist, dass nicht nur die Ergebnisse der Forschung, sondern auch die Forschenden selbst, zu einem Tummelfeld der Metrologie geworden sind. So hat all das Vorrang, was zu einem messbaren Erfolg beiträgt, und freiwillige Leistungen werden immer rarer. Eine unglückliche Verkettung von Missverständnissen und nicht erfüllten Erwartungen hat grössere Pläne durchkreuzt, und wir bitten um Nachsicht.

Umso dankbarer sind wir den Autorinnen und Autoren der vorliegenden drei Aufsätze, die sich mit dem «Managen» von Forschungsdaten befassen.

John Helliwell, einer der Pioniere für «Open Data» im Bereich der Kristallographie, schildert die Entwicklung und Situation aus der Sicht des Forschers. Die Beiträge von Bellanger *et al.* und Barillari *et al.* geben einen Einblick in die sich in starkem Aufschwung befindlichen Infrastrukturen zur Unterstützung der Datenspeicherung und –bewirtschaftung in der Schweiz. Bibliotheken und IT-Einrichtungen geben in der Tat eine gesunde, neutrale Grundlage für die Pflege und den Umgang mit Forschungsdaten.

«Open Data» als Konzept ist nicht neu: Von der Frühzeit des wissenschaftlichen Publizierens bis zum Einsetzen des elektronischen Zeitalters war es in den Naturwissenschaften nicht unüblich, die Messdaten unbearbeitet mitzuliefern. Die Publikation meiner Doktorarbeit (1968) in einer wissenschaftlichen Zeitschrift enthielt noch alle Messkurven im Original. Zur Erfüllung des Kriteriums der Reproduzierbarkeit gehörte auch eine Beschreibung der wesentlichen Umstände der Datenerhebung. Mit dem Anwachsen der Datenmengen und der fortschreitenden Kommerzialisierung des Publikationswesens wurde zunehmend auf Details verzichtet. Mit der Möglichkeit der elektronischen Speicherung wuchsen auch die Möglichkeiten, Zusatzinformationen zu deponieren, entweder direkt im Zusammenhang mit der Publikation beim Verlag oder in eigens angelegten «Repositories». Daneben begannen in den 1970er Jahren speziell an Grossforschungsanlagen Aktivitäten zur Standardisierung der Datenspeicherung zum Zwecke der Wiederauffindbarkeit und Transportierbarkeit, mit dem Nebeneffekt gesteigerter Zuverlässigkeit.

Die heutigen Forderungen gehen aber weiter: Daten sollen nicht nur der wissenschaftlichen Gemeinschaft, sondern

allen Interessierten zur Verfügung stehen, zumindest, wenn sie mit öffentlichen Mitteln gewonnen wurden. Damit die Daten gratis zugänglich sind, müssen organisatorische und technische Strukturen geschaffen werden, die Neutralität und Kompetenz beim Speichern und Verwalten der Daten gewährleisten. Dieser Prozess ist noch lange nicht abgeschlossen, ja noch gar nicht voll überschaubar.

Begleitend zu den Entwicklungen, die mit den international anerkannten FAIR (für Findable, Accessible, Interoperable and Reusable)-Prinzipien im Einklang stehen, bleibt noch weitgehend unklar, wer für die Qualität der gelieferten Daten eintritt. Wenn John Helliwell berichtet (siehe S.7), dass in einem Fall von 41 eingereichten Manuskripten nach der Aufforderung, die Originaldaten zur Verfügung zu stellen, nur eines zufriedenstellend vervollständigt wurde, muss man nicht unbedingt schliessen, dass bei allen anderen keine oder nur mangelhafte Daten vorhanden waren. So trägt z.B. die Automatisierung der Messapparaturen, die teilweise mit verschlüsselten Routinen arbeiten, dazu bei, dass nicht alles berichtet werden kann. Es werden teils sogar Anlagen angepriesen, die wissenschaftliche Auswertungen produzieren, ohne dass eine fachlich ausgebildete Person mitwirken muss... Dennoch stimmen solche Ergebnisse nachdenklich und unterstützen die Forderung, dass Originaldaten verpflichtend geliefert werden müssen. Bei Publikationen wird dadurch die Begutachtung wesentlich zeitraubender, denn bei diesem Schritt muss eine Qualitätsprüfung stattfinden. Eine Datenspeicherung ohne Begutachtung durch kompetente Fachleute könnte zu einer Überflutung und Proliferation von Falschinformationen führen und sollte nicht unterstützt werden.

Mit diesen wenigen, persönlich gefärbten Bemerkungen wünsche ich Ihnen eine anregende Lektüre.

Mit den besten Wünschen

Ihr Gernot Kostorz

Open science and reproducibility of the scientific record

John R Helliwell*

Abstract

It is true to say that open science is a definite trend and reproducibility a growing concern. The former should however counter the latter. There are then complex features to the science research landscape. So, where is the process of science working well and where is it less than optimal? I will explore this landscape from my science base in crystallography and describe our ongoing initiatives. I will describe what can already be copied from crystallography as examples of best practice. The experiences of other fields of science that have come to my attention on these matters, such as at joint workshops, I will describe with some key conclusions.

1. Introduction

A great fraction of funded research is by the taxpayer and most funding agencies now espouse the principle that this research should be open. Also published research is generally assumed to be trustworthy, i.e. reproducible.

The USA National Academies of Science, Engineering and Medicine (2019) published an authoritative report on *Reproducibility and Replicability in Science*, a 256 page analysis and survey including a final chapter of 20 pages on Confidence in Science. Whilst naturally USA focused, it also had international participation in the study. Data and software transparency, i.e. openness, featured prominently in the Report in realising the best processes for achieving as high a scientific standard as possible. They made the important distinction of the two terms:-

- *Reproducibility* meaning obtaining consistent computational results using the same input data, computational steps, methods, code, and conditions of analysis (as the original researchers).
- *Replicability* meaning obtaining consistent results across (different) studies aimed at answering the same scientific question, each of which (research team) has obtained its own data.

In between these two procedures there is possible what I call robustness checks so that an existing set of data can be assessed with a different software package. Thus a variance of results estimate is obtained indicating further the reliability of a study.

These National Academy of Science definitions are at the core of the scientific method and the definition of science itself (Latin word *scientia* meaning to know).

If we know something it is no longer subjective. The closest we can come to objectivity as our evidence of what we know is the primary experimental data (the raw data). The slight caveat is that a piece of apparatus has to be calibrated by an experienced person. The publication contains a narrative, and although it includes a summary of what was done, it is necessarily the words and interpretation of the authors, i.e. subjective. Processing raw data with different softwares, thereby deriving models such as a crystal structure in my scientific field, has this possible *variance of results in the processed and derived data*. In the absence of a suitable word I offer a new word 'subjectivity', namely between the objectivity of the raw data and the subjectivity to some degree of the writer of the analysis software. Variation in analysis software performance can arise in cases such as estimating weak signals of various kinds and their uncertainty values. Sharing data to compare software packages can

* Department of Chemistry, University of Manchester, M13 9PL, UK.

E-mail: John.helliwell@manchester.ac.uk

[https://www.research.manchester.ac.uk/portal/en/researchers/john-helliwell\(8e39a79d-aa6d-4301-8b62-bf443581c009\).html](https://www.research.manchester.ac.uk/portal/en/researchers/john-helliwell(8e39a79d-aa6d-4301-8b62-bf443581c009).html)

ORCID 0000-0002-0520-7540



John R Helliwell, DSc (Physics, University of York), DPhil (Molecular Biophysics, Oxford University) is Emeritus Professor of Chemistry at The University of Manchester, where he served as Professor of Structural Chemistry from 1989 to 2012. Academic teaching from 1979 till 1988 was at the Universities of Keele and York. Researcher in the fields of Crystallography, Biophysics, Structural biology, Structural Chemistry and Data science. Also based at the SRS Daresbury Laboratory, in various periods of appointment between 1979 to 2008, including in 2002 as Director of Synchrotron Radiation Science based at the Council for the Central Laboratory of the Research Councils (CCLRC), Daresbury Laboratory. He is a Fellow of the Institute of Physics, the Royal Society of Chemistry, the Royal Society of Biology, and the American Crystallographic Association, an Honorary Member of the British Crystallographic Association and of the British Biophysical Society. He is a Corresponding member of the Royal Academy of Sciences & Arts of Barcelona, Spain and Honorary Member of the National Institute of Chemistry, Slovenia. His awards include the European Crystallographic Association Eighth Max Perutz Prize 2015, the American Crystallographic Association Patterson Award 2014, and the 'Professor K Banerjee Endowment Lecture Silver Medal' of the Indian Association for the Cultivation of Science (IACS) 2001. He published over 200 scientific research papers and several books, e.g. "Macromolecular Crystallography with Synchrotron Radiation" Cambridge University Press (1992), published in paperback 2005 and "Macromolecular Crystallization and Crystal Perfection" (by N E Chayen, J R Helliwell and E H Snell), Oxford University Press - International Union of Crystallography Monographs on Crystallography (2010).

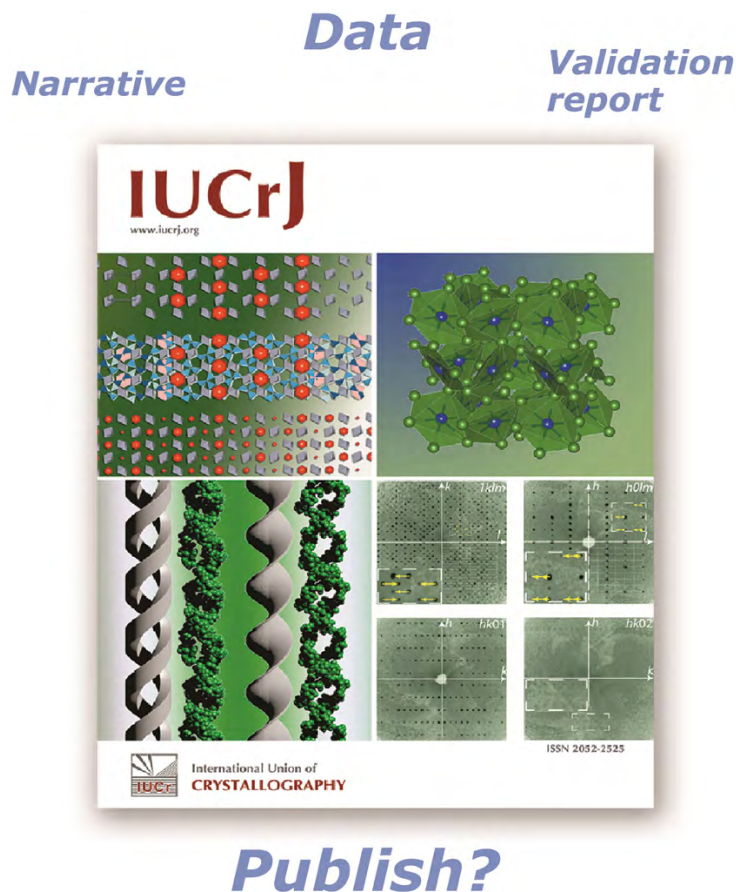


Figure 1. The Editor as the gatekeeper of reproducible science knowledge and on which basis replicability by others, and ultimately wisdom, can rely.

improve them and reduce variations in analysis. An early example of software improvement for processing of raw data (oscillation camera diffraction images recorded on film) was the comparative “round robin” project of Helliwell *et al.* (1981).

The National Academies Report also cogently states:

Unhelpful sources of non-replicability can be minimized through initiatives and practices aimed at improving research design and methodology through training and mentoring, repeating experiments before publication, rigorous peer review, utilizing tools for checking analysis and results, and better transparency in reporting.

2. The open science and reproducibility landscape in crystallography

In my own field of crystallography our community prides itself on its archiving of processed diffraction data and associated derived atomic coordinates. It is therefore one of the pioneers of the FAIR principles of data use (Findable, Accessible, Interoperable and Reusable, Wilkinson *et al.* 2016). An inspirational approach for striving for reproducibility is the International Union of Crystallography's (IUCr) chemical crystallography journals' refereeing processes. IUCr publishes nine journals; an example front cover is from

IUCrJ (see Figure 1). Thus an article narrative, a *checkCIF* report (more than 400 general checks of the data and metadata), and the underpinning processed diffraction data and derived coordinates data are assessed together by referees and editor. This is done in order to arrive at the *proper version of record* for each of these aspects of a study. In other areas of my science, this *openness* during the pre-publication peer review has not been so easy to secure, in my experience due to anxieties of the authors who can insist on confidentiality at that stage. Also, besides other areas of crystallography, how far might this exemplary *open* chemical crystallography procedure be extended into all other areas of science; e.g. do we just need to advertise it?

In another direction, we can ask:- are our crystallography laboratory leaders, who tend to be the ones asked for referee reports, still data analysis active? Clearly then they need to be if they agree to be a referee. Continual professional development should then include data analysis as a core learning objective. So, the IUCr Committee on Data has started running data science skills workshops such as the one at the European Crystallography Meeting held in Vienna in 2019.¹ The Vienna Workshop of course accepted participants who were at any stage of their careers. Aiming more at early career crystallographers the Associazione Italiana Cristallografia, along with several IUCr Committees, ran the School on Crystallographic information linked with data.²

A natural development is to extend the approach of refereeing of processed and derived crystallographic data with submitted articles to peer review of the primary, *i.e.* raw diffraction data. Thus IUCr Journals (Strickland, Helliwell and McMahon 2008) stated:

The IUCr is beginning to consider longer-term approaches to archiving the raw data, since these often contain additional information that is not fully utilized in the process of data reduction.

I define raw data as follows:

Those primary experimental data that have emanated from a properly calibrated device, *i.e.* detector, and which have been properly corrected for detector non-uniformities, non-linearities or spatial distortions, *etc.* Where it is a totally novel detector, of some level of unquantified performance and stability, then the data before such calibrations have been applied might need to be preserved.

¹ <https://www.iucr.org/resources/data/commdata/vienna-workshop>

² <http://www.cristallografia.org/contenuto/aics2019-crystallographic-information-fiesta/3311>

Raw data sets can however be very large compared with processed and derived data based on those raw data. Therefore this approach is pushing the limits of data archiving capacities and use of a referee's time. This also requires an expanded suite of metadata. Experiences with making diffraction image data available in crystallography are described by Kroon-Batenburg and Helliwell (2014). From 2011 to 2017 I chaired the IUCr Diffraction Data Deposition Working Group (DDDWG) and since 2017 I am Chairman of the IUCr's Committee on Data. The final report of the DDDWG³ had as its three top recommendations:

- Authors should provide a permanent and prominent link from their article to the raw data sets which underpin their journal publication and associated database deposition of processed diffraction data (e.g. structure factor amplitudes and intensities) and coordinates, and which should obey the 'FAIR' principles, that their raw diffraction data sets should be Findable, Accessible, Interoperable and Re-usable.⁴
- A registered Digital Object Identifier (DOI) should be the persistent identifier of choice (rather than a Uniform Resource Locator, URL) as the most sustainable way to identify and locate a raw diffraction data set.
- An archive of raw diffraction data sets for currently unsolved crystal structures should be pursued.

The first two are quite general and can be applied to any field. The third can be generalised by saying that in all fields of science there are likely to be problematic cases for which opening up the data for these would unleash the knowledge and expertise of a whole research community and maybe lead to a research breakthrough.

The IUCr has provided for open discussion of all matters to do with diffraction data deposition; the DDDWG from 2011 to 2017 and Committee on Data from 2017 onwards can be found at <https://forums.iucr.org/>. Individual postings have reached several thousands of downloads each.

A well-meaning effort by funding agencies is to require preprints so as to speed up research and discovery but does not include, thus far, attachment of the data, which is de facto problematic for ensuring reproducibility. However, if preprint servers would strongly encourage a preprint to include a DOI to the

underpinning archived raw data that would make preprints much more effective.

