Thomas Seyffertitz

Research Data Repositories and What to Consider When Choosing One for Deposit

Handbuch Repositorienmanagement, Hg. v. Blumesberger et al., 2024, S. 21–45
https://doi.org/10.25364/97839033742323

Thomas Seyffertitz, Wirtschaftsuniversität Wien, Universitätsbibliothek, thomas.seyffertitz@wu.ac.at |
ORCID iD: 0000-0002-7444-6780

Abstract

The main purpose of the present contribution is to provide practical guidance with regard to selecting a suitable data repository for research data. It might be helpful for researchers, librarians, and research support staff. Choosing a research data repository (RDR) that is appropriate for research data can be challenging and may be influenced by various factors such as the specific character of the data, requirements imposed on the researchers by third parties, e.g. a funding agency or publisher. Therefore, different types of repositories are briefly characterized in recognition of the fact that scientific disciplines or research communities have different requirements for their data management – sometimes due to the characteristics of their data, and to some extent due to the specific academic culture that has evolved. Further, directories of data repositories are presented that may be helpful in finding an appropriate place for research data storage and/or data publication. Next, data policy frameworks of some scientific journal publishers are summarized pointing to a broad spectrum of potential data sharing requirements when submitting to a journal. The final section discusses some important issues and questions to be considered in the repository selection process.

Keywords: Research data; research data repository; data archive; data policy; scholarly journal; journal publisher

Zusammenfassung

Repositorien für Forschungsdaten und was bei der Auswahl eines Repositoriums zu beachten ist

Das vorliegende Kapitel soll Forscher:innen, Bibliothekar:innen und Mitarbeitenden von Forschungseinrichtungen als praktische Orientierungshilfe zur Auswahl eines Datenrepositoriums dienen. Die Wahl oder Empfehlung eines für Forschungsdaten geeigneten Repositoriums (RDR) kann eine Herausforderung sein und von verschiedenen Faktoren beeinflusst werden, wie z. B. dem spezifischen Charakter der Daten, oder den Anforderungen, die den Forscher:innen von Dritten, z. B. einer Fördereinrichtung oder einem Verlag, auferlegt werden. Zuerst werden verschiedene Arten von Repositorien kurz charakterisiert, um der Tatsache Rechnung zu tragen, dass wissenschaftliche Disziplinen oder Forschungsgemeinschaften unterschiedliche Anforderungen an ihr Datenmanagement stellen – manchmal aufgrund der Merkmale ihrer Daten und in gewissem Maße auch aufgrund der spezifischen akademischen Kultur, die sich entwickelt hat. Außerdem werden einige Verzeichnisse von Datenrepositorien vorgestellt, die bei der Suche nach einem geeigneten Ort für die Speicherung und/oder Veröffentlichung von Forschungsdaten

hilfreich sein können. Anschließend werden am Beispiel einiger großer wissenschaftlicher Zeitschriftenverlage Rahmen-Datenrichtlinien zusammenfassend vorgestellt. Die abschließende Zusammenfassung liefert Überlegungen, die zur Auswahl eines geeigneten Datenrepositoriums von Bedeutung sind, und formuliert Fragen, die von den jeweiligen Datenautor:innen zuvor beantwortet werden sollten.

Schlagwörter: Forschungsdaten; Datenrepositorium; Daten teilen; Datenpolitik; Forschungspublikation; Zeitschriftenverlag

1. Introduction

In recent years, depositing datasets associated with published research has become more common. Amongst other factors, such as proper research data management as an element of Good Scientific Practice (cf. DFG, ÖAWI and others), this is due to an increase in journals requiring data sharing[1]. It has been evidenced that journal-based data archiving policies could be very effective in ensuring that research data are available to the scientific community to enable reuse and reproducibility of results, especially when journals require a data accessibility statement to be provided in the published paper[2]. Key contributions towards the increasing availability and accessibility of research data have been the OECD Principles and Guidelines for Access to Research Data from Public Funding[3] as well as the formulation of the FAIR data principles[4]. In the case of publicly funded research projects (e.g. EU H2020 projects, FWF-funded projects), the funding bodies are increasingly demanding that the associated research data are stored in a data repository and openly accessible as far as possible (ERC, H2020, FWF, etc.).

The first objective of this contribution is to present different types of research data repositories to the reader, being one of the ways to publish and share research, and second, to provide some guidance when selecting a repository for the data deposit. The contribution primarily aims at staff generally offering research services or advising researchers on data repositories, but it may also serve as a helpful source of information for researchers themselves.

