Research Data Management

Jochen Apel

https://orcid.org/0000-0002-0395-4120

Abstract This chapter introduces the area of research data management. Based on the FAIR Data Principles, the chapter outlines the main features of various aspects and aims of data management – from project planning and data organization in the project to the publication and preservation of research data.*

Keywords Research Data, Research Data Management, FAIR principles, Humanities, Digital Humanities

1. Research Data and Research Data Management

Structured and planned handling of research data is a central requirement for every research project, of course also in the field of theology. With the increasing use of digital tools and methods, the requirements for effective and sustainable data management that ensures the quality, traceability, and reusability of research results are growing. This chapter explains the basic principles of research data management (RDM). The term research data is understood in the following way:

> Research data are digital data generated, collected, or compiled in the research process and on the basis of which scientific hypotheses, models, or theories are formed and confirmed or refuted.[1]

According to this definition, digital data become research data through the specific *epistemic roles* they play in the research process. This seems particularly relevant for research in the humanities, where, in many cases, data is not generated as part of a research project, but is already available, *only becoming research data* through scien-

* This chapter, including quotations in foreign languages, was translated from German by Brandon Watson, and revised by Kevin Wunsch.

1 The author of this chapter first formulated proposed definition for a local website in 2014. Cf. https://web.archive.org/web/20230320185206/https://data.uni-heidelberg.de/faq.html (Accessed: 19 June 2024).

Jochen Apel: Research Data Management, in: Christopher A. Nunn and Frederike van Oorschot (eds.): Compendium of Computational Theology 1. Introducing Digital Humanities to Theology, Heidelberg: heiBOOKS, 2024, S. 393–405. DOI: https://doi.org/10.11588/heibooks.1521.c21961

tific inquiry. If source material becomes the subject of research, the corresponding digital data becomes research data. The proposed definition is purposefully relatively broad. It can encompass the following types of data: digital sources and digitized versions of sources (text, image, video, and video files or 3D data), edits of source (e.g., an OCR-generated text or an edition encoded in TEI-XML), results of analyses (e.g., statistical results of a quantitative analysis), or databases in which information is structurally compiled.

The relevant literature shows many attempts to define research data. These attempts are summarized by Geiger (2023). The landscape of research data in the humanities is a heterogeneous field in which different data types, formats, and structures are relevant and in which there are only established standards to a certain extent or in certain sub-areas (cf. Pempe 2012).

Thus, one can more precisely define what is meant by RDM. Two definitions from the relevant literature are worth mentioning, which together span the spectrum of meaning of the term:

> Research data management refers to all measures ensuring digital research data is usable. However, the requirements for doing so vary depending on the different purposes for which research data is to be used. Four types of purpose can be distinguished:
>
> - Use as a working copy for research;
> - the subsequent use of research data for later research;
> - storage as documentation of sound academic practice; and
> - storage to comply with legal or other non-research requirements (Enke & Ludwig 2013, 13).

> Research data management concerns the organization of data, from its entry to the research cycle through to the dissemination and archiving of valuable results. It aims to ensure reliable verification of results, and permits new and innovative research built on existing information (Whyte & Tedds 2011).

While the first definition particularly identifies the usability of data as the goal of RDM, the essential aspect of the data lifecycle image used in the second definition is the emphasis on the integration of the different phases of the research process. The handling of research data in later phases of a project depends on setting the agenda in the earlier phases. For example, if one works with data to which third parties hold rights and wants to publish this data, one must clarify the necessary rights already during the process of data collection. Assigning resources to the comprehensive documentation of the research data is imperative, if the data should be re-used by future projects.

2. FAIR Data Principles and the Aims of Research Data Management

The previous section, following Enke and Ludwig, argued that RDM is concerned with keeping research data usable. However, what does usability mean in detail? What characteristics must research data have to be usable? Wilkinson et al. (2016) attempt to answer these questions by providing the FAIR Data Principles.