Some not-for-profit crystallography data repositories, namely the Cambridge Structure Database and the International Center for Diffraction Data, have survived through a user pays approach but is that open data? Open access has nuanced meanings here. The Protein Data Bank (PDB) and the Crystallographic Open Database (COD) are free to use databases including downloads of the whole PDB or COD. Single structures are open in the CSD. The completeness of databases, derived from the scientific literature, also varies. The value of a download of a whole database is evident for example with the PDB Redo project, now more than ten years old, which provides updated and optimised versions of existing entries of the PDB from its own DataBank. It provides then a remediation of historical entries, or of deficient in some way entries, or users can optimise their own structure model using the PDB Redo server even before deposition (Joosten *et al.* 2009, 2014). Joosten *et al.* 2009 also envisaged ultimately a role for the primary data in that process noting that

the Joint Center for Structural Genomics archive (<http://www.jcsg.org>) is an excellent prototype of what can be achieved in this respect, and its often-acknowledged value to software developers is a clear indication of the potential benefits of such an extended deposition scheme through the dual improvements it would enable in both the results for structures already solved and in the ability to solve new more difficult ones in the future thanks to better tested and better validated software advances.

A detailed presentation and overview of the trust in crystallographic data, in powerpoint slides style, is available (McMahon, Helliwell and Hester 2018).

3. The situation in some other science fields and special cases

Allison *et al.* (2016), who are a group of researchers working on obesity, nutrition and energetics, have emphasised how pre-publication peer review of an article with data is far superior to their post-publication peer review.

Anonymity of medical data is obviously needed for respecting proper data ethics. This leads to the oft-quoted statement about open science and open data that they should be as open as possible but closed where necessary. In the reporting of clinical studies, medical journals can then employ their own data statistician in the editorial office so as to vet the data accompanying the article. A discussion of this was

³ <https://www.iucr.org/resources/data/dddwg/final-report>

⁴ <https://www.force11.org/group/fairgroup/fairprinciples>

held in 2018 at a workshop entitled “Transparency, Recognition and Innovation in Peer Review in the Life Sciences”.⁵ This Workshop considered the following issues:

- Should journal peer review become a transparent and citable form of scholarly communication?
- Should scientists receive credit for peer review and, if so, how might this be achieved?
- What are best practices in peer review, how can they be spread? How can we train scientists in scholarly review?
- Is it possible to overcome inefficiencies and redundancies in peer review?
- Should reviewers be expected to review supporting datasets and code?
- Using new tools (e.g. preprints and the internet), are there new models for feedback/evaluation that could augment traditional peer review?

In this Life Science workshop there were commonalities of approach to those I have described in detail for crystallography. The responses of attendees to some of those questions depended strongly on the age of the person. A most prominent point made was by early career researchers that they wished to retain their anonymity. Otherwise they feared retribution from their seniors.

My own contribution as a blog in preparation for the Life Sciences workshop can be found in Helliwell (2017). My conclusions included:

So, what should publishers and funders be doing?

Clearly, as well as requiring data availability after publication, they should make clear in their policies that they value peer reviewers who sincerely try to attest to a submitted article’s conclusions by truly having access to the underpinning data.

What should professional associations be doing?

They should ensure training courses are given which are properly examined at the end of a course. These should be available to early, mid and late career researchers. I realise now, in my retraining and reskilling I have undertaken in my retirement, that in my late career I had become a manager of research with old, rather out of date, data skills. Continual professional development for a researcher is important at all stages of one’s career not least if they are going

to accept to do **proper peer review including the data.**

My article attracted one comment Lawrence (2018):

Great piece John, and completely agree with you that it is very hard to see how a referee can judge the conclusions of a paper without seeing the underpinning data. Just to add a point of clarity about Wellcome Open Research (and the other Open Research platforms that F1000 operate – and to declare my (conflict of interest) COI, I am Managing Director of F1000): you are absolutely correct that we require the data to be made openly available for all our articles, but because our peer review process takes place after publication, this means that all our referees can indeed look at the data as part of their evaluation process. We also specifically prompt our referees to do exactly that in the referee report form, with some examples of really valuable reanalyses and subsequent insights.

Another Workshop organised by the Research Data Alliance (RDA) Europe that I took part in brought various science disciplines together with the theme for the meeting of “*The opportunities and challenges of open data for research*”. The Workshop looked at three specific issues:

How do researchers share and work with other people’s data within their discipline and across disciplines?

What are the incentives and rewards to researchers for sharing their data?

What infrastructure tools and services are of particular use to researchers in their disciplines?

These themes RDA stated “*are inspired by the emerging European Cloud Initiative, which is drawing together research infrastructures and e-infrastructures across Europe.*” This Workshop allowed the sharing of current and best practice between disciplines including plasma fusion research and materials research as well as crystallography which I spoke on.

Also the developing world researcher may prefer to avoid openness of research data linked for example with their agriculture if that openness leads to a developing world data theft by the more developed world’s researchers who may be better funded with better facilities. This is itself a complex landscape for realising fairness and responsible practice (Ferris and Rahman 2017).

In the neurosciences is the recent startling report of Miyakawa (2020), Editor of Molecular Brain. He

⁵ <https://asapbio.org/peer-review/agenda>

described the situation of manuscripts that were submitted to him since 2017. He stated:

As an Editor-in-Chief of *Molecular Brain*, I have handled 180 manuscripts since early 2017 and have made 41 editorial decisions categorized as “Revise before review,” requesting that the authors provide raw data. Surprisingly, among those 41 manuscripts, 21 were withdrawn without providing raw data, indicating that requiring raw data drove away more than half of the manuscripts. I rejected 19 out of the remaining 20 manuscripts because of insufficient raw data. Thus, more than 97% of the 41 manuscripts did not present the raw data supporting their results when requested by an editor, suggesting a possibility that the raw data did not exist from the beginning, at least in some portions of these cases.

4. Future developments to increase new and better science results and publication

Science in general is experiencing considerable expansion of data flows. This is true in crystallography as we have, and welcome, the ever-improving synchrotron radiation and X-ray laser, as well as neutron facilities, and enhancement of detectors that is still ongoing. This means that simple prioritising of which data to analyse would lead to delays. But the funding agencies are pressing all researchers with policies demanding quicker results and certainly within 3 years after the measurement. Artificial intelligence (AI) and machine learning (ML) must be brought increasingly to bear to assist the researcher to comply with such new policies. The alternative, again with funding agency pressure, is that the researchers' data must be made open after a period, such as 3 years since measurement (see for example the European Synchrotron Radiation Facility Data Policy⁶). This policy is in line with European Space Agency requirements with which my PhD student and I had to comply in our microgravity experiments. Indeed PhD students may need protection of their own data during their training but maybe for longer than 3 years if they are delayed for some reason in completing their studies.

It is a challenging world to the researcher. In section 2 I also described the situation where making the raw data open could lead to making a breakthrough.

5. Evidence of my progress with pre-publication peer review

In my article (Helliwell 2018) I documented my experiences with applying the chemical crystallography refereeing approach to my particular field of biolog-

ical crystallography over a 1½ years period: “I have successfully refereed approximately twelve such article submissions in this time on this basis. These were from seven different journals and four different publishers.” In one case, I received the following letter text from the editor of a specialist chemistry journal, which I have anonymised:

Dear Reviewer,

This is just a short letter of thanks for your review of the manuscript by Prof. xxx for our Journal of yyy. We recognize how much time careful reviewing requires. Your time and effort are appreciated.

The quality and prestige of the journal can only be enhanced by the knowledge and advice of reviewers such as you.

Many thanks!

Sincerely,

In another case, the journal policy was to publish the peer review reports, also a good initiative. The pre-publication peer review reports on Langan *et al.* (2018) are therefore available to view at Nature Communications (2018). Not long after, Nature (2018) published its editorial entitled “*Referees should exercise their rights: Peer reviewers should not feel pressured to produce a report if key data are missing*”.

6. The role of international organisations in commitment to standards

A leading organisation for all matters to do with scientific data is the International Science Council (ISC)'s Committee on Data, ‘CODATA’.⁷ An important recent document in terms of open science in general and data in particular is *Open Data in a Big Data World* (2015) published jointly by the International Council for Science (ICSU), the InterAcademy Partnership (IAP), The World Academy of Sciences (TWAS) and the International Social Science Council (ISSC). The IUCr endorsed this accord, which provided a detailed and reasoned account of responsibilities and opportunities for all stakeholders in the modern data-rich scientific enterprise. The endorsement was supported by a position paper which provided details of how the crystallographic community in general, and the IUCr in particular, respond to the challenges laid out in the Accord (Hackert *et al.* 2016) and offers the view:

Science is best served when access barriers to data (and publications) are low. A major barrier to access is cost, and the phrase ‘open access’ is often used to characterize access to data

⁶ <https://www.esrf.eu/datapolicy>

⁷ The ISC is the organisation arising from the recent merger of the International Council for Science (ICSU) and the International Social Science Council (ISSC).

and publications that involve no charge to the end-user. However, the maintenance of the highest levels of quality in collecting, evaluating, storing and curating data is a very expensive component of the scientific process, and care must be taken to understand how to obtain the maximum benefit from public funding of science.

And:

Crystallography has a diverse ecosystem of disciplinary databases, data repositories, experimental facilities and publishers (Bruno et al 2017). Several of these are sustained through subscription-based access; but the other side of the coin is that they ingest, evaluate and publish data and information at no charge to the author/depositor, and without imposing any additional charge on the public purse. At the present time, this variety of approaches to sustainability and quality assurance serves this discipline well.

7. So, who pays and who will accept to do refereeing of 'article with data'?

So, I get asked these two questions that I have reiterated in my title of this section. I could have equally well entitled it feasibility of all that I advocate here. Let's do an options' analysis:

- A favourite of the Russian army general Mikhail I. Kutuzov, as described in detail in Tolstoy's *War and Peace*, is do nothing or rather in his case save lives by the strategy of retreat, which he did even after the victory at Borodino. In achieving final victory over Napoleon's army the Russian winter did the trick and Napoleon retreated. So, will *irreproducibility of science* go away if we do nothing? I think not and if trust is increasingly lost will taxpayers continue to fund us scientists, through Government funding agencies as their proxy? Do nothing is not a realistic option.

The other options are then a matter of scale and level of commitment:

- Journals could, rather simply actually, strongly encourage the review of the submitted article with data approach. Increasingly the handling scientific editor's name is included at the start of the publication so editor pride can be harnessed. In addition a gold star on a publication would indicate a full review of article with data had been done. This would also offer an alternative to the "gold star by impact factor" chasing by authors to get published in the highest impact factor journal, and which can cause considerable delays in publication.

Deans of Universities could encourage this new vision instead of demanding high impact factor publications only from their staff, a trend at least in the UK. Internationally the DORA (Declaration on Research Assessment⁸) pursues the same goal of not slavishly following journal impact factors.

- Funding agencies could provide a funding pulse for say three years to journals who would actively seek to employ a data specialist so that the burden would not be on referees but on the in house journal staff member. In that time frame hopefully journals pursuing such a methodology of pre-publication review would then self-fund the cost of the extra staff person. The UK's JISC (Joint Information Systems Committee, a membership organisation, owned by the sectors they serve such as UK higher education⁹), offers funding pulses such as helping a publisher promote open access publications.
- Funding agencies could guide its researchers away from journals that do not undertake article with data pre-publication review. A list, not unlike the list of predatory journals, could be readily assembled.
- Funders could also recognise laboratories that do early data sharing in urgent situations like the covid-19 pandemic. A marvellous example of data sharing with an explanatory, detailed, blog post, before even a preprint, is from the FraserLab and collaborators (2020).

The above four points offer options in a plan of action harnessing an increased openness to referees for enhanced reproducibility in science. In sections 2 and 4 I described the situation of opening up the raw data to a whole community where the research was at a stop or proceeding too slowly or speed was of the essence. An example of the latter situation is immediately to hand with the covid-19 pandemic and where in crystallography a great deal of prompt action and data sharing worldwide has been on the way (Arago et al. 2020, Kramer 2020, Blakeley and Schorber 2020, FraserLab2020), including prompt reanalyses based on those shared data being offered (Wlodawer et al. 2020). In such research challenges working together as a whole community has so much more to offer than a fragmentation of individual parts.

⁸ <https://sfidora.org/>

⁹ <https://www.jisc.ac.uk/about>

Acknowledgements

As Editor in Chief of IUCr Journals 1996 to 2005 I saw in detail the exemplary checks undertaken by Acta Crystallographica Section C Coeditors, which included Dr Madeleine Helliwell, using the checkCIF report as well as the chance to scrutinize with their own calculations the underpinning processed structure factors and derived atomic coordinates of each submitted article. Thus I am very grateful to all colleagues involved with Acta Crystallographica Section C at the time notably

Professor Syd Hall, University of Perth, Australia and also to Dr Madeleine Helliwell for many discussions.

As IUCr Representative to CODATA (2011 onwards) and to ICSTI (the International Council for Science and Technical Information) from 2005 to 2014 I met colleagues from many science disciplines and associated agencies which greatly enriched my experience of the ecosystem that is the amazing world of science. ■

References

- Allison, D. B., Brown, A. W., George, B. J. & Kaiser, K. A. (2016). *Reproducibility: A tragedy of errors*. *Nature*, 530, 27–29.
- Aragao, David, Brandao-Neto, Jose, Carbery, Anna, Crawshaw, Adam, Dias, Alexandre, Douangamath, Alice, ... Winter, Graeme. (2020). *Raw diffraction data for structure of SARS-CoV-2 main protease with Z44592329* (ID: mpro-x0434 / PDB: 5R83) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.3730610> [This is just one example of around 50 such raw data set depositions from this research team.]
- Blakeley, M. P. and Schorber, H. (2020). CERN Courier <https://cerncourier.com/a/neutron-sources-join-the-fight-against-covid-19/>
- Bruno, I., Gražulis, S., Helliwell, J.R., Kabekkodu, S.N., McMahon, B. and Westbrook, J. (2017). *Crystallography and Databases. Data Science Journal*, 16, 38. DOI: <http://doi.org/10.5334/dsj-2017-038>
- Ferris L and Rahman Z. *Responsible Data in Agriculture* [2017; version 1; not peer reviewed]. *F1000Research* 6:1306 (document) (<https://doi.org/10.7490/f1000research.1114555.1>)
- FraserLab and collaborators (2020). "Identifying new ligands for the SARS-CoV-2 Macrodome by Fragment Screening and Multi-temperature Crystallography" <https://fraserlab.com/macrodome/> with processed and derived data available at https://zenodo.org/record/3932380#.Xwg2Euco_tS.
- Hackert, M.L., van Meervelt, L., Helliwell, J.R. and McMahon, B. (2016). *Open Data in a Big Data World: A position paper for crystallography* available at <https://www.iucr.org/iucr/open-data>
- Helliwell, J. R. (2017). *Should reviewers be expected to review supporting datasets and code?* Blog post for the meeting entitled "Transparency, Reward, and Innovation in Peer Review in the Life Sciences" <https://asapbio.org/helliwell-datasets>
- Helliwell, J. R. (2018). Data science skills for referees: I biological X-ray crystallography, *Crystallography Reviews*, 24, 263–272.
- Helliwell, J. R., Achari, A., Bloomer, A.C., Bourne, P.E., Carr, P.D., Clegg, G.A., Cooper, R., Elder, M., Greenhough, T.J., Shaanan, B., Smith, J.M.A., Stuart, D.I., Stura, E.A., Todd, R., Wilson, K.S., Wonacott, A.J. and Machin, P.A. (1981). *Protein crystal oscillation film data processing: A comparative study* Acta Cryst. A37, C311–C312.
- Joosten, R.P., Womack, T., Vriend, G. and Bricogne, G. (2009). *Re-refinement from deposited X-ray data can deliver improved models for most PDB entries*. Acta Cryst. D65, 176–185.
- Joosten, R.P., Long, F., Murshudov, G.N. and Perrakis, A. (2014). *The PDB_REDO server for macromolecular structure model optimization*. IUCrJ 1(Pt 4):213–220.
- Kramer, D. (2020). *World's physics instruments turn their focus to COVID-19* Physics Today <https://physicstoday.scitation.org/doi/10.1063/PT.3.4470>
- Kroon-Batenburg, L. M. J. and Helliwell, J. R. (2014). *Experiences with making diffraction image data available: what metadata do we need to archive?* Acta Cryst. D70, 2502–2509.
- Langan, P.S., Vandavasi, V.G., Weiss, K.L., Afonine, P. V., El Omari, K., Duman, R., Wagner, A. and Coates, L. (2018). *Anomalous X-ray diffraction studies of ion transport in K⁺ channels*. Nat. Commun 9, 4540. <https://doi.org/10.1038/s41467-018-06957-w>
- Lawrence, R. (2018). *Comment to Helliwell, J. R. (2017)*. <https://asapbio.org/helliwell-datasets#comments>
- Miyakawa, T. (2020) Editorial: No raw data, no science: Another possible source of the reproducibility crisis *Molecular Brain*. <https://molecularbrain.biomedcentral.com/articles/10.1186/s13041-020-0552-2> DOI: 10.1186/s13041-020-0552-2.
- McMahon, B., Helliwell, J. R. and Hester, J. (2018). *The element of trust: validating and valuing crystallographic data*. https://www.iucr.org/_data/assets/pdf_file/0012/141222/6-mcmahon.pdf
- National Academies of Sciences, Engineering, and Medicine (2019). *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>.
- Nature Communications (2018) *Peer review report on Langan et al (2018)* https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-018-06957-w/MediaObjects/41467_2018_6957_MOESM1_ESM.pdf
- Nature Editorial (2018). *Referees should exercise their rights: Peer reviewers should not feel pressured to produce a report if key data are missing* <https://www.nature.com/articles/d41586-018-06006-y>

Science International Accord (2015). <http://www.science-international.org/#accord>

Strickland, P.R., McMahon, B. and Helliwell, J.R. (2008). *Integrating research articles and supporting data in crystallography*. Learned Publishing 21 (1), 63-72.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). *The FAIR Guiding Principles for scientific data management and stewardship* Sci. Data, 3, 160018.