1 Resnik, D. B. et al. (2019)
2 Vines, T. H. et al. (2013 and 2020)
3 OECD (2007)
4 Wilkinson, M. et al. (2016)

In the next section (section 2), by including definitions concerning data and data repositories, the different types of RDR are described in more detail. Additionally, examples of some RDR are given. Where available, links to the relevant websites or related content are provided. However, the development of the landscape of RDR is dynamic, and finding an appropriate data repository may be a daunting task. The emerging of directories of research data repositories can be helpful for finding a suitable place to deposit research data (section 3). Data (sharing) policies and recommendations on data deposits by publishers and research funders are another important aspect in the publishing process. Hence, in section 4, some examples from important publishers of scholarly journals are presented. Finally, section 5 summarizes and considers issues of choosing a proper repository for data deposit.

2. Types and characteristics of research data repositories (RDR)

Data repositories not only serve as a storage infrastructure but can also help to make a researcher's research data more discoverable, accessible, and hence leading to potential reuse and an increased citation of scientific work. Furthermore, they can enable the reproducibility of results, making research outcomes more transparent. There is also evidence for a statistically well-supported citation benefit from open data[5]. Research tradition or discipline requirements imposed by publishers and funders, as well as institutional or national policies, may influence if, how and where researchers archive or publish the data that underlie their published research in dedicated repositories. There are different types of RDR. Looking at the scope of research areas a repository covers, two core types of data repositories can be distinguished: discipline-specific RDRs and general-purpose (or generic) RDRs. Following the US Geological Survey[6], throughout this contribution, the terms "data repository" and "data archive"[7] will be used synonymously. The same will apply to the term "research data repository" (RDR) and "research data archive".

A professionally operated data repository should always be the first choice for depositing research datasets. There, the data published in a thematically appropriate ecosystem, such repositories are recognized in their respective discipline, and the

5 Piwowar, H. A. et al. (2013)

6 U.S. Geological Survey (2022)

7 From different perspectives, the two terms may represent different concepts. Further details will not be discussed as it is beyond the scope of this contribution. The USGS approach will be followed here. For more details, see the terms archive and repository in the online dictionary of the Society of American Archivists (2022) or the German RDM information platform forschungsdaten.info. https://www.forschungsdaten.info/praxis-kompakt/glossar

visibility of the data is high. A domain- or discipline-specific repository is likely a good choice for data that can be publicly shared[8]. If there is no suitable repository available, or the requirements for the datasets cannot be met, a generalist repository such as Zenodo[9] could be an alternative[10]. If these services are not an option, e.g. because the amount of data is too large, or external storage is not possible for legal reasons, institutional research data repositories are a good place for publication. For the institution (and its researchers), providing its own repository can have several advantages. By having control over the published datasets, it is possible to specify the quality of the metadata and to ensure that the data remain in-house. Additionally, the publication process can be designed and controlled independently of third parties. It also increases the visibility of the institution and improves the overview of the published datasets of an institution. A general discussion and a good overview of these topics can be found in some recently published books[11].

There are various options available for the storage of research data. They can be stored in data archives or data repositories. An important step is to find out whether your institution operates such a platform (an institutional data archive or repository) or collaborates with a data centre where it may be compulsory to store your data. If this is not the case, it is advisable, particularly with a view to allowing other researchers to reuse your data, to find out whether any subject-specific repositories exist for your particular discipline where your data can be stored. Directories of repositories are useful tools for finding a suitable repository, an example being re3data[12], which currently lists more than 2.600 repositories (see section 3 for more details). Repositories can make data accessible in different ways: open (no access barriers), with restricted access (i.e. only metadata are accessible; or there is an embargo on a particular dataset for a certain period of time), or closed (no access rights at all)[13].

2.1. Research data

Research data are generated by various methods – depending on the research question. Data used or created within a research project may be primary data generated in the course of source research, experiments, measurements, interviews,

8 Whyte, A. (2015)
9 https://about.zenodo.org/policies/
10 See for example the data repository guidance for the Journal Scientific Data https://www.nature.com/sdata/policies/repositories
11 Cox, A. M. et al. (2018) and Corti, L. et al. (2020)
12 http://www.re3data.org/
13 re3data.org (2021)

surveys or polls. If data have been collected or produced by others than the author(s) and are already available (e.g. census data), they are termed secondary data. Depending on the research design, datasets may consist of both. These research data usually form the basis for scientific publications. Beside the different perspectives and concepts of what data generally are[14], quite a lot of different definitions of research data can be found in the relevant literature[15]. The definition by the German Research Foundation[16], and that by the Office of Management and Budget at the US Whitehouse (2020)[17] exemplify the diversity of viewpoints on research data. The OECD[18] defines research data "as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. (...)". For reasons of simplicity, the term "research data" refers to (digital) data that, depending on the subject context[19], are the subject of a research process, and are generated during a research process or are its result.