FAIR stands for the four elements: *Findable, Accessible, Interoperable and Reusable.* These concepts are detailed out as follows:

> Findable
> F1. (Meta)data are assigned a globally unique and persistent identifier.
> F2. Data are described with rich metadata (see R1 below).
> F3. Metadata clearly and explicitly contain the identifier of the data they describe.
> F4. (Meta)data are registered or indexed in a searchable resource.
>
> Accessible
> A1. (Meta)data are retrievable by their identifier using a standardized communications protocol.
> A1.1 The protocol is open, free and universally implementable.
> A1.2 The protocol allows for authentication and authorization procedure, where necessary.
> A2. Metadata are accessible, even if the data are no longer available.
>
> Interoperable
> I1. (Meta)data use a formal, accessible, shared and widely applicable language for knowledge representation.
> I2. (Meta)data contain vocabularies that follow FAIR principles.
> I3. (Meta)data contain qualified references to other (meta)data.
>
> Reusable
> R1. (Meta)data are richly described with a plurality of accurate and relevant attributes.
> R1.1. (Meta)data are released with a clear and accessible data usage license.
> R1.2. (Meta)data are associated with detailed provenance.
> R1.3. (Meta)data meet domain-relevant community standards.[2]

Without discussing the various aspects of the FAIR principles in detail, the following key characteristics can be identified. *Findability* is based on comprehensive documentation and description using metadata and the use of persistent digital identifiers

2 See https://www.go-fair.org/fair-principles (Accessed: 19 June 2024).

(e.g., *Digital Object Identifier* (DOI)), which form the basis for stable referencing in publications, reference systems, and search engines. *Accessibility* is about the attempt to make data available as openly as possible and as protected as necessary. Ideally, research data should be published as Open Research Data, under the most permissive license possible. Should doing so not be possible, however, the data can be made available to authorized users using suitable authentication methods. If this is also not possible, at least the descriptive metadata should be publicly available. *Interoperability,* in turn, is essentially based on standards for data and metadata. The use of standardized data structures, non-proprietary data formats, or standardized vocabularies are the basis for the easy usability of the data. If data is available in machine-readable form, its reuse can be automated as well as the possible integration of the data with other data sets. *Reusability* is based on the rich description and documentation of the content, including provenance information, as well as on legal stipulations as to how the data can be reused. This is accomplished by using suitable open content licenses (e.g., Creative Commons licenses).

Appropriate research infrastructures can support researchers in providing their data in accordance with the FAIR principles. While the description and documentation of the data must be carried out by the researchers themselves; the data will only be findable (in the sense of the FAIR principles) if the data repositories used to provide the data have suitable functionalities, e.g., support suitable metadata standards or offer persistent identifiers.

3. Frameworks and Guidelines for Data Management

In addition to the requirements for RDM arising from the research context, as well as those arising from the respective research practices, underlying conditions and guidelines must be maintained. Researchers should consider these underlying conditions when planning a project. Third-party funding bodies now universally formulate requirements for the RDM of supported projects. These bodies do so either in the form of central guidelines, such as those issued by the DFG and the EU, and/or in the form of specific prerequisites within the individual program lines, as is the case with the BMBF.[3]

However, not only third-party funders but also universities and other research institutions formulate rules for handling research data.[4] In some cases, this is done in

3 See https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/forschungsdaten/index.html; https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/for schungsdaten_checkliste_de.pdf; https://www.openaire.eu/rdm-in-horizon-europe-proposals; and https://forschungsdaten.info/themen/informieren-und-planen/foerderrichtlinien. All addresses were accessed on 19 June 2024.

4 Cf. https://www.forschungsdaten.org/index.php/Data_Policies (Accessed: 19 June 2024).

dedicated data policies or in codes of conduct on data protection in order to safeguard good academic practice. In addition, the scientific communities themselves – often by their respective learned societies – as well as publishers and academic journals formulate frameworks for handling research data.[5]

4. Data Management Plans

A data management plan can serve as starting point and basis for RDM. A data management plan is a document that sets out which data is collected or used and how this data is to be handled in the project as well as afterwards. Ideally, the data management plans should be regularly consulted and updated as needed throughout the project.