Wlodawer, A., Dauter, Z., Shabalin, I. G., Gilski, M., Brzezinski, D., Kowiel, M., Minor, W., Rupp, B., & Jaskolski, M. (2020). *Ligand-centered assessment of SARS-CoV-2 drug target models in the Protein Data Bank*. The FEBS journal, 10.1111/febs.15366. Advance online publication. <https://doi.org/10.1111/febs.15366>

Inserat - Annonce

Sie müssen auf Englisch schreiben oder vortragen? Sie fragen sich, wie deutsche Wendungen, z.B. „aus Platzgründen muss ... ausgeklammert werden“ oder „...sei zunächst angemerkt, dass ...“ korrekt und geschliffen ins Englische übertragen werden? Oder Sie suchen nach Formulierungsalternativen?

Das „**Wörterbuch der allgemeinen Wissenschaftssprache / Wörter, Wendungen und Muster-texte**“ von Dirk Siepmann bietet Ihnen in komprimierter Form zahlreiche griffige und direkt umsetzbare Beispiele.

Gebundene Ausgabe, 352 Seiten, 34,90 Euro (D) inkl. Porto, für DHV-Mitglieder zum Sonderpreis von 29,90 Euro inkl. Porto. **Zu bestellen über:** Deutscher Hochschulverband, Rheinallee 18-20, 53173 Bonn, Tel. 0228 9026666, Fax 0228 9026680 oder per Mail: dhv@hochschulverband.de

Dirk Siepmann ist Professor für Fachdidaktik des Englischen an der Universität Osnabrück. Er verfügt über eine jahrzehntelange Erfahrung in Fremdsprachendidaktik, Übersetzungswissenschaft und Lexikographie.



**DEUTSCHER
HOCHSCHUL
VERBAND**

Forschungsdatenmanagement und Bibliotheken

Silke Bellanger*, Gérard Bagnoud**, Eliane Blumer***, Carmen Jambé**, Simone Rosenkranz^a, René Schneider^b, Florian Steurer^c, Stefanie Strebel^c, Clemens Trautwein^d, Matthias Töwe^e

Zusammenfassung

In den letzten 10 Jahren haben die meisten wissenschaftlichen Bibliotheken das Thema Forschungsdatenmanagement (FDM) aufgegriffen. Im Kontext ihrer Open Science Aktivitäten planen viele bereits früh, Services nicht nur für die freie Publikation von Forschungsartikeln zu bieten, sondern auch Angebote für die Veröffentlichung weiterer Forschungsergebnisse wie Forschungsdaten zu schaffen. Mittlerweile sind international und national die Vorhaben zu Dienstleistungen geworden, doch wie das Thema Forschungsdaten an sich, sind auch die Services der Bibliotheken weiterhin in Entwicklung und die jeweiligen Profile stark von den lokalen institutionellen und fachspezifischen Gegebenheiten geprägt.

Der in kurzer Zeit zusammengestellte Artikel bietet einen kursorischen Überblick über Forschungsdatenmanagement sowie die Angebote von Schweizer Bibliotheken in diesem Zusammenhang. Das Kollektiv der Autoren und Autorinnen stellt anhand ihrer jeweiligen Bibliotheken vor, mit welchen Dienstleistungen Bibliotheken Forschende beim Forschungsdatenmanagement unterstützen. Ähnliche oder weitergehende Angebote bieten auch die Teams der weiteren Schweizer Bibliotheken an, die aus zeitlichen Gründen nicht bei der Zusammenstellung dabei sein konnten.

1. Forschungsdatenmanagement – was ist das?

Die Europäische Kommission definiert Forschungsdaten als: «Dokumente in digitaler Form, bei denen es sich nicht um wissenschaftliche Veröffentlichungen handelt und die im Laufe von wissenschaftlichen Forschungstätigkeiten erfasst oder erzeugt und als Nachweise im Rahmen des Forschungsprozesses verwendet werden oder die in der Forschungsgemeinschaft allgemein für die Validierung von Forschungsfeststellungen und -ergebnissen als notwendig erachtet werden»¹. Dabei kann es sich um Statistiken, Versuchsergebnisse, Messungen, Beobachtungen aus der Feldarbeit, Umfrageergebnisse, Befragungsaufzeichnungen und Bilder oder Audio handeln. Auch Metadaten, Spezifikationen und andere digitale Objekte wie Software sind Teil davon.²

In den letzten Jahren ist die Verfügbarkeit und die Weiterverwendbarkeit eben dieser Forschungsdaten,

welche zum grössten Teil aus öffentlichen Mitteln finanziert werden, Gegenstand unterschiedlichster politischer Initiativen geworden. Die Hauptziele sind ein möglichst schneller, offener und digitaler Zugang und die kostenlose und uneingeschränkte Ver-

* Universitätsbibliothek Basel, Schönbeinstrasse 18-20, 4056 Basel.

E-mail: silke.bellanger@unibas.ch

<https://ub-easyweb.ub.unibas.ch/de/open-science-publizieren-digitale-forschung/>

ORCID 0000-0003-1952-1933



Silke Bellanger, M.A., MAS, Leiterin des Open Science Teams, ist wissenschaftliche Bibliothekarin und seit 2006 an Universitätsbibliotheken tätig. Die Schwerpunkte ihrer Arbeit sind Open Access, Forschungsdaten und Informationskompetenz. Sie hat einen M.A. in Soziologie, Geschichte und öffentlichem Recht und einen Master of Advanced Studies in Informations- und Bibliothekswissenschaften. Sie ist aktuell Mitglied des Schweizerischen Arbeitskreises Open Access und der Arbeitsgruppe für eine nationale Open Research Data Strategy.

** Université de Lausanne, UNIRIS – Ressources Informationnelles et archives, Quartier UNIL-Centre, Bâtiment Unicentre, 1015 Lausanne.

E-mail: gerard.bagnoud@unil.ch, carmen.jambe@unil.ch
www.unil.ch/uniris



Gérard Bagnoud, Diplômé de l'Institut de Hautes Etudes Internationales et du Développement, a suivi de nombreuses formations certifiantes en archivistique (de l'Association des archivistes suisses), en gestion électronique de documents et en Electronic Records Management (certification AIIM). Il a notamment travaillé aux Archives d'Etat de Genève, pour l'entreprise horlogère Rolex ainsi qu'au Pouvoir judiciaire genevois. Il est actuellement directeur du service des Ressources informationnelles et archives (UNIRIS) de l'UNIL depuis 2013. Au sein d'UNIRIS, on développe depuis 2015 un secteur d'activité autour de la gestion des données de la recherche et de l'Open Science. En parallèle, Gérard Bagnoud est professeur invité à la Haute Ecole de Gestion de Genève, filière information documentaire, et membre du Comité scientifique de la HES-SO en science de l'information. Il est également membre du Conseil consultatif du programme de formation continue en archivistique, bibliothéconomie et sciences de l'information des universités de Lausanne et de Berne. Als Absolvent des Graduate Institute of International and Development Studies hat Gérard Bagnoud zahlreiche Zertifizierungskurse in den Bereichen Archivierung (vom Verband Schweizerischer Archivare), elektronische Dokumentenverwaltung und Electronic Records Management (AIIM-Zertifizierung) absolviert. Er hat insbesondere im Genfer Staatsarchiv, für die Uhrenfirma Rolex und für die Genfer Justiz gearbeitet. Derzeit und seit 2013 ist er Direktor des Dienstes für Informationsressourcen und Archive (UNIRIS) an der UNIL. Innerhalb von UNIRIS wurde seit 2015 ein Aktivitätsbereich rund um Forschungsdatenmanagement und Open Science entwickelt. Gleichzeitig ist Bagnoud Gastprofessor an der Hochschule für Wirtschaft Genf (HES-SO) im Bereich Dokumentarinformation und Mitglied des wissenschaftlichen Ausschusses der HES-SO für Informationswissenschaft. Er ist auch Mitglied des Beirats des Weiterbildungsprogramms in Archiv-, Bibliotheks- und Informationswissenschaft der Universitäten Lausanne und Bern.
<https://ch.linkedin.com/in/gerardbagnoud>

¹ <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32019L1024>, Artikel 2, Punkt 9

² <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32019L1024>, Richtlinie, Punkt 27



Carmen Jambé, Bachelor of Science HES-SO en Information documentaire, diplômée en 2015 de la Haute école de gestion de Genève. Sur mandat du Service des ressources informationnelles et archives (UNIRIS) de l'Université de Lausanne (UNIL) et de son responsable M. Gérard Bagnoud, elle a réalisé le travail «La gestion des données de recherche à l'Université de Lausanne: enjeux transdisciplinaires». Suite à ce travail, elle a été engagée au sein du Service UNIRIS afin de développer le pôle Gestion des données de recherche, secteur dont elle est responsable. Depuis 2019 et suite à l'adoption par la Direction de l'UNIL de la Stratégie Open Science et plan d'actions 2019–2021, ce secteur, incluant notamment la thématique de l'Open research Data est intégré à l'Open Science Team de l'UNIL. Ses activités s'inscrivent dans les cinq axes prioritaires définis pour soutenir la communauté des chercheur-e-s de l'UNIL, fournissant ainsi conseil et soutien durant l'entier du projet de recherche et cycle de vie des données.

Bachelor of Science HES-SO in Dokumentarischer Information, Abschluss im Jahr 2015 an der Hochschule für Wirtschaft Genf. Im Auftrag des Service des ressources informationnelles et archives (UNIRIS) der Universität Lausanne (UNIL) und seines Leiters, Herrn Gérard Bagnoud, hat sie die Arbeit «La gestion des données de recherche à l'Université de Lausanne: enjeux transdisciplinaires» (Verwaltung der Forschungsdaten an der Universität Lausanne: transdisziplinäre Fragen) durchgeführt. Im Anschluss an diese Arbeit wurde sie innerhalb des UNIRIS-Dienstes eingestellt, um die Einheit für die Verwaltung von Forschungsdaten zu entwickeln, ein Bereich, für den sie verantwortlich ist. Seit 2019 und nach der Verabschiedung der Offenen Wissenschaftsstrategie und des Aktionsplans 2019–2021 durch das UNIL-Management ist dieser Bereich, insbesondere auch das Thema Offene Forschungsdaten, in das Offene Wissenschaftsteam der UNIL integriert. Seine Aktivitäten stehen im Einklang mit den fünf Prioritätsachsen, die zur Unterstützung der UNIL-Forschungsgemeinschaft definiert wurden, und bieten Beratung und Unterstützung während des gesamten Forschungsprojekts und Datenlebenszyklus.

*** Bibliothèque de l'Ecole Polytechnique Fédérale Lausanne, Rolex Learning Center, Station 20, 1015 Lausanne.

E-mail: eliane.blumer@epfl.ch
<https://www.epfl.ch/campus/library/services/services-researchers/rdm-quick-start/>

ORCID 0000-0002-0972-5396



Eliane Blumer, B. Sc. und M. Sc., ist seit 2017 Koordinatorin Forschungsdatenmanagement und Fachreferentin Life Sciences an der Bibliothek der EPFL Lausanne. Nach einem Bachelor in Informationswissenschaften und einem Master in Informationssystemen arbeitete sie in unterschiedlichsten Bibliotheken.

wendung der Daten. Die Vorteile davon sind eine verbesserte Qualität, mehr Transparenz und Effizienz in der Forschung. Für Forschende kann die Veröffentlichung der Daten zudem zusätzlich die Verbreitung und Wertschätzung ihrer Arbeit fördern.

Damit die Daten jedoch auch für andere Personen als den ursprünglichen Forschenden wiederverwendbar sind, müssen sie ein gewisses Mass an Qualität und Verständlichkeit vorweisen. Aus diesem Grund werden wissenschaftspolitisch und institutionell Bemühungen unternommen, dass Forschende bereits vor und während ihres Forschungsprojekts mittels

spezifischer Datenmanagementpläne überlegen, was mit ihren Daten passieren soll. Sie definieren dabei die Erhebung, die Aufbereitung und Analyse, die Interpretation und Dokumentation, die Publikation sowie die Langzeitarchivierung der Daten und halten ethische und rechtliche Probleme fest. Neben den Datenmanagementplänen unterstützen Grundsätze für gutes Forschungsdatenmanagement, wie sie z.B. die FAIR Prinzipien³ vorschlagen, die Daten so aufzubereiten, dass sie nicht nur verständlich, sondern auch für Menschen und Maschinen wiederverwendbar sind. So empfehlen die FAIR-Prinzipien Praktiken und Technologien, mit deren Hilfe Daten findbar, zugänglich, interoperabel und nachnutzbar gemacht werden können. Wie weiter unten deutlich wird, werden Forschenden je nach Organisation bei diesen Arbeitsschritten von unterschiedlichsten Abteilungen, wie zum Beispiel der Bibliothek, dem Forschungszentrum, der Informatik oder der Rechtsabteilung unterstützt.

2. Forschungsdatenmanagement und Data Science

Es gibt aktuell verschiedene Ansätze zum Umgang mit Daten – wie z.B. Big Data, Data Science oder eben Forschungsdatenmanagement. Da es nicht immer eine offiziell validierte Definition der jeweiligen Konzepte gibt, ist es schwierig, eine exakte Trennlinie zu ziehen. Etwas einfacher ist es, die nötigen Kompetenzen anzuschauen, die es jeweils braucht. «A broad range of data skills are needed across the research life-cycle. These cover data science skills to ensure research communities handle data appropriately and can exploit FAIR data resources, as well as stewardship skills to ensure data are FAIR, properly managed, preserved and shared. [...]»⁴ Wie Hodson et al. (2018) beschreiben, braucht es ein breites Skill-Set, um den gesamten Data Life Cycle abzudecken. Es werden sowohl Fähigkeiten zur Organisation der Daten gebraucht als auch technische Kompetenzen benötigt, wobei z.B. letztere klar in der Informatik anzusiedeln sind, wie eine Studie von Modulhandbüchern von Neuroth et al. gezeigt hat.⁵ Die benötigten Kompetenzen sind im Forschungsprozess besonders im Moment der Datenanalyse eng miteinander verschränkt. Hier braucht es sowohl technische Kompetenzen für Data Processing, Datenversionierung, -visualisierung oder Data Mining als auch Fähigkeiten, die Daten gut zu organisieren und zu kuratieren.

3. Internationale und nationale Entwicklungen

Wer mit Daten und deren Management zu tun hat, kommt auf dem internationalen Teppich sowohl an

³ «Findable, Accessible, Interoperable and Reusable», siehe <https://www.go-fair.org/fair-principles/>

⁴ Hodson, Simon et al., 2018

⁵ Neuroth et al. p. 424, 2019

der Research Data Alliance (RDA)⁶ sowie am Digital Curation Centre (DCC)⁷ nicht vorbei.

Die RDA wurde 2013 von der Europäischen Kommission, der National Science Foundation und dem National Institute of Standards and Technology der US-Regierung sowie dem Innovationsministerium der australischen Regierung als eine von der Gemeinschaft getragene Initiative ins Leben gerufen, um die soziale und technische Infrastruktur für die offene gemeinsame Nutzung und Wiederverwendung von Daten aufzubauen.⁸

Das DCC ist ein von mehreren englischen Universitäten geführtes Kompetenzzentrum für die Kuratierung digitaler Informationen mit Schwerpunkt auf dem Aufbau von Fähigkeiten und Fertigkeiten für die Verwaltung von Forschungsdaten.⁹

Beide Organisationen bieten eine Reihe von Hilfestellungen und Empfehlungen zum offenen und FAIRen Umgang mit Daten.