2.2. Research data repository (RDR)

In general terms, a research data repository (or data archive) can be characterized as the technical infrastructure and final destination for research data that is intended to be stored for the long term. The structure and characteristics of scientific data – the content of an RDR – differ greatly depending on the research discipline. Legal regulations may also influence the decision if, where and how to archive research data. Another reason for the different requirements of data management is the diversity of research traditions and publication cultures that have developed over time in the numerous disciplines. In simple terms, a research data repository represents an institutionalized storage location for digital research data. If you look more closely, it is not that simple, and there are definitions that are more detailed. For example, a data archive can be defined as the organizational unit that responsibly takes on the task of data management within a defined frame of reference, and a data repository as the realization of a data storage facility in that data archive[20].

14 Kitchin, R. (2021)
15 Cox, A. M. and Verbaan, E. (2018), pp. 19.
16 DFG (2009), p. 2.
17 Office of Management and Budget [=OMB] (2020); see also https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A110/2cfr215-0.pdf and for further details on OMB's functions within the US Government, see https://www.whitehouse.gov/omb/
18 OECD (2007), p. 13.
19 Kindling, M.; Schirmbacher, P. (2013), p. 130.
20 Ludwig, J.; Enke, H. (2013), p. 47-48.

Types of research data repositories

Selecting a research data repository that is appropriate for the data in question can be quite challenging and sometimes depends on very specific requirements that may arise from the data itself, but also from requirements imposed on the researcher by third parties, such as publishers or funders (see section 4 below for data policy issues). Data repositories can be distinguished along different attributes: the disciplinary context of data, functional criteria, organizational embedding, how they are operated (commercial/non-commercial) and special purposes. Hence, a certain repository may fall into more than one of the below-mentioned categories. Nevertheless, the following types of data repositories can be distinguished[21]:

(i) Domain-specific research data repositories (dedicated to a certain discipline, or a certain research field, hence sometimes called subject-repository),
(ii) Generic data repositories usually open for all kinds of research data (for example Zenodo),
(iii) Institutional (data) repositories (e.g. operated by a university[22])
(iv) Software repositories (e.g. Github, specifically for depositing software or software projects)
(v) Commercially operated repositories, usually operated by a profit-oriented company (or part of it; e.g. Figshare)

It should be noted that repositories of the types (i), (ii), and (iv) can also be commercially oriented (i.e. they fall into category (v)). In general, the types of data repositories do not represent disjoint types or classes. In table 1, some typical characteristics of research data repositories are presented. Software repositories differ from the traditional research data archives for several reasons. They are rather generic in their nature, usually not focusing on specific scientific disciplines. Another characteristic is the availability of services for private or commercial projects as well as research projects. The initial or further development of a software is a process that may take months or years, and software may be updated over time by releasing new versions. In summary, this indicates a significant difference in use compared to traditional research data repositories, where the archived data have usually reached a final state. A recent short overview on software repositories provides some factors to be considered in the selection process[23]. It should be noted

21 See for example, Baker, K. S.; Duerr, R. E. (2017), pp. 139-144, for more details.
22 For example, GAMS, which is the Humanities' Asset Management System of the University of Graz and contains not only research data https://gams.uni-graz.at/archive/objects/context:gams/methods/sdef:Context/get?mode=about (Stigler J. H.; Steiner, E. (2018)).
23 Hong, N. (2020)

that the RDR landscape is dynamic in its nature: for example, a repository may be subject to changes in its business model or extend the scope of the data to be archived. Therefore, features listed in the table below may not exclusively apply to one type of repository. In the author's recent experience, researchers are sometimes looking for a quick fix for their data deposit, not attaching equal importance to the options listed below. Instead, they try to meet the minimum requirements set by funders or publishers.

Table 1. Data repository types[24]: characteristics, potential advantages and disadvantages. Features may not be available for all repositories listed in the example column, but rather indicate usual core features that can be expected for this class of repositories.