The purpose of a data management plan is to make well-founded decisions on how to handle the research data of one's own project, to identify risks and challenges at an early stage, to establish uniform procedures and standards for the joint use of data in cooperative projects, and thereby ensure the sustainability of the data, saving time and effort by designing efficient processes. William Michener has formulated the central topics that a data management plan should address in the form of a checklist with ten aspects to consider (Michener 2015)[6]:

1. Determine the research sponsor requirements.
2. Identify the data to be collected.
3. Define how the data will be organized.
4. Explain how the data will be documented.
5. Describe how data quality will be assured.
6. Present a sound data storage and preservation strategy.
7. Define the project's data policies.
8. Describe how the data will be disseminated.
9. Assign roles and responsibilities.
10. Prepare a realistic budget.

There are a variety of templates, checklists, and web-based tools for creating a data management plan for a project (Dierkes 2021, 310). Tools such as the RDMO and DMPonline services are particularly helpful, as they use comprehensive question-

5 Cf. https://forschungsdaten.info/themen/ethik-und-gute-wissenschaftliche-praxis/leitlinien-und-policies (Accessed: 19 June 2024).

6 See also https://forschungsdaten.info/themen/informieren-und-planen/datenmanagementplan and the Science Europe guide: https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf. Both addresses were accessed on 19 June 2024. In addition to the list below, they also address issues such as data reuse, the legal and ethical framework and data sharing.

naires to draw attention to all potentially relevant aspects to be considered in data management planning.

5. Data Management in the Project (hot data)

Key aspects of data management concern the planned handling of research data over the course of a project. In many respects, these aspects are specific to the subject, method, and data specific and thus cannot be adequately explained in an introductory overview. However, they are addressed in other chapters of this handbook, in which concrete case studies and specific methods are discussed. This subsection focuses on generic aspects relevant to any research project.

5.1 Data Collection

The way in which data is collected for a research project depends on the project. In humanities projects, the data is often not collected by the researchers themselves, but provided by third parties, e.g., as publicly available digital copies from libraries and archives. This information is important not only because of the required attribution information or possible requirements for use and transfer of the data of those providing the data, but also because questions regarding the preservation and publication of the data may be answered differently. If the data on which a project is based has been published by a trustworthy provider, such as a library, then this institution will also ensure the long-term preservation of the data. Thus, the project does not need to find a solution for these aspects and can focus primarily on dealing with working copies, which are only needed during the project. Of course, when using third-party data, it may also be the case that the project generates its own research data derived from original data sets, e.g., annotations, tabular evaluations, or statistical analyses; the long-term preservation and publication of this data is thus subject of the project's own research data management.

Regardless of how data is collected, the documentation required for the long-term usability of the data should be carried out as early as possible, ideally using relevant metadata standards. The documentation can take various forms, which can also be combined, for example in Read Me files, structured metadata databases, a wiki, or directly in the data management plan.[7]

7 See https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation (Accessed: 19 June 2024).

5.2 Storage, Backup, and Deletion

For every research project, the question arises as to where and how the project-relevant research data is stored, such as on the hard disk of a local PC, on an institutional server, a university's central storage, or commercial provider. The use of central services is particularly recommended for large amounts of data. Every research project requires a suitable backup strategy. The so-called 3-2-1 rule serves as a rough guide here: three copies of data on two different storage media, one of which is at an external location (see Krogh 2009, Chapter 6).

If the data is stored on a personal computer, the owners are required to set up their own backup. If the institutional servers or other service providers are used, the provision of storage and backup solution falls within the scope of these providers. In principle, these solutions should be prefered. The latter solutions should therefore be given preference in principle. For small data volumes, the sync-and-share services now offered relatively widely by research institutions offer a low threshold solution.

In addition, research data management may already include selection processes regarding the deletion of files at this stage. Raw data on which the research is based should never be deleted during the project. Processed versions of the data can be deleted if they are no longer required. However, each individual step should always be documented to enable retracing the corresponding processing steps on the basis of the raw data if necessary. Processed versions of the data should be retained if they are essential building blocks of a publication. When the research project reaches its end, additional regulations may need to be made (e.g., regarding the deletion period for certain data).