Auf europäischer Ebene spielt die Europäische Kommission eine wichtige Rolle für das Forschungsdatenmanagement. Die grossen Forschungsunterstützungsprojekte «H2020»¹⁰ und «Horizon Europe»¹¹ stützen sich auf Richtlinien, welche den öffentlichen Zugang der in den Projekten generierten Daten voraussetzen.¹²

Mit Blick auf wissenschaftliche Bibliotheken arbeitet eine Arbeitsgruppe des Verbandes Ligue des Bibliothèques Européennes de Recherche (LIBER) ebenfalls an Forschungsdatenmanagement-Richtlinien und steht in engem Austausch mit den bereits erwähnten Zentren und Verbänden.¹³

4. Was passiert in der Schweiz?

Seit 2017 verlangt der Schweizerische Nationalfonds (SNF) bei Projektanfragen einen Datenmanagementplan und den offenen Zugang zu Forschungsdaten, die einer Publikation zugrunde liegen. Der SNF folgt damit den europäischen Entwicklungen. Er unterstützt die Publikation von Daten mit bis zu 10'000 CHF pro Projekt. Auch wenn der Datenma-

^a Hochschule Luzern, Fachstelle Open Access, Leitung E-Science, Zentral- und Hochschulbibliothek Luzern, Frohburgstrasse 3, 6002 Luzern

E-mail: Simone.Rosenkranz@zhbluzern.ch
<https://www.unilu.ch/fakultaeten/tf/professuren/judaistik-und-theologie/mitarbeitende/simone-rosenkranz/>



Simone Rosenkranz, Dr. phil., studierte Judaistik, Islamwissenschaften und Philosophie an den Universitäten Luzern und Basel sowie an der Hebräischen Universität Jerusalem. 1999 promovierte sie mit einer Arbeit zu den jüdisch-christlichen Beziehungen im frühislamischen Orient. Von 1999 bis 2009 war Simone Rosenkranz Assistentin am Institut für Jüdisch-Christliche Forschung der Universität Luzern. Seit 2002 ist sie zusätzlich Fachreferentin an der Zentral- und Hochschulbibliothek, wo sie zu Beginn die Fächer Judaistik und Bibelwissenschaft, seit 2005 zusätzlich die Fachbereiche Religionswissenschaften, Wissenschaftsforschung (ab 2009) und Health Sciences (ab 2013) betreut. Simone Rosenkranz hat ausserdem den Lehrauftrag für Modernhebräisch I am Institut für Jüdisch-Christliche Forschung inne.

^b Haute école de gestion Genève, Rue de la Tambourine 17, 1227 Carouge.

E-mail: rene.schneider@hesge.ch
<https://www.hesge.ch/heg/annuaire/rene-schneider/actualites>

ORCID 0000-0003-4897-8561



René Schneider, Dr. phil., ist Professor und Studiengangsleiter für das Fach Informationswissenschaft an der Haute Ecole de Gestion in Genf. In seiner Forschung widmet er sich den Fragen des Forschungsdatenmanagements und insbesondere der Verbindung von Linked Data und persistenter Identifikation

^c Universität Zürich, Hauptbibliothek, Data Services & Open Access, Strickhofstrasse 39, 8057 Zürich.

E-mail: data@hbz.uzh.ch
<https://www.hbz.uzh.ch/de/open-access-und-open-science.html>



Florian Steurer, MA und Informatiker mit eidg. Fähigkeitsausweis, ist wissenschaftlicher Mitarbeiter Data Services IT & Infrastruktur. Florian Steurer hat eine Grundbildung als Informatiker und studierte Geschichte und Lateinische Philologie an der Universität Zürich. Seit 2016 arbeitet er im Bereich Forschungsdaten im Umfeld der Universität Zürich und seit 2018 im neu etablierten Team Data Services.



Stefanie Strebel, Dr. phil., studierte an der Universität Zürich Anglistik, Publizistik und Geographie und promovierte 2018 in Englischer Literaturwissenschaft. Sie ist seit 2019 im Team der Data Services der Universität Zürich tätig und dort für den Bereich Geistes- und Sozialwissenschaften zuständig

nagementplan (DMP) aktuell nicht Teil der Evaluation der Projektanträge ist, so muss er dennoch bei Einreichung zwingend vorhanden sein, ist während des Projektes zu aktualisieren und wird am Ende der Projektfinanzierung veröffentlicht.¹⁴ Mittlerweile sind Open Research Data, FAIR Data und Datenmanagementpläne ein Thema für alle Forschungs-

⁶ <https://rd-alliance.org/>

⁷ <https://www.dcc.ac.uk/>

⁸ <https://rd-alliance.org/about-rda>

⁹ <https://www.dcc.ac.uk/about>

¹⁰ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

¹¹ https://ec.europa.eu/horizon_europe_en/investing_to_shape_our_future

¹² <https://eur-lex.europa.eu/legal-content/DE/TEXT/?uri=CELEX:32019L1024>

¹³ <https://libereurope.eu/strategy/research-infrastructures/rdm/members/>

¹⁴ http://www.snf.ch/de/derSnf/forschungspolitische_positionen/open_research_data/Seiten/default.aspx

^d Hochschule Luzern – Technik & Architektur,
Technikumstrasse 21, Trakt V, 6048 Horw.

E-mail: clemens.trautwein@hslu.ch
<https://www.hslu.ch/en/lucerne-university-of-applied-sciences-and-arts/about-us/people-finder/profile/?pid=4520>



Clemens Trautwein, DI, Dr. nat. techn., erwarb 2013 an der Universität für Bodenkultur Wien das Doktorat in Umweltingenieurwissenschaften (Dr. nat. techn.), nachdem er 2007 an selbiger Universität das Diplom für Landschaftsplanung und Landschaftspflege abgeschlossen hatte. Seine fachliche Expertise vertiefte er in der Gewässerökologie und setzt nun seine Erfahrungen im analytischen Denken, Forschen und Publizieren als Fachreferent an der Hochschule Luzern, Bibliothek Technik & Architektur ein. Er berät und schult Forschende im wissenschaftlichen Publizieren, insbesondere für Open-Access-Publikationen, sowie im Forschungsdatenmanagement.

^e ETH Zürich, ETH-Bibliothek, Rämistrasse 101, 8092 Zürich.

E-mail: matthias.toewe@library.ethz.ch
<https://www.library.ethz.ch/FDM>

ORCID 0000-0001-9671-083X



Matthias Töwe, Dr. phil. nat., leitet die Gruppe Forschungsdatenmanagement und Datenerhalt an der ETH-Bibliothek. Nach dem Studium der Chemie in Hamburg und dem Doktorat an der Universität Basel (2001) absolvierte er die Ausbildung zum Wissenschaftlichen Bibliothekar an der Universitätsbibliothek Basel. Seit 2003 arbeitet er an der ETH-Bibliothek, zunächst für das Konsortium der Schweizer Hochschulbibliotheken im Projekt E-Archiving, später als Koordinator für das Programm «Elektronische Bibliothek Schweiz – e-lib.ch».

projekte in der Schweiz. Eine Zwischenbilanz des SNF zeigt, dass etwa 70% der Datenmanagementpläne, die eingereicht wurden, gut sind und 16% das Funding in Anspruch nehmen.¹⁵

swissuniversities und seine Förderprogramme spielen ebenfalls seit mehreren Jahren eine wichtige Rolle für die FDM-Bewegung in der Schweiz.¹⁶ Einige der geförderten Projekte, wie z.B. das Projekt DLCM zu Langzeitarchivierung von Daten,¹⁷ haben die Institutionalisierung von FDM an den Hochschulen und Universitäten sowie die Zusammenarbeit zwischen den Institutionen gefördert.

Aktuell erarbeitet swissuniversities im Auftrag des SBFJ eine nationale Open Research Strategie und einen dazugehörigen Aktionsplan. Komplementär zu der nationalen Open Access Strategie und dem Open Access Aktionsplan erfährt damit ein weiterer Baustein im Bereich Open Science einen national verankerten Rahmen.¹⁸

¹⁵ <http://www.snf.ch/en/researchinFocus/newsroom/Pages/news-200219-open-data-positive-interim-result.aspx>

¹⁶ <https://www.swissuniversities.ch/themen/digitalisierung/p-5-wissenschaftliche-information>

¹⁷ <https://dlcm.ch/>

¹⁸ <https://www.swissuniversities.ch/themen/digitalisierung/open-science>

Seit 2018 besteht in der Schweiz nun ebenfalls ein nationales Forschungsdaten-Netzwerk, das sich mehrheitlich aus Mitarbeitenden der FDM-Services der wissenschaftlichen Bibliotheken zusammensetzt. Die mehr als 40 Teilnehmenden tauschen sich regelmässig aus, treffen sich zwei Mal pro Jahr und führen gemeinsame Projekte durch.

5. Forschungsdaten und Bibliotheken

Dienstleistungen im Bereich von Forschungsdatenmanagement schliessen an die Angebote von Bibliotheken im Bereich von Open Science an. Open Science wird zumeist definiert als «(...) *the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.*»¹⁹ Prominent werden Open Access, Open Research Data und Open Educational Resources unter dem Umbrella-Term Open Science zusammengefasst.²⁰ Für die Realisierung von Open Access haben viele Bibliotheken mit dem Aufbau von institutionellen Repositorien für Publikationen begonnen und ihre Dienstleistungen dann auf weitere Publikationsplattformen, Publikationsberatung und Finanzierungsmöglichkeiten ausgeweitet. Bei den Entwicklungen der Services für Forschungsdatenmanagement verhält es sich eher umgekehrt. Insbesondere aufgrund der bereits erwähnten Vorgaben der Drittmittelförderer wie dem SNF, Datenmanagementpläne mit den Forschungsanträgen einzureichen, haben die meisten Bibliotheken mit allgemeinen Beratungstätigkeiten zu Datenmanagementplänen und den verschiedenen Schritten des Lebenszyklus von Forschungsdaten begonnen.

Aktuell bieten die meisten Bibliotheken folgendes Set an Dienstleistungen an:²¹

- Beratungs-, Kommunikations- und Vermittlungstätigkeiten entlang des gesamten Forschungsdatenzyklus
- Publikationsinfrastrukturen und –unterstützungen
- Archivierung von Forschungsdaten
- Digitalisierung und Erschliessung von (historischen) Beständen als Datensätze
- Aufbau und Pflege von praxisbezogenen Netzwerken von Forschenden, in denen kollegial Erfahrungen und Empfehlungen weitergegeben werden.

¹⁹ Foster, n.D.

²⁰ Für weitere Elemente von Open Science, siehe item.

²¹ Carol Tenopir et al. 2019. "Research Data Services in Academic Libraries: Where are We Today?" ACRL/Choice, publisher. <http://choice360.org/librarianship/whitepaper>.

6. Exkurs Literacies – von «Information, Media und Digital Literacy» zu «Data Literacy»

Die Vermittlungstätigkeiten im Bereich Forschungsdatenmanagement schliessen an bereits gut etablierte Schulungs- und Kursangebote von Bibliotheken an. Wissenschaftliche Bibliotheken stützen ihre Vermittlungstätigkeit mehrheitlich auf Konzepte der «Information Literacy» ab, ein Konzept der 1990er Jahre. Die American Library Association definiert Information Literacy (IL) als «[...] a set of abilities requiring individuals to 'recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information.»²² IL kann also als übergreifende Kompetenz aufgefasst werden, die sich auf den allgemeinen Umgang mit Informationen bezieht. Seit der Digitalisierung genügt dieser Ansatz jedoch nicht mehr. Es spielen vermehrt die «Media Literacy» – also der kritische Umgang mit nicht-textuellen Medien – sowie die «Digital Literacy» – der ethische und sichere Umgang mit Computer oder anderen technischen Ressourcen – eine nicht wegzudenkende Rolle. In den letzten Jahren hat zudem die vermehrte Zunahme von datenorientierten Aufgabenfeldern eine Notwendigkeit aufgezeigt, die IL um den Umgang mit Daten zu erweitern. Die Schlüsselkompetenz heisst «Data Literacy» und «[...] befähigt Menschen, reale Probleme durch die Nutzung, Analyse und Interpretation von Daten, die zugrunde liegenden Phänomene zu messen sowie zu adressieren.»²³

Kürzlich verfasste Berichte im Bereich der digitalen Hochschulreform zeigen anschaulich auf, dass Data Literacy sehr viele Facetten hat und dass Angebote an Hochschulen oder Bibliotheken dementsprechend auch sehr unterschiedlich ausfallen können.

Der «Knowledge Synthesis Report» der Dalhousie University hat zum Beispiel fünf Hauptkompetenzbereiche entwickelt, nämlich den konzeptuellen Rahmen, Data Collection, Data Management, Data Evaluation und Data Application und verankert darin 23 Kernkompetenzen sowie 64 Hauptaufgaben.²⁴

Im Report «Future Skills: ein Framework für Data Literacy Report» beschreiben Schüller et al. einen Kompetenzrahmen, welcher in fünf Kompetenzdimensionen aufgeteilt ist: (a) Wissen, (b) Fertigkeiten, (c) Fähigkeiten, (d) Motivation und (e) (Wert-) Haltung. Sie sollen zur Lernzieldefinierung und zur Veranschaulichung der interdisziplinären Natur von Data Literacy beitragen.²⁵

Was heisst das jetzt konkret für eine wissenschaftliche Bibliothek, welche in einem grösseren Bildungskontext eingeordnet ist? Wichtig sind die Fragen: Was kann die Bibliothek alleine machen, was macht sie vielleicht schon, und wo ist institutionelle Zusammenarbeit gefragt? Die zuvor erwähnten Berichte zeigen auf, dass «Data Literacy» weit gefasst ist. Die Bibliothek hat sicher ihren Teil dazu beizutragen, ohne intensive Zusammenarbeit zwischen Bibliotheken und Nutzern ist eine sinnvolle Weiterentwicklung jedoch nicht möglich.

7. Wie weiter?

Der weitere Aus- und Aufbau von Dienstleistungen im Bereich FDM ist in den Bibliotheken stark abhängig von der jeweiligen Ausrichtung der Institution, den vertretenen Fachdisziplinen und der Zusammenarbeit mit weiteren Serviceprovidern, wie z.B. Informatikdiensten oder Datenzentren, lokal wie überregional. Denn zum einen können keine Service-Angebote ohne die sehr enge Zusammenarbeit mit Forschenden erfolgreich sein, aber jede Fachdisziplin hat ihr eigenes Verständnis, was Forschungsdaten sind und die Kulturen und Konventionen des Austauschs und der Veröffentlichung von Forschungsdaten sind verschieden. Zum anderen berührt das FDM viele Kompetenzen, die idealerweise ergänzend von verschiedenen Service-Stellen an Hochschulen adressiert werden. Denn FDM liegt quer zu Tätigkeits- und Zuständigkeitsbereichen wie Datenschutz, High Performance Computing, Speicherinfrastrukturen, Informationssicherheit, Publikations- und Dokumentationsinfrastrukturen, Förder- und Wissenschaftspolitik. Entsprechend werden an den meisten Hochschulen und Universitäten Services für FDM gemeinsam von Bibliotheken und Informatikdiensten angeboten. Teilweise sind dafür neue Organisationseinheiten oder Teams gebildet worden, häufig erfolgt die Arbeit aber auch in quer zu den bestehenden Einheiten organisierten Netzwerken. Und nicht zuletzt hängt der Ausbau und das Angebot eng mit den personellen Möglichkeiten der Bibliotheken und weiteren beteiligten Service-Provider zusammen. Dabei ist die Anzahl an Personen genauso relevant wie das Wissen und Können der Beteiligten. Häufig werden die Aufgaben im Bereich Forschungsdatenmanagement zusätzlich zu bestehenden Funktionen übernommen und nicht immer werden neue Stellen dafür geschaffen. Zudem zeichnen sich aktuell zwar neue Berufsbilder wie «Data Librarians» ab, für die Ausbildungsmöglichkeiten entstehen, doch in der Regel haben gerade zu Beginn die meisten Bibliotheken und IT-Dienste parallel zu den Diensten auch die entsprechenden Kompetenzen aufgebaut.