Repository type	**Description of features typically present/expected**		**Examples**
	Potential advantages	Potential problems or shortcomings	
Domain specific digital repositories/data archives (disciplinary repositories)	• meeting data quality standards • PID • long-term preservation • data catalogues for discovery • licensing arrangements • promotion of data • monitoring secondary use of data • management of • access to data • management of user requests on behalf of data owner • enhanced consulting services for data depositors • training offers	• may accept only specific data that are typical for the respective research domain • (possibly) costs involved in depositing data or curation services • data depositing process may be a little more complex, i.e. time and effort of the depositing process may be an entry barrier for potential submitters	UKDA (UK Data Archive)[26] AUSSDA (Austrian Social Science Data Archive)[27] DARIAH-DE (Digital Research Infrastructure for the Arts and Humanities)[28] CESSDA (Consortium of European Social Science Data Archives)[29] ICPSR (Inter-University Consortium for Political and Social Research)[30]

24 Parts of the table have been adopted from Haaker, M.; Corti, L. (2020), pp. 276-277.
26 https://www.data-archive.ac.uk/ (UK's largest digital collection of social sciences and population research data).
27 https://aussda.at/
28 https://de.dariah.eu/en/dariah-de-in-kurze or https://de.dariah.eu/en/home
29 https://datacatalogue.cessda.eu/ As a consortium CESSDA rather acts as a data-catalogue.
30 https://www.icpsr.umich.edu/web/pages/

	• data curation services • may provide self-deposit for partner institutions • additional tools and services along the research cycle[25] • usually highly acknowledged by funders' policies		
Special purpose repositories / data community driven repositories	• characteristics similar to those of domain specific repositories • homogeneous data and data ingest process[31]	• for smaller research communities, operating their own data repository may not be commensurate with the effort required	BMRB (Biological Magnetic Resonance Data Bank)[32] wwPDB[33] x-econ.org[34] (focus on experimental economics data)
Institutional digital repositories[35] (For more examples, use a query at re3data.org[36])	• Usually accepting all data collections created by their staff • Deposit usually at no cost to staff • Locality of data (may be important in case e.g. of highly sensitive data) • Visibility via institutional channels • Suitable for small datasets	• Probably lacking data management and diversity of metadata profiles for managing data from different disciplines • Limited data curation in smaller institutions due to limited staff resources or lack of expertise • Number of datasets sometimes	GAMS (Humanities Asset Management System)[37] IST Austria (Institute of Science and Technology – Austria) PHAIDRA[38] (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets)

25 See for example https://de.dariah.eu/web/guest/dienste-und-werkzeuge

31 This means usually the data that are to be deposited in such repositories are homogeneous in terms of their characteristics, contents and structure, which is advantageous to the ingest process.

32 https://bmrb.io/

33 https://www.wwpdb.org/

34 A repository on experimental economics, since 2018: https://x-econ.org/xecon/#!AGB.

35 Operated by a research institution and usually only available for faculty members of the entire institution or projects where at least one project member must be affiliated with the institution.

36 Example query at re3data.org using "European Union" AND "institutional" as filter terms: https://www.re3data.org/search?query=&types%5B%5D=institutional&countries%5B%5D=EEC.

37 GAMS is the Humanities' Asset Management System of the University of Graz and contains not only research data but also other digital objects https://gams.uni-graz.at/archive/objects/context:gams/methods/sdef:Context/get?mode=about; (see also Stigler, J. H.; Steiner, E. (2018), for more details on GAMS).

38 https://phaidra.univie.ac.at/

		rather small compared to efforts required to run this service	
Generic digital repository (ranging from general purpose[39] to cross-disciplinary data repositories)	• Accepting a wide range of data types • Suitable for cross-disciplinary data • Suitable for smaller datasets • Data owner usually controls publishing of and access to dataset	• No or less editorial control over quality of data or metadata • No or limited data curation services • Limited metadata profiles • Location of physical storage of data may be legally relevant[40]	Zenodo (non-profit) Figshare (commercial) Harvard dataverse (non-profit, based on the dataverse project[41]) Dryad[42] RADAR[43] (Research Data Repository)
Software repositories[44]	Exemplary characteristics[45]: • Available for private, commercial and research projects • Usually for open-source licensed[46] projects • Some allowing for data hosting (e.g. git-annex for very large data files) • Depersonalization of services • Different version control systems • Certain services are fee-based • Developer- vs. project-focused environments		Github (a more developer focused environment) Bitbucket Sourceforge CRAN[47] (Comprehensive R Archive Network)

A further distinction can be made between special purpose repositories and disciplinary repositories. While the latter have their emphasis on research data (and the related research papers) in a certain scientific discipline or sub-discipline as a whole, the former focus on a specific topic or the type of data collected in these repositories is very narrow and homogeneous in its structure. Especially in the STEM-fields[48] the emergence of such repositories can be facilitated by the formation and longer-term