5.3 Data Sharing

When choosing a suitable storage system, in addition to data security, backup, and costs, it is relevant whether the data in the research project is to be shared with other people and in which form this should take place. Will the data be shared within a research group or a project network? Does this require parallel access to and/or simultaneous work in files? Should cooperation partners have access to some or all the data on an ad hoc or permanent basis? Here, it is important to ensure that the selected services ideally already provide suitable rights management that can be adjusted as finely grained as possible and in which the corresponding access and editing permissions can be managed.

5.4 Data Organization

Like the aspects discussed above, data organization depends heavily on the respective research project and the relevant data material: image data must be handled differently than text corpora or data in tabular form; researchers working alone have different requirements than cooperative research projects in which a shared database is used; and large volumes of data must be handled differently than small volumes. Nevertheless, there are some generic recommendations for data organization.

Conventions on folder structure and file naming form the basis for a sensible form of data organization. The handbook *The Turing Way Community* (2022, chapter "Research Data Management") provides a straightforward introduction. Research data is usually stored in a folder structure. The naming conventions for folders should follow a systematic approach, such as chronological sorting, sorting according to the survey methods used, or assignment to individual sub-projects. Within the folders, the files should be named in a systematic way like using the creation date in the form YYYYMMDD as their respective names. Data types, data collection method, the name of the researcher or initials (especially in cooperative projects), or version numbers might be additional elements of the naming scheme.

The guiding principle should be that the file names should provide context for the respective file to distinguish it from similar files and from other versions of the same file. In addition, one might also consider suitable software tools to support data organization. Text-based data can be efficiently managed and versioned in a Git system. File naming tools can help name large quantities of files according to a uniform scheme.

5.5 Choice of Data and File Formats

Regarding the preparation of data in line with the FAIR principles, attention should also be given early on to the selection of suitable data and file formats. It is particularly important to differentiate between proprietary and non-proprietary (Open Source) formats as well as between binary and text formats.

When possible, non-proprietary formats should be chosen. However, a distinction must be made between binary and text formats. Binary formats are generally less memory-intensive, and many software products process and deliver binary formats. Nevertheless, with respect to the long-term preservation of data, one can (also) store their data in a textual format (if a corresponding migration is possible). Different formats exist for different means and for different purposes, namely, active work with the data in the project and later archiving and preservation. Wether data migration of the format is possible without loss of functionality and into any repository or archival software should be evaluated early on in the project phase. If conversion is only possible with a loss of information, it is important to consider whether this loss affects significant properties or not. For example, a table in Excel format may contain formatting such as bold column headings that are lost when converted to CSV format. However, if the formatting is not a significant feature for archiving, then a corresponding loss of information due to the migration can be acceptable. The Swiss Coordination Office for the Permanent Archiving of Electronic Records (KOST) provides a helpful overview of more than 50 common data formats and their suitability for long-term preservation.[8]

6. Archiving and Publication (cold data)

6.1 Repositories and Data Publication

By the end of a project, research data should be handed over to a trustworthy research data repository or data center, which takes over the sustainable long-term provision of the research data.

The following questions serve as a guide when selecting a suitable publication location for research data: Are there any special protection requirements for making the data available? Should the data be kept inaccessible to the public and only available upon request and under certain conditions? In this case, a service or repository must be found that can guarantee this form of *closed access.* If not, the next step should be to check whether there are suitable specialist repositories. These are usually the most suitable publication venues for research data due to their specialization in certain disciplines and/or data types. Data publications are visible in the subject-specific context in a joint collection with other relevant data from the subject. Subject repositories also support established metadata standards in the subject, offer specific search or visualization functionalities, and the operators of a subject repository may also be able to guarantee comprehensive curation and review of the data on the basis of relevant subject expertise.

Particularly if no suitable subject-specific service is available, institutional repositories or other generic, i.e., non-subject-specific services can be used. These do not offer the subject-specific functionalities described above but are nevertheless trustworthy publication venues that also comply with the FAIR principles. In particular, the strong connection to local research data support units and the benefit of direct on-site support might be strong arguments in favor of using institutional structures.