²² <https://literacy.ala.org/information-literacy/>

²³ Schüller et al., 2019, p. 24

²⁴ <https://dalspace.library.dal.ca/handle/10222/64578>, Appendix 1, P. 38

²⁵ Schüller et al. 2019

8. Exkurs: Was braucht es für eine Ausbildung für zukünftige Mitarbeitende in FDM-Services in der Schweiz?

Stellt man die Frage nach den Ausbildungen für zukünftige Mitarbeitende in FDM-Services in der Schweiz, wird schnell klar, dass sich – charakteristisch für die Schweiz – kein einheitliches, sondern ein vielschichtiges und differenziertes Bild ergibt.

Zum einen muss es Fachpersonal geben, das umfassende Kompetenzen und ein tiefes Verständnis der fachlichen Materie besitzt, im Fall des FDMs sind dies informationswissenschaftliche Kenntnisse zum Umgang mit Daten und Metadaten, insbesondere bezogen auf Repositorien oder Langzeitarchive. Diese Kompetenzen lassen sich am besten unter dem neuen Begriff der Datenkuratierung zusammenfassen und werden bereits in Ausbildungen im Bachelor- und Masterniveau der dafür vorgesehenen Fachhochschulen adressiert: zum einen für die Deutschschweiz in Chur an der FH Graubünden²⁶, zum anderen für die Romandie an der Haute Ecole de Gestion in Genf.²⁷

Datenkuratierung, die als Schlüsselkompetenz gilt, ist der präzise und im Hinblick auf die auszuwählenden Repositorien beziehungsweise Archive zielgenaue Umgang mit Daten und Metadaten, die auch legale und sehr datentechnische Aspekte (wie etwa Embargos, Retention Schedules, Formate, Identifikatoren, Datenmigration usw.) umfasst, letztlich Fähigkeiten, die mit der Datensicherheit zu tun haben. Zum gesamten Vorgang gehört auch die Kompetenz des Evaluierens: Welche Daten werden überhaupt erfasst und wie lässt sich deren Bedeutsamkeit abschätzen?

Das Fachpersonal muss zudem darin geschult werden, Aus-, Fort- und Weiterbildungen für verschiedene Zielgruppen geben zu können, auch für diejenigen Berufsleute, die nach Abschluss ihrer Ausbildungen in der Praxis mit FDM zu tun haben.

Weiterhin führt der Begriff FDM die Managementkompetenz ja schon mit sich. Management beweist sich in Zeiten des Wechsels und in Krisen, und «Change Management» ist angesagt, auch um Forschungsdatendesaster zu vermeiden. Allerdings können nicht alle Manager sein, sonst fehlen die Personen, die die gemanagten Prozesse ausführen (sollen).

Aufgrund des teilweise sehr weiten Zeithorizonts ist auch die Kompetenz des Planens, wie sie sich an

der konzeptionellen Wichtigkeit der Datenmanagementpläne ablesen lässt, von besonderer Wichtigkeit, so dass sich in der Summe die Kernkompetenzen der Planung, Kuratierung und Evaluierung, der Sicherung und des Managings ableiten lassen, alle im Hinblick auf Daten und Metadaten.

All dies betrifft auch den Bereich der zertifizierenden Weiterbildungsstudiengänge, genauer die CAS, DAS und MAS-Weiterbildungen. Hier gibt es bereits eine Vielzahl von Angeboten für wissenschaftliche Bibliothekare, Archivare usw.; ein spezifisches Angebot zum Bereich Datenkuratierung, das auch zukünftige Datenbibliothekare oder Datenarchivare betrifft, bleibt aber noch ein Desiderat und sollte alsbald – vorzugsweise in einer konzertierten Aktion – angeboten werden, zumal der Bedarf und die Nachfrage als hoch einzuschätzen sind.

9. Aktuelle Herausforderungen für Ausbildung, Forschende, Institutionen und Bibliotheken

Als grösste Herausforderungen für die Ausbildung sind sicherlich fehlende Standards und die grosse Divergenz zwischen generischen, d.h. auf alle Disziplinen und ihre Daten anwendbare Methoden und den sehr spezifischen, beinahe von Datensatz zu Datensatz oder Projekt zu Projekt, in jedem Fall aber von Disziplin zu Disziplin divergierenden Prozeduren für das Datenmanagement zu nennen. Gleichzeitig zeichnet sich derzeit ein Unterschied ab zwischen einem eher an Open Science orientierten und sehr pragmatischen Ansatz im Forschungsdatenmanagement sowie einem eher technischen und auf Langzeitarchivierung fokussierten Ansatz.

Betrachtet man die Herausforderungen aus Sicht der Forschenden und Institutionen, so üben die Forschungsförderer als externe Treiber mit zum Teil ehrgeizigen Plänen zur Umsetzung von Open Research Data und FAIR Data seit einigen Jahren erheblichen Druck auf Forschende und Institutionen aus. Insbesondere Forschende am Beginn ihrer akademischen Laufbahn finden sich in einem Spannungsfeld wieder zwischen den Anforderungen der Geldgeber, den Erwartungen ihrer Institution und den impliziten oder expliziten Zwängen für eine erfolgreiche Qualifikation in ihrem Fachgebiet. Forschenden hier eine gewisse Orientierung mit Hinweisen für praxistaugliche Lösungen zu bieten, ist eine wesentliche Herausforderung. Das Zusammenwirken unterschiedlicher Services einer Institution kann dabei ein Vorteil sein, da unterschiedliche Kompetenzen zusammenkommen. Es bedeutet aber auch Aufwand bei der Kommunikation und Abstimmung von Verantwortlichkeiten – untereinander und gegenüber den Forschenden, für die die Arbeitsteilung nicht immer verständlich ist.

²⁶ <https://www.fhgr.ch/fhgr/angewandte-zukunftstechnologien/schweizerisches-institut-fuer-informationswissenschaft-sii/>

²⁷ <https://www.hesge.ch/heg/en/core-programmes/bachelors-science/information-studies>

In einer noch grösseren Perspektive kommt sowohl auf der Ebene der Wissenschaftspolitik als auch der Entwicklung konkreter technischer Plattformen die Herausforderung hinzu, den Anschluss an die nationalen und internationalen Entwicklungen wie die European Open Science Cloud (EOSC)²⁸ zu behalten. Dabei ist es zumindest phasenweise gar nicht einfach herauszufinden, was die realen Ergebnisse der gross gedachten Projekte sein könnten und welche Auswirkungen diese haben könnten.

Nicht zu vernachlässigen ist auch die bereits erwähnte Schnittstelle zwischen dem FDM und den Datenwissenschaften bzw. Data Scientists. Letztere produzieren und prozessieren Daten, sie sind aber zuvorderst auf prozessierbare Daten angewiesen, wie sie das FDM liefern kann und muss.

Sollten diese Herausforderungen angenommen werden, erscheinen die Möglichkeiten nahezu unbegrenzt.

10. Wer macht was – einige Beispiele

Im Folgenden werden die Dienstleistungen einiger Schweizer Hochschul- und Universitätsbibliotheken vorgestellt. Ähnliche oder weitergehende Angebote bieten auch die Teams der weiteren Schweizer Bibliotheken an, die aufgrund zeitlicher Gründe nicht am Artikel mitwirken konnten. Die aufgeführten Beispiele sollten daher auch als Anregung gelesen werden, die Webseiten und Dienste der eigenen Einrichtung zu konsultieren.²⁹

10.1. Forschungsdatenmanagement und Serviceangebote an der Hochschule Luzern (HSLU)³⁰

Die Bibliotheken der HSLU bieten FDM-Services seit der Einführung von obligatorischen Datenmanagementplänen bei SNF-Projektanträgen an. Die Grundlage für FDM wurde bereits 2014 im Reglement zur wissenschaftlichen Integrität für die HSLU gelegt.³¹ Punktuell unterstützen die zentralen IT-Services seit längerem Forschende beim Datenmanagement.

Seit 2017 arbeitet ein Team aus Fachreferenten/innen und weiteren bibliothekarischen Fachverantwortlichen zusammen, um Kompetenzen aufzubauen und FDM-Dienstleistungen anzubieten. Diese Gruppe kooperiert ausserdem mit der zentralen Ab-

teilung für Hochschulentwicklung im Bereich Forschung, den IT-Services und Business Applications.

Insgesamt sind etwa zehn Personen im FDM-Service mit insgesamt ein bis zwei Vollzeitäquivalenten involviert. Die Forschenden werden über interne Informationskanäle zu FDM informiert, sie können an Infoveranstaltungen der Bibliotheken teilnehmen oder sich direkt an die Bibliothek oder die IT-Services wenden.

Für Datenpublikationen in Open Access steht das institutionelle Repositorium LORY – basierend auf Zenodo – zur Verfügung. Seitens IT wird hochmoderne Infrastruktur (z.B. HPC oder Cloud Computing) zur Verfügung gestellt. Testweise wurden auch andere Anwendungen eingerichtet (z.B. Lab Notebooks).

Der Ruf nach einem institutionellen FDM-Service wird immer grösser. Deswegen läuft an der HSLU ein internes Projekt zur Realisierung eines FDM Tools, das die Forschenden in der Organisation im Datenlebenszyklus unterstützt. Der Schwerpunkt liegt dabei auf der Evaluation des Data Science Tools RENKU³² und der Verankerung des FDM Service Netzwerks zwischen den beteiligten Stellen innerhalb der Hochschule.

Die Heterogenität in der Forschung an der HSLU ist gross. Die FDM-Services müssen daher Basisbedürfnisse abdecken und auf Spezifika eingehen können. Deswegen ist eine «one-size-fits-all»-Lösung wahrscheinlich sehr schwierig umzusetzen. Die sich momentan im Aufbau befindlichen Lösungen auf nationaler und internationaler Ebene beobachten wir daher sehr genau. Der Aufbau eines zentralen FDM-Services ist eine Chance, die Kommunikation zwischen den involvierten Stellen und den Forschenden zu stärken.

10.2. ETH Zürich³³

An der ETH Zürich sind die ETH-Bibliothek (Gruppe Forschungsdatenmanagement und Datenerhalt, FDD) und die Scientific IT Services (SIS) Träger der FDM-Unterstützung (siehe auch den Beitrag von Barillari et al. in diesem Heft). Beide Abteilungen haben sich gewissermassen von entgegengesetzten Enden her des Themas FDM angenommen.

An der ETH-Bibliothek wurde ein durchdachtes Datenmanagement spätestens im Laufe des Aufbaus eines digitalen Langzeitarchivs ab 2011 zunehmend als notwendige Voraussetzung für die spätere Publikation und Nachnutzbarkeit von Forschungsdaten erkannt. Der Aufbau von Dienstleistungen erfolgte schritt-

²⁸ <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

²⁹ Siehe hierzu auch: <https://www.forschungsdaten.org/index.php/FDM-Kontakte>

³⁰ <https://www.hslu.ch/de-ch/technik-architektur/campus/bibliothek/forschen-publizieren/>; <https://www.zhbluzern.ch/dienstleistungen/forschen-publizieren/#WzU2NTMsNTY1MI0=>

³¹ Fachhochschulrat, Reglement zur wissenschaftlichen Integrität und zur guten wissenschaftlichen Praxis der Hochschule Luzern, FH Zentralschweiz, no. 524. 2015, pp. 1–7.

³² «RENU | SDSC.» [Online]. Available:

<https://datascience.ch/renku/>. [Accessed: 24-Jul-2020].

³³ <https://ethz.ch/services/en/service/a-to-z/research-data.html>

weise, bis erstmals 2015 eine Mitarbeiterin mit wissenschaftlichem Hintergrund ausschliesslich für die Unterstützung des Forschungsdatenmanagements eingestellt wurde. Als wichtigstes Werkzeug wurde 2017 die Research Collection als Repositorium für Veröffentlichungen und Forschungsdaten in Betrieb genommen (<https://www.research-collection.ethz.ch>).

Die Sektion SIS hat ihre Wurzeln im Informatik-Support des Departements Biosystems Science and Engineering (D-BSSE) mit entsprechend hohen Anforderungen hinsichtlich Datenmengen und der Handhabung der zahlreichen Abhängigkeiten zwischen den im laufenden Forschungsprozess eingesetzten Komponenten. Mit der Datenmanagementplattform openBIS entwickelte die damalige Gruppe ab 2007 eine leistungsfähige Umgebung primär für quantitativ arbeitende Forschende. Auf der Software basieren heute mehrere Services mit unterschiedlicher Ausprägung.³⁴

In der ETH-Bibliothek sind aktuell etwa 2.5 Stellen der Unterstützung des FDMs gewidmet, hinzu kommen Stellenanteile aus anderen Teams, vorrangig aus der Gruppe E-Publishing, welche die Research Collection betreibt und die Publikationsunterstützung leistet. Ebenso bieten Kollegen, die hauptsächlich mit der Betreuung des ETH Data Archive beschäftigt sind, Unterstützung z.B. bei Fragen der Langzeitarchivierung oder der Wahl von Dateiformaten. Bei Bedarf werden weitere Expertinnen beigezogen, etwa vom Hochschularchiv in der Bibliothek, z.B. für Fragen des Records Management, oder aus den Fachreferaten, z.B. zur Einschätzung von externen Repositorien aus fachlicher Sicht. Ausserhalb der Bibliothek sind der Rechtsdienst der ETH Zürich sowie die Technologietransferstelle ETH transfer immer wieder gefragt für die Einschätzung konkreter Fragen. Nicht zu vergessen sind die in den Departementen angesiedelten Informatik-Support-Gruppen (ISG) als Ansprechpersonen.

Die mit dem Forschungsdatenmanagement befassten Personen spielen somit in vielen Fällen auch eine wichtige Rolle als Vermittler und Lotsen in einem gerade für junge Forschende unübersichtlichen Umfeld und können so den Weg zu relevanten Informationen nicht selten stark abkürzen. Gleichzeitig gewinnen sie Schritt für Schritt selbst Know-how dazu. Sichtbar wird die breite Zusammenarbeit auch daran, dass alle genannten Stellen aktiv zum Programm der ETH RDM Summer School für junge Forschende des ETH-Bereichs beitragen.

Selbstverständlich ist die Einbindung so vieler Stellen nicht ganz einfach. Aus Sicht der ETH-Bibliothek er-

öffnen diese vielfältigen Kontakte aber weitere Perspektiven auf neue Formen und Wege der Zusammenarbeit mit Partnern innerhalb und ausserhalb der Hochschule. In diesem Sinn kann FDM auch für die Bibliothek als Institution eine Möglichkeit zur Weiterentwicklung über die naheliegenden Dienstleistungen hinaus bieten.

Neben der Teilnahme an verschiedenen Formaten von Trainings und Workshops nehmen Forschende laufend per E-Mail oder telefonisch Kontakt mit der Gruppe FDD auf. Für diese Erstkontakte sind meist die Informationen auf der Webseite ausschlaggebend, inzwischen kommt vermehrt auch Mundpropaganda hinzu, wobei die Teilnehmenden von Veranstaltungen ihrerseits als Multiplikatoren wirken. Nicht selten resultieren daraus Nachfolge-Workshops speziell für einzelne Gruppen.

An der ETH Zürich spielen Kooperationen mit Industriepartnern eine noch grössere Rolle als an anderen Universitäten. Die entsprechenden Vereinbarungen begrenzen meist die Möglichkeit, Daten aus so einer Zusammenarbeit unverzüglich oder überhaupt zugänglich zu machen. Ebenfalls noch grösser als an anderen Hochschulen ist der Anteil von Forschungsgruppen, die mit grossen Datenvolumina im mehrstelligen Terabyte-Bereich umgehen. Die damit verbundenen technischen und auch wirtschaftlichen Fragen im Hinblick auf die Möglichkeit zur Erhaltung und zur Nachnutzung sind nur teilweise gelöst. Es ist jedoch ein Missverständnis zu glauben, dass das Vorhandensein von genügend preiswertem Speicher in grossem Umfang alle Probleme lösen würde. Viele Herausforderungen des Datenmanagements müssen auch dann erkannt und bewältigt werden, wenn die Datenmengen überschaubar bleiben oder die Speicherung kein Problem darstellt.