39 Generalist repositories accept data regardless of data type, format, content, or disciplinary focus.
40 For example, a research institution does not allow data deposits if the physical cloud-storage is located in the US.
41 For more details on the project, see King, G. (2007)
42 https://datadryad.org/stash/our_mission
43 https://www.radar-service.eu/en; for details, see Potthoff, J. et al. (2014) and Kraft, A. et al. (2016) and https://radar.products.fiz-karlsruhe.de/en/radarabout/ueber-radar.
44 Hong (2020)
45 (The +/- descriptions do not apply for SW-repositories.)
46 See, for example, the GNU General Public Licenses (https://www.gnu.org/licenses/gpl-3.0.html).
47 https://cran.r-project.org/
48 Abbr. for Science, Technology, Engineering and Mathematics.

existence of so-called data communities that focus on lively data exchange or sharing[49]. In a blogpost[50] at Springer Nature, Matthews (2022) states: "Each community repository is built around a specific data type, and this data-specificity yields several advantages", this term refers to the more common terms subject-specific or domain-specific repository. The following two databanks serve as examples for special purpose archives: The Biological Magnetic Resonance Databank[51] is a special purpose repository for experimental and derived data gathered from nuclear magnetic resonance spectroscopic studies of biological molecules[52]. As a second example, the Protein Data Bank[53] archive (PDB) serves as the only repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies. The Worldwide PDB (wwPDB) organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global community[54]. At the same time, these two repositories serve a highly specialized data community.

3. Directories of RDR – or where to search for an appropriate repository

Directories of repositories are internet portals hosted by different organizations or bodies and provide information about available repositories as well as related information. Based on the respective functionalities, this information can be discovered by browsing and search options. The scope of searchable content may vary quite a lot. In addition, repository descriptions may differ in comprehensiveness and detail among the registries. In the following, some of the most prominent directories are provided:

re3data.org[55] is by far the most comprehensive registry of RDRs. Since 2021, it has been providing information about more than 2.600 data repositories from different

49 Cooper, D.; Springer, R. (2019)
50 Matthews, T. (2022)
51 BMRB (https://bmrb.io/) is a special purpose repository for experimental and derived data gathered from nuclear magnetic resonance (NMR) spectroscopic studies of biological molecules.
52 Ulrich, E. L. et al. (2007)
53 https://www.wwpdb.org
54 wwPDB consortium (2018)
55 http://re3data.org/; online since 2012

academic disciplines all over the world[56], offering a broad set of functions and features[57]. It uses an icon system for visually indicating important features of a repository[58]. See table 2 for exemplary entries in this registry.

OpenDOAR (since 2005)[59] is a quality-assured, global directory of open-access repositories. It is not limited to data repositories. Thousands of registered repositories can be searched and browsed, based on a range of features, such as location, software or type of material held.

FAIRsharing[60] provides curated information on data repositories, data and metadata standards, as well as data policies[61].

Datacite's Repository Finder[62] can help to find an appropriate repository to deposit research data. The tool is hosted by DataCite[63] and queries the re3data registry. It further provides two predefined queries via the re3data registry resulting in a list of RDR that meet the criteria of the Enabling FAIR Data Project and the FAIRsFAIR Project[64]. Recently, DataCite has further developed its search tool: the "Repository Search" feature replaces the previous tool and merges metadata from re3data and DataCite[65].

Core Trust Seal[66] provides a (searchable) list of data repositories that are certified by at least one of the following data service providers: WDS (World Data Systems)[67], DSA (Data Seal of Approval). DSA is a certification body for repositories and was merged into CoreTrustSeal (CTS) in 2018, together with the ICSU Word Data System (WDS).

The Federal Ministry for Education, Science and Research in Austria (abbr. BMBWF) provides a search interface on its research infrastructure website, which

56 Strecker, D.; Weisweiler, N. L. (2021)
57 Vierkant, P. et al. (2018) and Pampel, H. et al. (2013)
58 Pampel, H. et al. (2015); for a more detailed description of the icons' meanings, see the FAQs at https://www.re3data.org/faq.
59 https://v2.sherpa.ac.uk/opendoar/
60 https://fairsharing.org/
61 Sansone, S.-A. et al. (2019)
62 https://repositoryfinder.datacite.org/: It is a pilot project of the Enabling FAIR Data Project led by the American Geophysical Union (AGU) in partnership with DataCite and the Earth, space and environment sciences community.
63 DataCite is a global provider of DOIs for research data.
64 https://www.fairsfair.eu/
65 Vierkant, P. (2022)
66 https://www.coretrustseal.org/
67 https://www.worlddatasystem.org

can be searched for data repositories in Austria[68], with the option to display the search results as a map.