Unless there are specific reasons to the contrary, the data should be published as open research data, whereby open content licenses should be used wherever possible to ensure the broadest possible and lowest-threshold reuse of the data. Similar to Open Access publications, Creative Commons licenses have been established in

8 See https://kost-ceco.ch/cms/kad_intro_de.html (Accessed: 19 June 2024).

the field of research data as the *de facto* standard. This standard is questioned by researchers from big data fields, such as automated text analysis, who argue that research data should be placed in the public domain as far as possible, for example, by using a CC0 waiver (Brettschneider et al. 2021).

The international directory of research data repositories, *re3data*, lists a total of 22 subject repositories from the field of theology in May 2023.[9] However, depending on the field of research, other intellectual repositories in the humanities may also be suitable publication venues, e.g., the existing offerings and those currently under development of the humanities NFDI consortia NFDI4Culture, NFDI4Memory, NFDI4Objects, and Text+,[10] or the repositories of the CLARIAH network.[11] The data centers working group in the DHd – *Association for Digital Humanities in the German Speaking Areas* – also offers assistance in the search for suitable platforms.[12] Individual websites or web-based databases, which are a frequent result of humanities projects as "presentation layers" of research data, are a special case. The choice of these individual presentation formats may often make sense due to the heterogeneity of research questions in the humanities; however, the problem of sustainability arises immediately: how and by whom should such data products be able to be operated in the long term beyond the duration of the project? Sustainability is only possible if an infrastructure partner is involved at an early stage and should include the fallback option to turn off the web presentation while keeping the raw data available through a repository.

6.2 Long-Term Archiving

The problem of long-term digital preservation is particularly pressing for relevant research data in the humanities that remains relevant over very long periods of time. Digital long-term archiving has three aspects (cf. Liegmann et al. 2010):

- Bitstream Preservation
- Preservation of Functionality
- Preservation of Usability

Digital preservation cannot be ensured by individual researchers or research projects; rather, it requires technically and organizationally elaborate, sustainable infrastructures carrying out this overarching task. Nevertheless, researchers can con-

9 See https://www.re3data.org/search?subjects[]=107%20Theology (Accessed: 19 June 2024).

10 See https://nfdi4culture.de, https://4memory.de, https://www.nfdi4objects.net and https://www.text-plus.org. All addresses were accessed on 19 June 2024.

11 See https://www.clariah.de/publizieren-archivieren (Accessed: 19 June 2024).

12 See https://dhd-ag-datenzentren.github.io (Accessed: 19 June 2024).

tribute directly to the preservability of their generated research data, e.g., by using open, non-proprietary data formats or converting their data into such formats where possible. Doing so enables the maintenance of the functionality of the data in the long term. These formats are more likely to be supported in the long run and operators of preservation services are more likely to be able to transfer the data to new formats if the existing formats are no longer supported. Through adequate documentation and the provision of rich metadata, researchers also help to ensure that the data remains not only functional, but also usable, since only in this way can the data be understood, adequately interpreted, and contextualized.

7. Conclusion

Research data management is, without a doubt, a crucial element of the research process. The Digital Humanities (like other fields of research) can only fully develop the potential with high-quality research data organized according to the FAIR principles. This crucial role is summarized by Alma Gold as follows:

> Data is the currency of science, even if publications are still the currency of tenure. To be able to exchange data, communicate it, mine it, reuse it, and review it is essential to scientific productivity, collaboration, and to discovery itself" (Gold 2007).

RDM is essentially an activity of the researchers themselves, thus an intrinsic part of the research process (cf. also Lemaire 2018, 245). However, there are broad and diverse range of services that support researchers by providing advice and the necessary infrastructure and tools. Such services include the NFDI consortia, the CLARIAH network, and the DHd, but also institutional service centers for RDM remain key contact points. RDM that is consistently designed to implement the FAIR Data Principles improves the quality of research and its results and is essential for the seamless execution of a research project as well as for follow-up research based on the current project. Plainly stated, there is no digital research without RDM. You might manage your data efficiently or inefficiently, more FAIR, or less FAIR, but there never is no RDM.