10.3. Das Team Data Services der Hauptbibliothek der Universität Zürich³⁵

Die an der Hauptbibliothek angesiedelten Data Services der Universität Zürich werden seit 2018 in enger Kooperation mit der Zentralbibliothek und der S3IT der Zentralen Informatik angeboten. Das Team besteht neben der Leitung aus drei Mitarbeitenden, die jeweils einen unterschiedlichen Schwerpunkt (Naturwissenschaften und Medizin / Geistes- und Sozialwissenschaften / IT und Infrastruktur) betreuen. Für die Forschenden beinhaltet der Service hauptsächlich individuelle Beratung, Reviews von Datenmanagementplänen, sowie Schulungen und Workshops im Bereich FDM.

³⁴ <https://sis.id.ethz.ch/services/rdm/>

³⁵ <https://www.hbz.uzh.ch/de/open-access-und-open-science.html>

Seit Januar 2019 ist die Universität Zürich gemeinsam mit der Universität Lausanne Partnerin von SWISSUbase. Ziel ist der Aufbau eines schweizweiten Repositories. Die Projektleitung liegt bei FORS in Lausanne und verfügt über langjährige Erfahrung im Betrieb eines Repositories für die Sozialwissenschaften. Seitens der Universität Zürich möchte man das Projekt nutzen, um linguistische Daten zu verzeichnen und nachnutzbar zu machen. Dazu wurde eine enge Zusammenarbeit mit der an der UZH angesiedelten Linguistic Research Infrastructure (LiRI) beschlossen. Schwerpunkt wie Herausforderung stellte das Erstellen eines disziplinären Metadaten-schemas dar. Von Beginn weg orientierte man sich an den konkreten Bedürfnissen der linguistischen Community und diskutierte die Ergebnisse mit der Projektleitung von SWISSUbase, so dass inzwischen ein Kernmodell für linguistische Metadaten vorliegt. Ebenso arbeitete man bei der Lösungsfindung mit, wie diese Metadaten visuell angeordnet und dargestellt werden sollen. Da man die linguistischen Daten zu einem späteren Zeitpunkt im internationalen Virtual Language Observatory von CLARIN, dem europäischen Netzwerk für Linguistik, auffindbar machen möchte, stehen nun Abklärungen an, wie man ein geeignetes Metadaten-Profil für CLARIN erstellt. Gegenüber den Forschenden von LiRI nehmen die Data Services primär zwei Rollen wahr. Nebst den bereits geschilderten Tätigkeiten, bei denen es um konkrete Realisierungen wie dem Erstellen von Metadaten-schemata geht, agierten sie auch im Bereich des Anforderungsmanagements. So stellte sich früh heraus, dass man seitens der Linguistic Research Infrastructure an einer API für den automatisierten, skriptgesteuerten Ingest von Daten interessiert ist.

10.4. Das FDM-Team an der EPFL-Bibliothek in Lausanne³⁶

Erste Überlegungen zu einem Forschungsdatenmanagement-Service begannen 2012 in der Bibliothek, und zwei Teilzeit-Stellen unterstrichen die Wichtigkeit des Themas und verstetigten ein erstes Angebot. Im Jahr 2014 kam der Service dank des Open-Data-Pilotprojekts der Europäischen Kommission und neuer, ausschließlich FDM-bezogener Stellenbesetzungen richtig in Schwung. Der interne Start eines offiziellen FDM-Services erfolgte 2015 zusammen mit dem Research Office, dem Technology Transfer Office und den zentralen IT-Diensten. Im Jahr 2016 wurde das von swissuniversities P-5 unterstützte Projekt DLCM lanciert, welches zu einer Konsolidierung der FDM-Dienste in der Bibliothek führte. Nur ein Jahr später

unterstrich der neue Präsident der EPFL, Martin Vetterli, die strategische Bedeutung von Open Science und unterstützte somit das bereits existierende Angebot, genauso wie die Richtlinien des SNF rund um Open Data. Zu diesem Zeitpunkt entschied sich die Bibliothek, eine komplett neue Website einzurichten, um der Nachfrage gerecht zu werden. 2018 konnten weitere zwei Stellen für Data Librarians im FDM-Service besetzt werden. Eine Data Champions Community ermöglicht seit 2019 einen noch breiteren Austausch, und das laufende Jahr konzentriert sich auf die Entwicklung der technischen Publikations- und Archivinfrastruktur.

Gegenwärtig arbeiten fünf Mitarbeitende, insgesamt 2.7 FTE im FDM-Team. Alle haben einen unterschiedlichen Hintergrund, sind entweder ehemalige Forschende oder aber Informationswissenschaftler und Informationssystemexperten. Die enge interne Zusammenarbeit erstreckt sich auf die Zentral-IT / Fakultäts-ITs, das Forschungsbüro, die Technologietransferstelle sowie das Open Science Strategic Committee auf Präsidentenebene und dessen interne Projektausschreibungen. Zudem arbeitet die Bibliothek aktuell im swissuniversities-Projekt EasyFAIR³⁷ mit.

Das aktuelle FDM-Dienstleistungsangebot umfasst Beratung, Vermittlung und Unterstützung für den gesamten Data-Life-Cycle, einschliesslich Datenmanagementplänen, Datendokumentation und Metadaten, Datenpublikation, Lab Notebooks, Datenformate, Datenpublikation, Datenlizenzierung, Datenkuratierung und -archivierung. Alle Angebote können je nach Bedarf durch spezifische, personalisierte Schulungen auf einzelne Anfragen und Bedürfnisse oder durch die Buchung eines Datenbibliothekars ausgeführt werden. Darüber hinaus bietet die Bibliothek gut 15 regelmässige RDM-Schulungen pro Jahr an, die sich an verschiedene Zielgruppen richten (z.B. Doktoranden oder allgemeine Mitarbeiter) und unterschiedliche pädagogische Ansätze verfolgen (z.B. lange oder kurze Workshops oder Demos).

Um die Forschenden bei ihren täglichen Aufgaben zu unterstützen, wurden Datenstrategie- und DMP-Templates ausgearbeitet. Die «Data Champions Community» mit mehr als 30 Forschenden aus verschiedenen Disziplinen³⁸ steht für den fachlichen Austausch und als «Community of Practice» zur Verfügung und wird von der Bibliothek koordiniert. Aktuell arbeitet die EPFLBibliothek mit der EPFL-IT

³⁶ <https://www.epfl.ch/campus/library/services/services-researchers/rdm-quick-start/>

³⁷ <https://www.swissuniversities.ch/themen/digitalisierung/p-5-wissenschaftliche-information/projekte/swiss-mooc-service-1-1-1>

³⁸ <https://www.epfl.ch/campus/library/services/services-researchers/rdm-contacts-communities/epfl-data-champions/>

an der Weiterentwicklung der technischen Infrastruktur und hat sich zum Ziel gesetzt, demnächst ein Datenrepository und -archiv, als Erweiterung des Open Access Repository anzubieten.

10.5. Netzwerk Forschungsdatenmanagement an der Universität Basel³⁹

An der Universität Basel bietet seit 2018 ein Netzwerk universitärer Stellen Unterstützung zu FDM an. Zu dem Netzwerk gehören u.a. sciCORE – das universitäre Zentrum für wissenschaftliche Informatik, die Informatikdienste, die Datenschutzbeauftragte, das Departement für Klinische Forschung, das Grants Office, Ressort Forschung und die Universitätsbibliothek. Das Netzwerk ist ausserdem eng verbunden mit dem Center for Data Analytics und dem Digital Humanities Lab. Insgesamt sind rund 12 Personen regelmässig mit FDM befasst. Alle übernehmen diese Aufgaben zusätzlich zu ihren bestehenden Funktionen.

Das Netzwerk pflegt die gemeinsame Informationsplattform, bietet individuelle Beratungen an und führt regelmässige Trainings und Workshops durch. Neben wiederkehrenden fachübergreifenden Angeboten werden Workshops zu speziellen Problemen durchgeführt. So wurde gemeinsam mit dem Open Science Team der Universitätsbibliothek Bern⁴⁰ ein Workshop zu Datenschutz und qualitativen Forschungsdaten⁴¹ durchgeführt, in dessen Zentrum die Diskussion von Forschungsprojekten von Geistes- und SozialwissenschaftlerInnen standen. Aufbauend auf die Ergebnisse des Workshops werden Best Practices zu Datenschutzfragen für weitere Forschende zusammengestellt.

Zu den nächsten Schritten gehört zum einen, eine Plattform für den Austausch zwischen Forschenden zu schaffen. Als Anregung dienen hier die Data Conversations der Universität Lancaster,⁴² bei denen Forschende ihre Fragen und Lösungen zum Forschungsdatenmanagement vorstellen. Zum anderen sollen fachspezifische Ansprechpersonen im Rahmen eines internen Weiterbildungsprogramms in ihren Kenntnissen zu FDM gestärkt werden, um forschungsnaher Hilfestellungen leisten zu können. Dabei sind die

Programme für Data Stewards, wie sie in den Niederlanden bereits entwickelt werden, Vorbild.⁴³

10.6. Forschungsdatenmanagement an der Universität Lausanne⁴⁴

Seit 2015 ist es hauptsächlich die Abteilung UNIRIS, die FDM-Dienstleistungen anbietet. Unterstützt wird sie durch ein internes Netzwerk von Forschungsberatern (Fakultäten) und zentralen Diensten. Gemeinsam versuchen die Stellen, eine den Bedürfnissen entsprechende Unterstützung über den gesamten Lebenszyklus von Forschungsdaten anzubieten. Seit 2019 und der Verabschiedung der Open Science Strategie⁴⁵ und des Aktionsplans 2019–2021 der UNIL bilden UNIRIS und die Forschungsberatenden nun das Open Science Team.⁴⁶ Andere interne und externe Partner bieten ebenfalls FDM-Dienstleistungen an und arbeiten mit uns zusammen:

- BiUM⁴⁷,
- DARIS⁴⁸ (FORS),
- PlaTec⁴⁹ (DaSCH),
- Rechenzentrum & Computing and Research Support Division – DCSR; Fragen der Speicherung, Datensicherung, Datensicherheit, Forschungsinfrastrukturen, etc.⁵⁰
- PACTT UNIL-CHUV; geistiges Eigentum, Patente, Technologietransfer usw.
- Forschungsethikkommission - CER-UNIL⁵¹; Ethik, ethische Grundsätze und Regeln, Bewertung der Einhaltung von Normen, Konformitätsbescheinigungen usw.
- Juristischer Dienst; rechtliche Aspekte,
- Disziplinäre Forschungszentren, z.B. UNIL-EPFL dhCenter⁵²; spezifische FDM-Fragen.

Bei UNIRIS sind zwei unbefristete Stellen mit insgesamt 120% dem FDM gewidmet. Seit Juni 2020 vervollständigt zudem eine auf ein Jahr befristete 80%-Stelle das FDM-Team. Alle Fakultäten (mit Ausnahme der HEC) haben einen Forschungsberater, bei welchem ein kleiner Prozentsatz seines Pflichtenheftes dem FDM zur Verfügung gewidmet

³⁹ <https://researchdata.unibas.ch/en/home/>

⁴⁰ https://www.unibe.ch/universitaet/dienstleistungen/universitaetsbibliothek/service/open_science/index_ger.html

⁴¹ Iris Lindenmann (2019): Datenschutz und Anonymisierung von qualitativen Forschungsdaten. UB Basel Blog. <https://blog.ub.unibas.ch/2019/06/18/datenschutz-und-anonymisierung-von-qualitativen-forschungsdaten/>; Bellanger, Silke, Besmer, Christina, Kaufmann, Danielle, Keller, Anna, Lindenmann, Iris, Morger, Jennifer und Schreier, Gero. 2020. Handling Sensitive Data in Social Sciences and Humanities: Researchers' Questions, Legal and Ethical Considerations. Zenodo. doi:10.5281/zenodo.3669957.

⁴² <https://www.lancaster.ac.uk/library/research-data-management/data-conversations/>

⁴³ Siehe hierzu: Salome Scholtens, Petronella Anbeek, Jasmin Böhmer, Mirjam Brullemans-Spansier, Marije van der Geest, Mijke Jetten, Celia W G van Gelder. (2019). Function and competencies matrices for three types of data stewards 'policy', 'research' and 'infrastructure' (Version Version 2.1). Zenodo. <https://zenodo.org/record/3490855>.

⁴⁴ <https://www.unil.ch/uniris/fr/home.html>

⁴⁵ https://www.unil.ch/openscience/files/live/sites/openscience/files/OpenScience/Files/UNIL_Strategie_Open_Science_Plan_D_Actions_VF_2019.pdf

⁴⁶ <https://www.unil.ch/openscience/home/menuinst/who-are-we.html>

⁴⁷ <https://www.bium.ch/en/publication-open-access/data-management/>

⁴⁸ <https://forscenter.ch/data-services/>

⁴⁹ <https://platec.unil.ch/>

⁵⁰ <http://unil.ch/ci/dcsr>

⁵¹ <https://www.unil.ch/cerunil/home.html>

⁵² <https://dhcenter-unil-epfl.com/en/>

ist. Was das FDM in den Fakultäten, anderen zentralen Diensten und Partnern betrifft, ist es schwierig, die Stellenprozente zu schätzen.

UNIRIS bietet den Forschenden personalisierte Beratung (one-to-one, Telefon, E-Mail), insbesondere zu DMP (z.B. DMP des SNF), Recherche, Zusammentragen und Wiederverwendung vorhandener Daten, Dokumentation und Organisation von Daten, Datenspeicherung, -sicherung und -sicherheit (allgemeine Beratung und Weiterleitung an das Datenzentrum für spezifische Anfragen), Datenschutz und -sicherheit (allgemeine Beratung und Weiterleitung an das Datenzentrum bei spezifischen Anfragen), Veröffentlichung und gemeinsame Nutzung von Daten (offene Forschungsdaten), langfristige Speicherung, Archivierung und Aufbewahrung von Daten.

Zudem werden Workshops, Schulungstage oder Sensibilisierungsschulungen zu verschiedenen FDM- oder OS-Themen organisiert, wie zum Beispiel Informationsveranstaltungen in den Fakultäten zum DMP des SNF oder Workshops für Doktorierende. Ein Katalog von FDM- und OS-Online-Schulungskursen wird in den kommenden Monaten veröffentlicht.

Gegenwärtig gibt es technische Infrastrukturen für High Performance Computing (HPC) und Speicherung von sensiblen und nicht sensiblen Daten, Datensicherheitslösungen, elektronisches Dokumentenmanagement (EDM) administrativer Art, SWITCH-Infrastruktur und -Tools. Lücken gibt es noch bei der Entwicklung von elektronischen Lab Notebooks, virtuellen Forschungsumgebungen sowie einem Repository für die gemeinsame Nutzung und Langzeitarchivierung von Forschungsdaten. Dies ist auch ein Grund, weshalb UNIL Partner des swissuniversities-Projekts SWISS-Ubase⁵³ ist. Diese zukünftige Archivierungsinfrastruktur und Dienste zur Speicherung, gemeinsamen Nutzung und Förderung von Daten» werden der UNIL und ihren Forschenden ein Repository zur Verfügung stellen, das den Anforderungen des SNF entspricht. Die Beteiligung von FOR⁵⁴ an diesem Projekt ist ein grosser Vorteil, da FOR seit vielen Jahren ein ausgewiesener Experte für Datenmanagement und Archivierung ist. ■

⁵³ <https://info.swissubase.ch/fr/>

⁵⁴ <https://forscenter.ch/>

Literatur

Bellanger, Silke, Christina Besmer, Danielle Kaufmann, Anne Keller, Iris Lindenmann, Jennifer Morger, and Gero Schreier. 2020. Handling Sensitive Data in Social Sciences and Humanities: Researchers' Questions, Legal and Ethical Considerations. Zenodo. doi:10.5281/zenodo.3669957.

Coan, Matthew. 2014. "Research Guides: Information Literacy: Guide for Students: What Is Information Literacy?" 2014. <https://libguides.madisoncollege.edu/c.php?g=485946&p=3323409>.

Foster. n.D. "Open Science Definition." <https://www.fosteropenscience.eu/taxonomy/term/100>.

Hodson, Simon, Sarah Jones, Sandra Collins, Françoise Genova, Natalie Harrower, Leif Laaksonen, Daniel Mietchen, Rūta Petrauskaitė, and Peter Wittenburg. 2018. "Turning FAIR Data into Reality: Interim Report from the European Commission Expert Group on FAIR Data," June. <https://doi.org/10.5281/zenodo.1285272>.

"Information Literacy." 2017. *Welcome to ALA's Literacy Clearinghouse* (blog). February 7, 2017. <https://literacy.ala.org/information-literacy/>.