Table 2. Selected examples of RDR (taken from the re3data.org directory of RDR)

re3data record	About	Content focus
re3data.org: **Comprehensive R Archive Network (CRAN)**; editing status 2021-09-02; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R3J88J	CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R.	Statistical computing
re3data.org: **AUSSDA** Dataverse; editing status 2021-11-25; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R39G72	Data archive for the social science community in Austria, offers a variety of research support services, primarily data archiving and help with data re-use[69].	Social sciences
re3data.org: **DARIAH-DE** Repository; editing status 2020-11-26; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R30G8N	Digital long-term archive for human and cultural-scientific research data.	Digital humanities
re3data.org: **RADAR**; editing status 2021-03-18; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R3ZX96	RDR for archiving and publishing research data from completed scientific studies and projects, focusing on data from subjects that do not yet have their own discipline-specific infrastructures for research data management.	Cross-disciplinary
re3data.org: **DRYAD**; editing status 2021-09-03; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R34S33	All material is associated with a scholarly publication.	General purpose

68 https://forschungsinfrastruktur.bmbwf.gv.at/en
69 https://aussda.at/en/

For researchers intending either to deposit data or search for reusable data, access is an important feature. Concerning the different types of access, for example, the re3data directory groups repositories according to four different levels[70]:

open: There are no access barriers. Both the data and their respective metadata are accessible.

embargoed: External users cannot overcome access barriers until the data are released and openly accessible. This refers only to the data themselves.

restricted: External users can overcome access barriers. Metadata can be accessed, but not the related data set.

closed: External users cannot overcome access barriers. Restriction refers to both metadata and the data themselves.

Restrictions may include the requirement of fees, registration or institutional membership. After the identification of one or more potential RDR for deposit, it is recommended to look at the respective homepages for specific deposit conditions (policies or the general terms and conditions) and for detailed information on the depositing process. This is important in so far as a data deposit to a data archive with a highly professional data curating process (including a moderated deposit process, such as AUSSDA or ICSPR) may take more time and cost than simply uploading data files to, for example, Zenodo. On the other hand, these larger disciplinary data archives usually comply with data policies provided by funders and publishers. Furthermore, they are often approved by relevant repository certification organizations, hence guaranteeing high quality standards.

70 How open are repositories in re3data? https://coref.project.re3data.org/blog/how-open-are-repositories-in-re3data

4. Publishers' data (sharing) recommendations and policies

In recent years, in the scholarly publication landscape, journal publishers have started developing (data) policies around the sharing or publication of research data underlying the manuscripts they publish. Some publishers or journals also refer authors to different data repositories in their policies or guidelines, or recommend searching a directory of data repositories (see section 3 in this contribution) – such as re3data – to find a suitable data archive for the relevant research data[71]. As this may be important to the authors submitting their publication, it is worth looking at some journal data policies here. For example, Springer Nature developed a framework for the research data policies of all its journals[72]. The Data Policy Standardisation and Implementation Interest Group (IG) of the Research Data Alliance further developed this framework around existing scholarly publishers' research data policies of Springer Nature, Elsevier, Wiley, and PLOS[73]. An overview of the research data policies of universities and other research institutions is omitted here, as it would go beyond the scope of this contribution. Also, research data policies provided by public funders (e.g. Austrian Science Fund, ERC[74], NSF[75] or ESRC[76]) will not be discussed in detail. For an example, the reader is referred to a comparison on requirements of the Austrian Science Fund and the European Union's Horizon 2020 Programme[77]. The summary section provides a few noteworthy tips for the repository selection decision considering funding agency guidelines. Journal publishers' data policies are probably the most relevant ones for researchers in everyday scholarly life. Therefore, some key elements and characteristics of these policies are summarized below. However, the reader should bear in mind that this represents only a small section and does not reflect the entire scientific journal publisher universe. In table 3, four prominent scholarly publishers and their research data guidelines (or policies) are exemplified in a condensed manner.

71 See for example https://authorservices.taylorandfrancis.com/data-sharing/share-your-data/repositories/
72 Hrynaszkiewicz, I. et al. (2017)
73 Hrynaszkiewicz, I. et al. (2020)
74 European Research Council: Data Guidelines and Open Data Policy https://open-research-europe.ec.europa.eu/for-authors/data-guidelines
75 See the National Science Foundation Proposal & Award Policies & Procedures Guide: https://new.nsf.gov/policies/pappg/23-1
76 See the Economic and Social Research Council's Research Data Policy https://www.ukri.org/publications/esrc-research-data-policy/
77 Spichtinger, D.; Blumesberger, S. (2020)

Table 3. A condensed summary of the policy framework spectrum of four large publishers

Publisher	Weakest (all features encouraged)	.....	Strongest (all features required)
Wiley[78]	• DAS*	.....	• DAS • Peer review of data
Taylor & Francis[79]	• DAS	.....	• DAS • PID for data • Data citation
Elsevier[80]	• Data deposit in a relevant data repository • Citing this dataset in the article	....	• Data deposit • Data citation and linking (or a DAS) • Peer review of data prior to publication
Springer[81]	• Data sharing • Data citation • DAS	.....	• Data sharing • Evidence of data sharing • Peer review of data

Note: *DAS = Data Availability Statement/Data Access Statement; PID = Permanent Identifier (e.g. DOI)

There may be certain differences in the detailed wording and expressions, but their policy frameworks show some common core features:

They provide a general framework – with a spectrum ranging from encouraging recommendations to strictly mandatory data (sharing) policies.