References

Brettschneider, P., Axtmann, A., Böker, E., & Suchodoletz, D. v. (2021). Offene Lizenzen für Forschungsdaten, *o-bib. Das offene Bibliotheksjournal,* 8(3), 1–22. DOI: https://doi.org/10.5282/O-BIB/5749 (Accessed: 19 June 2024).

Dierkes, J. (2021). Planung, Beschreibung und Dokumentation von Forschungsdaten. In M. Putnings, H. Neuroth, & J. Neumann (Eds.), *Praxishandbuch Forschungsdatenmanagement* (pp. 303–326). Berlin/Boston: De Gruyter Saur. DOI: https://doi.org/10.1515/9783110657807-018 (Accessed: 19 June 2024).

Enke, H., & Ludwig, J. (Eds.). (2013). *Leitfaden zum Forschungsdaten-Management.* Boizenburg: Verlag Werner Hülsbusch. URL: https://www.forschungsdaten.org/images/b/bo/Leitfaden_Data-Management-WissGrid.pdf (Accessed: 19 June 2024).

Geiger, J. D. (2023). Daten/Forschungsdaten. In AG Digital Humanities Theorie des Verbandes Digital Humanities im deutschsprachigen Raum e. V. (Ed.), *Begriffe der Digital Humanities. Ein diskursives Glossar.* Wolfenbüttel: Herzog August Bibliothek [= *Zeitschrift für digitale Geisteswissenschaften. Working Papers,* 2]. DOI: https://doi.org/10.17175/WP_2023_003 (Accessed: 19 June 2024).

Gold, A. (2007). Cyberinfrastructure, Data, and Libraries, 1, *D-Lib Magazine,* 23(1/2), no pag. DOI: https://doi.org/10.1045/september20september-gold-pt1 (Accessed: 19 June 2024).

Kailus, A. (2023). Handreichung für ein FAIRes Management kulturwissenschaftlicher Forschungsdaten. V. 1.0.3. URL: https://nfdi4culture.de/go/E3625 (Accessed: 19 June 2024).

Krogh, P. (2009). *The DAM Book. Digital Asset Managament for Photographers.* 2. ed. Sebastopol: O'Reilly Media.

Lemaire, M. (2018). Vereinbarkeit von Forschungsprozess und Datenmanagement in den Geisteswissenschaften. Forschungsdatenmanagement nüchtern betrachtet, *o-bib. Das offene Bibliotheksjournal,* 5(4), 237–247. DOI: https://doi.org/10.5282/O-BIB/2018H4S237-247 (Accessed: 19 June 2024).

Liegmann, H., & Neuroth, H. (2010). Einleitung. In H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, & K. Huth. (Eds.), *nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung.* Version 2(3) (pp. 1–10). Boizenburg: Verlag Werner Hülsbusch. URL: https://nbn-resolving.de/urn:nbn:de:0008-2010071949 (Accessed: 19 June 2024).

Michener, W. K. (2015). Ten Simple Rules for Creating a Good Data Management Plan, *PLOS Computational Biology,* 11(10), 1–9. DOI: https://doi.org/10.1371/journal.pcbi.1004525 (Accessed: 19 June 2024).

Pempe, W. (2012). Geisteswissenschaften. In N. Heike, S. Strathmann, A. Oßwald, R. Scheffel, J. Klump, & J. Ludwig (Eds.), *Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme* (pp. 137–160). Boizenburg: Verlag Werner Hülsbusch.

The Turing Way Community. (2022). The Turing Way. A handbook for reproducible, ethical and collaborative research. Online: Zenodo. DOI: https://doi.org/10.5281/ZENODO.7625728 (Accessed: 19 June 2024).

Whyte, A., & Tedds, J. (2011). Making the Case for Research Data Management. In *DCC Briefing Papers.* Edinburgh: Digital Curation Centre. URL: https://www.dcc.ac.uk/guidance/briefing-papers/making-case-rdm (Accessed: 19 June 2024).

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, Ph.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, Ch.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., C't Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, Ph., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, Th., 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, *Scientific Data*, 3(1), 1–9. DOI: https://doi.org/10.1038/sdata.2016.18 (Accessed: 19 June 2024).