Lindenmann, Iris. 2019. "Datenschutz und Anonymisierung von qualitativen Forschungsdaten." UB Basel Blog. <https://blog.ub.unibas.ch/2019/06/18/datenschutz-und-anonymisierung-von-qualitativen-forschungsdaten/>

"Members of the Research Data Management Working Group." n.d. *LIBER* (blog). Accessed July 29, 2020. <https://libereurope.eu/strategy/research-infrastructures/rdm/members/>.

Neuroth, Heike, Laura Rothfritz, Vivien Petras, and Maxi Kindling. 2019. "Digitales Datenmanagement als neue Aufgabe für wissenschaftliche Bibliotheken," October. <https://doi.org/10.18452/20680>.

Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors. 2019. 172. Vol. OJ L. <http://data.europa.eu/eli/dir/2019/1024/oj/deu>.

Ridsdale, Chantel, James Rothwell, Michael Smit, Hossam Ali-Hassan, Michael Bliemel, Dean Irvine, Daniel Kelley, Stan Matwin, and Bradley Wuetherick. 2015. "Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report." Report. <https://doi.org/10.13140/RG.2.1.1922.5044>.

Scholtens, Salome, Petronella Anbeek, Jasmin Böhmer, Mirjam Brullemans-Spansier, Marije van der Geest, Mijke Jetten, Celia W G van Gelder. 2019. "Function and competencies matrices for three types of data stewards 'policy', 'research' and 'infrastructure' (Version Version 2.1). Zenodo. <https://zenodo.org/record/3490855>.

Schüller, Katharina, Paulina Busch, and Carina Hindinger. 2019. "Future Skills: Ein Framework Für Data Literacy." Zenodo. <https://doi.org/10.5281/zenodo.3349865>.

Tenopir, Carol, Jordan Kaufmann, Robert Sandusky, and Danielle Pollock. 2019. "Research Data Services in Academic Libraries: Where are We Today?" ACRL/Choice, publisher <http://choice360.org/librarianship/whitepaper>.

Supporting Scientists with FAIR Research Data Management – Initiatives at ETH Zurich and beyond

Caterina Barillari*, Henry Lütcke**, Ana Sesartic Petrus***, Matthias Töwe^a, Thomas Wüst**, Bernd Rinn**

Abstract

The FAIR (Findable, Accessible, Interoperable, Reusable) guiding principles for research data management have been established to advance data-driven science and are by now widely adopted among journals, funders and institutions. In order to comply with the FAIR principles, researchers must ensure that data are correctly documented and stored throughout the lifetime of a project. However, dealing with the ever-increasing amount and complexity of data and information generated in research labs requires new data management skills that researchers need to acquire at the beginning of their career. To this end, the commit-

ment of the academic institutions and funding agencies is of critical importance so that scientists can be provided with tools, education and support for data management. Here we present the ongoing efforts in this area at ETH Zurich, where the Scientific IT Services and the ETH Library support scientists with tools and services in data management. Similar efforts at the national level, with initiatives funded by swissuniversities, are also presented.

1. Introduction

Over the last decade, scientific research in many disciplines has become increasingly data-driven. By implementing novel data science approaches, such as artificial intelligence and machine learning, research processes are expected to become more efficient and effective. To fully leverage the power of data science approaches across different disciplines, research data must be accessible and interpretable, ideally in machine-readable formats. To this end, the FAIR guiding principles for scientific data management and stewardship have been established¹. The FAIR principles² outline key attributes and properties to make digital research outputs more Findable, Accessible, Interoperable, and Reusable. The principles focus on machine-actionability to enable automated computational exploitation of large and complex datasets. The FAIR principles have been globally endorsed and are widely adopted by journals, funders and institutions.

In order to comply with the FAIR principles, researchers must ensure that data are correctly documented and stored throughout the lifetime of a project. However, with the ever-increasing amount and complexity of data and information generated in scientific research labs, most of the traditional approaches to Research Data Management (RDM), based on file systems, spreadsheets and paper notebooks are no longer sustainable. Thus, management of research data is becoming a complex task that requires key skill sets that researchers should now acquire at the beginning of their career.

In most academic institutions, principal investigators are responsible for the data produced in their groups,

* ETH Zurich, Scientific IT Services, Mattenstrasse 26, 4058 Basel.

Corresponding author

E-mail: caterina.barillari@id.ethz.ch



Caterina Barillari, PhD, graduated in Chemistry and Pharmaceutical Technology at the University of Milan in 2001. After a short experience as medicinal chemist in the private sector, she pursued a PhD in Computational Chemistry at the University of Southampton, UK, studying the effects of water molecules in proteins' binding sites. She then held a Marie Curie fellowship in chemoinformatics at the CNRS in Strasbourg, France, and she later worked as chemoinformatician at the Institute of Cancer Research in Sutton, UK. Caterina Barillari joined ETH Zurich in 2011, where she currently holds a position as Manager of Research Data Management Services in the Scientific IT Service of ETH.

** ETH Zurich, Scientific IT Services, Weinbergstr. 11, 8092 Zurich.

E-mail: hluetcke@ethz.ch



Henry Lütcke, Dr. rer. nat., is a Scientific Computing and Research Data Management specialist at Scientific IT Services of ETH Zurich. He was trained as a neuroscientist at the Universities of Edinburgh (UK) and Göttingen (Germany). After a postdoc at the University of Zurich, he joined SIS in 2014. Besides supporting researchers in different aspects of data management and scientific computing, he is also project manager of the national openRDM.swiss project.

E-mail: twuest@ethz.ch

<https://sis.id.ethz.ch/>

ORCID: 0000-0001-6901-9277



Thomas Wüst, Dr. phil.-nat., is head of the group Computational & Data Science Support in Scientific IT Services (SIS) of ETH Zurich. He studied physics at ETH Zurich and did a PhD in computational physics at the University of Bern. Before joining SIS in 2014, he held positions as a computational scientist and high-performance computing specialist at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) and as a postdoctoral researcher at the Center for Simulational Physics of the University of Georgia, USA.

¹ Wilkinson, M. D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018 (2016).

² <https://www.go-fair.org/fair-principles/>

but often there are no precise guidelines in place for managing the data. Providing tools, education and support for RDM at the institutional level is thus becoming a necessity. At ETH Zurich (ETH), the executive board has recognized this need and has invested in initiatives for providing RDM services to ETH scientists, with the aim to facilitate the publication of data that adhere to the FAIR data principles. To cover the complete data life cycle, i.e. the different stages of data generation and re-use, this work is primarily led by two units: Scientific IT Services (ITS SIS³) and ETH Library.⁴ RDM services include adequate planning of which data will be generated and how they will be handled, data acquisition, data analysis, long-term data preservation and data sharing. While ITS SIS help and support scientists with RDM activities during an ongoing research project, ETH Library supports the later stages of data publication, long-term preservation and sharing. Both units also provide joint consulting on RDM planning as well as training.

The need for improving RDM expertise is recognized also on a national level in Switzerland. From 2015–2018, swissuniversities, the umbrella organisation of Swiss Higher Education Institutions, established the national project DLCM (Data Life Cycle Management),⁵ to organize and promote the development of tools and services for FAIR RDM in Swiss academic institutions. ETH Zurich, with ITS SIS and ETH Library, has been a partner in this project. Based on the success and the recognized importance of the DLCM project, two successor projects have been approved in 2018, one on active RDM (openRDM.swiss, led by ETH ITS SIS) and one on a national data repository (DLCM Phase 2, led by University of Geneva IT).

In this article we give an overview of services at ETH Zurich for RDM support along the stages of the research data life cycle with some conclusions for these functions in universities in general.

2. Overview of RDM services at ETH Zurich

At ETH Zurich, ITS SIS and ETH Library provide tools and services to help researchers with RDM during the different stages of the data life cycle (see Fig 1). Both units provide joint training and consulting in the early stages of a research project, with respect to data management planning. ITS SIS provides tools, services and trainings in active RDM, i.e. the daily management of data generated throughout a research project, from initial data acquisition to data processing and analysis. ETH Library provides tools, services and trainings

for the later stages of data publication and data sharing. Moreover, both units run joint trainings in RDM aimed at students and research staff.

3. Data management planning

In order to successfully manage research data, adequate planning of methodologies and resources is essential. Data management plans (DMPs) are now

E-mail: brinn@ethz.ch
<https://sis.id.ethz.ch/>



Bernd Rinn, Dr. rer. nat., is head of Scientific IT Services of ETH Zurich and a group leader at the Swiss Institute of Bioinformatics. He received his PhD in theoretical physical from the University of Konstanz. After finishing his PhD, he joined Genedata AG in Basel. End of 2006, he joined the Center for Information Sciences and Databases (C-ISD) at the Department of Biosystems Science and Engineering of ETH Zurich. There, he focused his work on designing and building the research data management system “openBIS”, which is widely used as a FAIR data management solution by research labs and projects today. From 2009 to 2013, Bernd Rinn led C-ISD. In 2013, he joined the Executive Board of ETH IT Services to establish the new section Scientific IT Services (SIS). SIS focuses on enabling compute- and data-intensive research at ETH Zurich and covers a wide range of scientific infrastructures and services, including high-performance computing infrastructure, scientific software engineering services, data science services for research projects, FAIR research data management, workflow solutions, as well as consulting and training on these topics.

*** No postal address available.

E-mail: ana.petrus@alumni.ethz.ch
<https://www.linkedin.com/in/anapetrus/>
 ORCID: 0000-0002-0928-8894



Ana Sesartic Petrus, Dr. sc. ETH, works as a freelance Sustainability Consultant. She studied Environmental Sciences at ETH Zurich and performed her doctoral studies at the Institute for Atmospheric and Climate Science until 2011. Postdoc positions in Berne (Institute of Geography, 2011/2012) and at the Center for Climate Systems Modeling at ETH Zurich (C2SM, 2012/2013) brought her to the Terrestrial Systems Ecology Group at ETH in 2013. From 2015 to 2020 she was strongly engaged in the creation of the services of the Research Data Management and Digital Curation group at ETH Library, where she acted as Data Management Consultant.

^a ETH Zurich, ETH Library, Sägmstrasse 101, 8092 Zurich.

E-mail: matthias.toewe@library.ethz.ch
<https://www.library.ethz.ch/FDM>
 ORCID: 0000-0001-9671-083X



Matthias Töwe, Dr. phil. nat., is head of the group Research Data Management and Digital Curation at ETH Library. After his studies in chemistry in Hamburg and his doctoral studies at the University of Basel until 2001, he was trained as a Scientific Librarian at the University Library Basel. Since 2003 he has been working for ETH Library, first for the Consortium of Swiss Academic Libraries in its project E-Archiving on digital long-term preservation, later as coordinator for the programme Swiss Electronic Library – e-lib.ch. Since 2010 he has been engaged in the implementation of the ETH Data Archive and in the creation of RDM services.

³ <https://sis.id.ethz.ch/>

⁴ <http://www.library.ethz.ch/en/>

⁵ <https://dlcm.ch/>

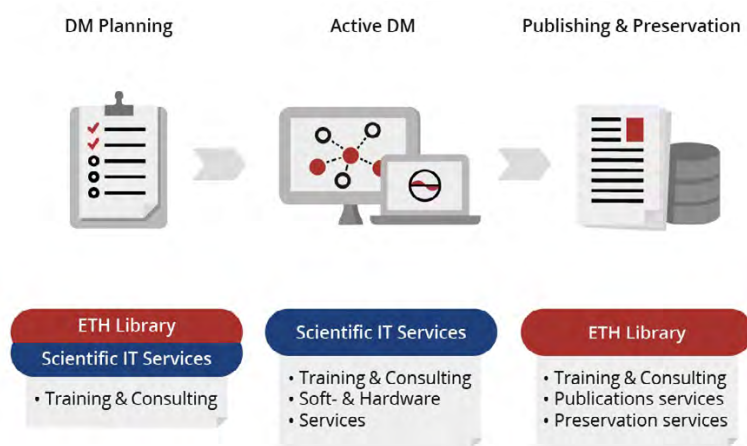


Figure 1. Stages of data life cycle support at ETH Zurich. Data management (DM) planning is supported by both library and IT services, the active DM phase is covered by the Scientific IT services and, finally, the publishing and preservation are supported by ETH Library.

increasingly requested by most funding agencies. In Switzerland, the Swiss National Science Foundation (SNSF) made DMPs mandatory for grant applications from October 2017. Horizon 2020 has required a DMP since January 2017. In the context of the DLCM project, the libraries of ETH and EPFL have elaborated a DMP template document with explanations and example answers to all questions of the SNSF DMP document, available to all Swiss scientists.⁶ ETH Library subsequently adapted this template to ETH specifications, for use by ETH scientists. In addition, ITS SIS provide a partially pre-filled version of the DMP for users of the ETH active RDM service to cover the non-project specific parts of the template. This has been greatly welcomed by researchers who had to fill in DMPs for their grant applications. Furthermore, ETH Library and ITS SIS regularly provide consultancy to researchers on DMP writing (see Fig2), a service that is highly valued by scientists.

4. Active Research Data Management

In order to publish research data in accordance with the FAIR data principles, data should be treated as FAIR from the start of a research project. This implies that data need to be annotated, organized and backed up from initial acquisition through to final results. Moreover, all steps and procedures to obtain the data must be well documented. Typically, researchers rely on data organization using a defined folder structure and file naming convention, possibly adding README files with data annotations if needed. Experimental and computational procedures recorded on paper or in a digital notebook are linked to the data stored on the file system. All other relevant information should be appropriately stored and referenced, as well. While

this approach may work well for individual researchers managing modest amounts of data, it does not scale well to very large and complex datasets as well as big research consortia. For these cases, RDM platforms have been developed that combine a digital notebook with a sample and protocol inventory. The main advantage of such platforms over traditional file-based approaches is that all data and related information is stored (or at least catalogued) in one central location with a management layer for tracking and retrieval on top.

Since 2007, ITS SIS have developed an open-source software solution for active RDM: openBIS.⁷ The solution covers storage, annotation and backup of data during the lifetime of a project. openBIS is a client-server application with a strong focus on providing an easy user interface, and well-designed APIs for system integration. The standard user client is web-based. The application was originally developed for managing large amounts of life-science data, such as microscopy, high content screening, proteomics, metabolomics, sequencing, as well as their derived processed and analysed data. In recent years the development has been focused on adding electronic laboratory notebook capabilities to document experiments and annotate data of labs of any size.

Furthermore, openBIS has been integrated with Jupyter notebooks and the JupyterHub server to allow direct analysis of data managed by the system.⁸ Analysis notebooks are connected to these data when they are stored in the database, thereby ensuring provenance tracking of the analysis results. In addition, openBIS can be used as an inventory of lab materials and protocols. Due to its generic underlying data structures, the usage of openBIS is now expanding to other quantitative research disciplines, such as environmental and material sciences. ITS SIS are working closely with these scientific communities to understand their needs and, when needed, adapt the software to accommodate them.

In accordance with the FAIR principles, all items in openBIS get a unique identifier, and all entities are indexed and searchable via the web-interface and the API. For use in the life sciences, a default metadata model is provided that has been developed with the user commu-

⁶ See <https://www.dlcm.ch/resources/dlcm-dmp>

⁷ Bauch, A. et al.: openBIS: a flexible framework for managing and analyzing complex data in biology research. *BMC Bioinformatics* 12, 468 (2011). Barillari, C. et al.: openBIS ELN-LIMS: an open-source database for academic laboratories. *Bioinformatics* 32, 638–640 (2016). <https://openbis.ch/>.