They provide a set of data policies focusing on different groups of journals. The features within a certain type of policy range from a simply recommending to a highly encouraging nature to absolutely mandatory features.

It may also depend on the journal editors/editorial board what type of policy is implemented at journal level. Hence, the respective journal author guidelines should be read (see, for example Nature's editorial and publishing policies[82]).

78 John Wiley & Sons Inc. (2022); As of July 2023 the provision of a DAS has become mandatory for all original articles.
79 Taylor & Francis (2018)
80 Elsevier (2022a, 2022b)
81 Springer Nature (n.d.)
82 https://www.nature.com/palcomms/journal-policies/editorial-and-publishing-policies

The "Data Availability Statement" (DAS) is the feature most often provided as mandatory instrument. It states where data supporting the results reported in a published article can be found. It may contain links (e.g. a DOI) to publicly archived datasets.

Data journals[83] that specialize in publishing (research) datasets (so-called data papers[84]), are amongst those with the most rigid data policies. As datasets themselves are the main subject of publication, these journals require a peer review of the data. The data must be made available to editors and referees at the time of submission.

For more details on the scope and design of the policies summarized in table 3, see the respective pages at the publishers' websites. Two further examples of data sharing policies are those by SAGE[85] and PLOS[86]. Other journals, like those published by the American Economic Association (AEA), have a so-called data editor[87] who defines and monitors the journal's approach to data and reproducibility. Similarly, INFORMS's[88] Management Science Journal[89] has a data editor installed and adopted some of the data policy features from the AEA. In general, it is strongly recommended to take a close look at editorial guidelines, author guidelines and – if explicitly mentioned – a journal's data policy (sometimes termed data disclosure policy) before submitting the paper. In case of ambiguity, one should contact the journal editor or any data editor in advance.

5. Summary

Within research projects or multi-authored research articles that include or generate research data, certain questions are very important and should be addressed at the earliest possible point of time – at its best at the beginning of such a research endeavor: Which data policies may apply, whether and where which data shall be archived. As mentioned in the previous section, the publisher's or funder's guidelines may include specific requirements regarding the provision or archiving of research data (e.g. open data by default, certified repository, etc.). Even if there are no legal or contractual requirements regarding data archiving, there are a number

83 Candela, L. et al. (2015)
84 For a detailed description of this type of journal articles see for example Chavan, V.; Penev, L. (2011).
85 https://uk.sagepub.com/en-gb/eur/research-data-sharing-policies
86 https://journals.plos.org/plosone/s/data-availability
87 See https://aeadataeditor.github.io/ and American Economic Association (2019).
88 The Institute for Operations Research and the Management Sciences is an internationally recognized association for professionals in operations research, analytics, management science, economics, and many other related fields.
89 https://pubsonline.informs.org/page/mnsc/datapolicy

of issues that should always be considered when selecting a suitable data repository.

Only recently, in the United States, a set of desirable characteristics of online, public access data repositories have been formulated to help ensuring that research data are findable, accessible, interoperable, and reusable (i.e. FAIR-principles[90]) to the greatest extent possible, while integrating privacy, security, and other preventive measures.

Table 4. Desirable characteristics of repositories for managing and sharing data from federally funded or supported research[91]

Organizational infrastructure	Digital Object Management	Technological issues	Human Data related considerations[92]
• Free and easy access • Clear use guidance • Risk management • Retention policy • Long-term organizational sustainability	• Unique persistent identifiers • Metadata • Curation and Quality assurance • Broad and Measured reuse • Common formats • Provenance	• Authentication • Long-term technical sustainability • Security and Integrity	• Fidelity to consent • Security • Limited use compliant • Download control • Request review • Plan for breach • Accountability

Usually it is preferable to use a domain-specific or disciplinary repository, or alternatively an institutional data repository. If neither of these is available, a generalist repository may be an option (see section 2 above). There are some key issues to consider when depositing research data[93] and there are different options of where and how data can be archived, published and disseminated. The choice may depend on the project, the nature and characteristics of the dataset itself, the kind of access control needed for the data, the length of preservation desired, costs associated with publishing data and many other factors.