⁸ Kluyver, T. et al.: J. D. T. Jupyter Notebooks – a publishing format for reproducible computational workflows. *Position. Power Acad. Publ. Play. Agents Agendas* 87–90 (2016). doi:10.3233/978-1-61499-649-1-87

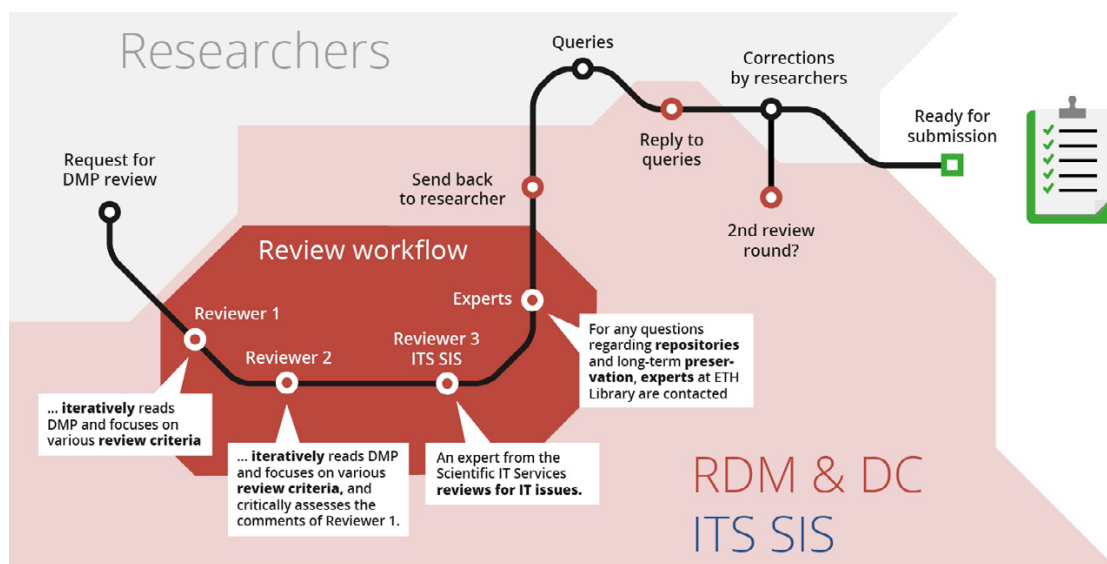


Figure 2. DMP review workflow at ETH Zurich, covered by the Scientific IT services (ITS SIS) and the RDM and Digital Curation (RDM & DC) team of ETH Library.

nity. Users can modify the defaults and customize the metadata information according to their needs. ITS SIS are collaborating with other scientific communities at ETH Zurich for implementing default metadata fields in openBIS for non-life-sciences disciplines too.

openBIS requires authentication and every operation performed by any user is registered in the database and is thus traceable. All data and metadata stored in openBIS can be accessed by authorized users with sufficient rights via HTTPS and/or SFTP via built-in servers. It is possible to delete bulk data without deleting the corresponding metadata. Moreover, even when metadata are deleted in the user interface, they are still stored in the database in 'history tables', and can be retrieved from there if need be. Data and data models can be annotated using native openBIS mechanisms and semantic annotations and they can be retrieved via the JSON-RPC API.

Adoption of a proper RDM solution in a research lab is time-consuming, implies a steep learning curve and requires a change in established procedures and work routines. To ease the process and increase adoption, training and support from technical experts with scientific knowledge and understanding are fundamental. For this reason, ITS SIS provide openBIS as a service to ETH research groups in two ways: The *ETH Research Data Hub* is available to all ETH research groups and provides a low entrance barrier (e.g. no extra cost). As an alternative for groups who need a more customized solution, a research group can also have their group-private openBIS installation, fully maintained by ITS SIS. In both cases, initial training and support for customization is provided to users and continuous support is offered afterwards.

5. Data publication and long-term preservation

One of the central roles of ETH Library is the support of the processes for publication and preservation of scientific data. RDM is a key prerequisite for effective data sharing and publication, which, in turn, can increase the visibility of scholarly work and its citation rates.⁹ Managing research data throughout their lifecycle is furthermore key to efficient long-term preservation, because the latter must rely on technical, administrative and rights metadata, as well as sufficient context information being available to make sure that data remain usable and understandable in the long run.¹⁰

ETH Library provides several tools for the publication and long-term preservation of data. The Research Collection is a FAIR repository for publications and data acknowledged by funding agencies, where ETH researchers can share and publish their work and/or their data free of charge.¹¹

In accordance with the FAIR principles, every item in the Research Collection receives a DOI. The rich metadata used ensure that data are findable, interoperable and reusable.¹² Data and metadata are accessible via the OAI-PMH protocol.¹³ When data are deleted, the metadata are still accessible. It is possible to determine access rights for datasets and the desired Creative Commons license can be selected.¹⁴

⁹ Piwowar, H. A., Day, R. S. & Fridsma, D. B.: Sharing detailed research data is associated with increased citation rate. *PLoS One* 2 (2007).

Piwowar, H. A. & Vision, T. J.: Data reuse and the open data citation advantage. *PerrJ Comput. Sci.* 25 (2013). doi:10.7717/peerj.175

¹⁰ Sesartic, A. & Töwe, M.: Research Data Services at ETH-Bibliothek. *IFLA J.* 42, 284–291 (2016).

¹¹ <https://www.research-collection.ethz.ch/>

¹² <https://documentation.library.ethz.ch/display/RC/Metadata+Schema>

¹³ <https://documentation.library.ethz.ch/display/RC/OAI-PMH+interface>

¹⁴ <https://documentation.library.ethz.ch/display/RC/Access+rights>

The ETH Data Archive on the other hand, is ETH Zurich's long-term preservation solution for digital information, such as research data, documents, software, images.¹⁵ All data that are uploaded to the Research Collection are automatically preserved in the Data Archive and additional tools exist for archiving data from other sources.

IT SIS and ETH Library also provide a close integration between openBIS and the Research Collection. This allows researchers to use openBIS in their daily work for annotation and processing data and finally select what should be published at the end of a research project. Selected objects can be automatically transferred to the Research Collection for publication. This will allow scientists at ETH to ensure that their data are Findable, Accessible, Interoperable and Re-usable.

6. Overview of Swiss initiatives towards FAIR data

Between 2015 and 2018, the Swiss Conference of Higher Education Institutions ("swissuniversities") funded the national DLCM (Data Life Cycle Management) project, to organize and promote the development of tools and services for FAIR RDM in Swiss academic institutions. Several partners from academic institutions across Switzerland were involved, including ETH Zurich, with ITS SIS and ETH Library. The project had tracks covering the different aspects of the data life cycle. Each track saw collaborative work between some of the different partner institutions.

One track focused primarily on guidelines and policies and this working group elaborated the DMP template document described above. Another track focused on active RDM, and part of the work involved a market analysis of existing commercial and open-source solutions for active RDM.¹⁶ This working group also prepared a guideline document on how to introduce an Electronic Laboratory Notebook (ELN) and Laboratory Information Management System (LIMS) in a laboratory, available from the same webpage. A third track focused on the development of a Swiss national data repository platform that adheres to the FAIR data principles and will be available to all scientists working in Swiss higher academic institutions. Finally, a fourth track was dedicated to education and training. Several trainings in RDM were carried out at the different partner institutions during the project and both materials and expertise were collected and exchanged. As an example, the RDM – Educational Material Collection by ETH Zurich showcases presentations developed at ETH.¹⁷

The main aim of the DLCM project was to set the basis for establishing national services in RDM that can be self-sustainable, which is often challenging for academic services. As follow up of the DLCM project, two more projects aimed at establishing national services were funded in the 2018 – 2020 funding round by swissuniversities: openRDM.swiss and DLCM Phase 2.

openRDM.swiss is led by ITS SIS and aims to establish a service for providing the active RDM platform openBIS to the Swiss academic community.¹⁸ Similar to the service provided at ETH, ITS SIS provide openBIS installations on the cloud infrastructure of the Swiss academic network provider SWITCH, and also provides user trainings and support. For institutions with more IT infrastructure and expertise, the project provides support and training to institutional IT administrators for setting up an on-premise installation of openBIS.

DLCM Phase 2 continues the development of the national data repository started in DLCM Phase 1 which is now being made available under the name OLOS.¹⁹ Moreover, it also established a national platform for RDM training.

7. Conclusion

Several factors are key to successful RDM. First, general education in RDM is of fundamental importance, in order to raise awareness of the topic. It is essential to target trainings at different organizational levels, i.e. management, professors, staff and students. ITS SIS and ETH Library run joint workshops on RDM, that cover various topics such as ETH regulations, RDM planning, active RDM, data sharing, data publication and long-term preservation.²⁰ Events are scheduled throughout the year, but often workshops are organized on demand for specific research groups, institutes or departments. Since summer 2019, an ETH Summer School in RDM is also jointly run by the two units. In addition, both units organize separate workshops and trainings more specific to their areas of competence. Second, adopting a software solution or a new workflow for the daily management of data generated in a research laboratory often requires a change in working mode and is associated with a steep learning curve. Key to this is a strong commitment from the principal investigators, acknowledging and accepting that the process requires a time investment. This time investment is still often perceived as time subtracted from research, and as such not valuable. In fact, time spent on setting up a professional RDM solution should be considered as an investment for the future, since hav-

¹⁵ <http://www.library.ethz.ch/en/ms/Research-Data-Management-and-Digital-Curation/ETH-Data-Archive>

¹⁶ A list is available at <https://www.dlcm.ch/services/dlcm-eln>

¹⁷ <https://www.research-collection.ethz.ch/handle/20.500.11850/296468>

¹⁸ <http://openrdm.swiss/>

¹⁹ <https://www.dlcm.ch/olos/about>

²⁰ <https://www.ethz.ch/services/en/service/a-to-z/research-data/support-and-training.html>

ing well annotated and organised data, easily searchable and retrievable, can save time in the longer run, i.e. when retrieving data for a publication or thesis.

Current efforts at ETH Zurich are aiming at overcoming this misconception. In our opinion, one of the main obstacles to the adoption of RDM solutions is that the benefits of using them are often not immediate, but become only evident with time. For this reason, we believe that having training and support from system experts can lower the barrier of adoption of such solutions, and reduce the time investment from the research labs, thus encouraging more people to use them. To this end, ITS SIS provide openBIS as a service to ETH research groups. This includes operating and maintaining the openBIS software on ETH infrastructure, as well as providing user training and support. Through the openRDM.swiss project, a similar service has now been established at the national level.

In addition, researchers should be supported with data publication and sharing, e.g. when making informed decisions on which data they can and should make available as FAIR data. ETH library has set up an extensive service portfolio in the field of open access and open data in order to support researchers in the open sharing of their articles and data.²¹

RDM is part of the quality assurance of research. Ideally, researchers should acquire the necessary skills during their training to empower them to implement appropriate RDM for their research projects. In the long run, this would be more sustainable than communicating RDM as an issue which is seemingly separate from the core scientific education. To this end, ETH Library, in collaboration with ITS SIS, will intensify its efforts to anchor the topic of RDM as early as possible in scientific careers and will also try to sensitize students to the topic. This includes offering a Summer School since 2019 that is rewarded with ECTS credits.²² During the Summer School, young research-

ers from the ETH Domain get to know the principles of RDM along the data life cycle and deepen their knowledge in practical examples and exercises. ETH Library and ITS SIS also strive to develop advisory services that are specially tailored to the needs of project leaders. Based on experience and exchanges with colleagues, we know that RDM can only work if it is supported by the management level. According to the Guidelines for Integrity in Research of ETH Zurich²³, it is the responsibility of research project leaders to ensure RDM within teams and that all team members are informed of what is expected from them.

Professional RDM is a means to an end and not a goal in itself. The aim must be to make it easier for researchers to access the methods and tools that can help them to carry out their tasks, as they are ultimately the experts on their own data. The main interest of researchers is to underpin excellent research with adequate RDM – and not vice versa. It is therefore central to mediate between the RDM requirements and the concrete needs of the researchers. This balancing act will challenge all those who support researchers in managing their data in the coming years.²⁴

The advent of interdisciplinary, data-driven research powered by novel data science methods promises important new scientific insights. To fully leverage the potential of this approach, however, will require the challenging implementation of RDM practices according to the FAIR data principles. ETH Library and Scientific IT Services are committed to supporting researchers at ETH Zurich and beyond with this challenge.

Acknowledgements

We thank Andres Bucher for his help in creating the figures. ■

²¹ <http://www.library.ethz.ch/en/ms/Open-access-at-ETH-Zurich>

²² <https://www.library.ethz.ch/en/ms/Conferences/ETH-Research-Data-Management-Summer-School-2020>

²³ ETH Zürich. Guidelines for Research Integrity. 31 (2011).

²⁴ Sesartic Petrus, A. & Töwe, M.: Forschungsdatenmanagement an der ETH Zürich: Ansätze und Wirkung. BIBLIOTHEK – Forschung und Praxis (2018). doi:10.18452/19633

Herausgeber und Verlag/Editeur: Vereinigung der Schweizerischen Hochschuldozierenden
Association Suisse des Enseignant-e-s d'Université
Associazione Svizzera dei Docenti Universitari

Generalsekretariat: Prof. Dr. Gernot Kostorz
Buchhalden 5, 8127 Forch
Tel.: 044 980 09 49 oder/ou 044 633 33 99 (ETHZ)

NEU → E-mail: info@vsh-aeu.ch
NEU → Homepage: www.vsh-aeu.ch

PC-Konto / ccp 80-47274-7 IBAN: CH87 0900 0000 8004 7274 7

Redaktion/Rédaction: Prof. Dr. Gernot Kostorz, Buchhalden 5, 8127 Forch
NEU → E-mail: info@vsh-aeu.ch

Layout: Print + Publish ETH Zürich, 8092 Zürich, E-mail: print-publish@services.ethz.ch

Druck/Imprimerie: Print + Publish ETH Zürich, 8092 Zürich

Anzeigen/Annonces: Generalsekretariat
Preise: Stellenanzeigen/Postes à pourvoir: CHF 250 (1/2 Seite/page), CHF 500 (1 Seite/page),
andere Annoncen/autres annonces: CHF 500/1000

**Mitgliederbetreuung, Adressen/
Service membres, adresses:** Generalsekretariat

Das Bulletin erscheint drei- bis viermal im Jahr und wird gratis an die Mitglieder versandt.
Abonnements (CHF 65 pro Jahr inkl. Versand Schweiz) können beim Verlag bestellt werden.
Le Bulletin apparaît trois à quatre fois par an et est distribué gratuitement aux membres.
Des abonnements sont disponibles auprès de l'éditeur (CHF 65 par an, frais de port compris en Suisse).

Vorstand/Comité directeur am 1. August / au 1^{er} août 2020

Präsident/Président: Prof. Dr. sc. nat. Christian Bochet, Université de Fribourg, Département de Chimie,
Chemin du musée 9, 1700 Fribourg, Tel.: 026 300 8758, E-mail: christian.bochet@unifr.ch

Vorstandsmitglieder/Membres du comité: Prof. Dr. Nikolaus Beck, Università della Svizzera italiana, Institute of Management,
Via G Buffi 13, 6900 Lugano, Tel.: 058 666 44 68, E-mail: nikolaus.beck@usi.ch

Prof. (em.) Dr. phil. Hans Eppenberger, Wiesenweg 5, 5436 Würenlos,
Tel.: 056 424 3256, E-mail: hans.eppenberger@cell.biol.ethz.ch

Prof. Dr. Norbert Lange, Université de Genève, Université de Lausanne, Ecole de Pharmacie, Sciences II,
Quai Ernest Ansermet 30, 1211 Genève 4, Tél.: 022 379 33 35, E-mail: norbert.lange@unige.ch

Prof. Dr. (Ph.D.) Stephan Morgenthaler, Ecole Polytechnique de Lausanne (EPFL),
Fac. Sciences de base (SB), Inst. de mathématiques (IMA), MAB 1473 (Bâtiment MA),
Station 8, 1015 Lausanne, Tél.: 021 6934232, E-mail: stephan.morgenthaler@epfl.ch

Prof. Dr. med. Dr. phil. Hubert Steinke, Universität Bern, Institut für Medizingeschichte,
Bühlstrasse 26, 3012 Bern, Tel.: 031 631 84 29, E-mail: hubert.steinke@img.unibe.ch

Prof. Dr. Helmut Zander, Université de Fribourg, Theologische Fakultät, Dep. Glaubenswissen-
schaft und Philosophie,
Av. de l'Europe 20, 1700 Fribourg, Tel.: 026 300 74 38, E-mail: helmut.zander@unifr.ch

Herausgegeben mit Unterstützung der Schweizerischen Akademie der Geistes- und Sozialwissenschaften (SAGW)

Publié avec le soutien de l'Académie suisse des sciences humaines et sociales (ASSH)



Mitglied der Schweizerischen Akademie
der Geistes- und Sozialwissenschaften
www.sagw.ch

Vereinigung der
Schweizerischen Hochschuldozierenden



Association Suisse
des Enseignant-e-s d'Université



Membre de l'Académie suisse
des sciences humaines et sociales
www.assh.ch

***Die Stimme
der Hochschuldozierenden***



***La voix
des enseignant-e-s d'université***