90 For details see https://www.go-fair.org/fair-principles/ or Wilkinson, M. et al. (2016).
91 For a detailed description of the features, see The National Science and Technology Council (2022), p. 4-6.
92 Very important to human data protection (see Wilkins 2021).
93 Haaker, M.; Corti, L. (2020), p. 301.

Table 5. Data-related characteristics[94]

Data format	• Hosting of common file formats (e.g. csv, xlsx, doc, pdf etc.) • Hosting of proprietary file formats (e.g. raw image files)
Data size	• Limits to size per file or to total dataset size • Small/medium datasets (e.g. spreadsheets): datasets may be uploaded by the researcher, or transferred through university network drives to a server or the cloud, and/or uploaded by a data archivist into a repository. • Medium-to-large datasets (i.e. requiring terabyte/petabyte drives): there is a larger weight toward data curation, adding robust metadata for access points and considering logical divisions of datasets/fields in consultation with researchers • Very large datasets may require consortial services or national data preservation and archiving infrastructure. At this level, the storage costs and cost for data curation may become an important issue.
Data licensing	• Types of licenses available (CC0, waiver; software license etc.) • Take care in advance whether the data are licensed (e.g. from a third-party, like a commercial data-provider)
Data attribution and citation tools	• Assignment of dataset DOIs, PIDs • References to related publications
User access controls	• Tiered access (e.g. administrator-level, collaborator-level, curator-level) • Journal-integrated, anonymous access (for peer review pre-publication) • Optional embargo to data release following publication
Data access tools	• Comprehensive data and metadata search tools • Data access via direct download • Data downloading via API • Built-in tools for reading proprietary file formats • Integrated data analysis tools

While there are many important issues to be considered when choosing a repository or archive for depositing data, some of these issues may be more relevant in the ongoing research project, whereas others may be important at the end of the project when research data are archived for the long term. Therefore, in the following table a list of some important questions is presented. These questions should be considered at the beginning of the project. However, some of them may not be answered in advance and have to be tackled again at a later stage in the project.

94 Uzwyshyn, R. (2016), p. 21.

Table 6. Important questions with regard to the choice of a research data repository

- Who operates the repository? A major publicly funded organization may be a safer choice than a commercial provider, particularly where ***long-term availability*** is concerned. Is the archive well recognized within the respective research field?
- What does the ***collection policy*** of the archive look like, i.e. what data are collected?
- What ***legal requirements*** (data protection, location of the repository, etc.) need to be taken into account? Which GDPR rules apply?
- Are the research data or parts thereof subject to ***special restrictions*** (e.g. requirement of de-identification, copyright protection)?
- What recommendations or ***data policies*** are provided by the journals where the research output is supposed to be submitted?
- What specifications regarding the provision of the research data are required by a ***funder's policy*** (e.g. is there an open data mandate; what type of repositories are recommended)?
- Is there a suitable choice of different ***licensing models*** defining how data can be reused?
- Does the repository ***enable collaborative work*** with the data? This may be of interest during the project, especially in projects with researchers from different institutions (e.g. access to data; versioning of data sets).
- Do ***charges*** apply for data storage? What does the funder's policy say about cost coverage (e.g. the SNSF does not cover costs for data deposits if the repository is commercial[95])?
- Do you need a ***trustworthy data infrastructure***; is the data archive certified (e.g. CoreTrustSeal)?
- Are the search functions useful, and is it easy to cite the data (e.g. for data reuse), for instance through the assignment of ***persistent identifiers***? What about versioning of data sets (especially during a project)?
- Does the repository provide appropriate ***metadata*** schemes for data description?

Without claiming to be exhaustive, the above questions offer some first clues for the selection of a data repository. There are quite a few tools, provided either by a special interest group in the respective research community, or by a university's research support office. For example, in 2018, the AGU Enabling FAIR Data project's Repository Guidance Targeted Adoption Group developed a detailed "decision tree" for researchers in the earth, space and environmental sciences[96], as a tool to determine an appropriate repository in which to deposit their data. The tree is applicable to most domains of research, funding scenarios and project stages (i.e., proposal vs. project completion), and considers many of the above-listed issues with which the researcher may need to comply. If neither a domain-specific repository nor an institutional repository is a solution or none of them is available, a

95 Swiss National Science Foundation (2022), p. 15.
96 Enabling FAIR Data Community et al. (2018); see also https://doi.org/10.5281/zenodo.1475430

"generalist repository comparison" chart[97] may assist researchers in finding a generalist repository. Such a comparison chart can be a first starting point, as it provides some basic feature information on, for example, size limits for data, supported metadata standards, versioning support, potential costs etc.

Finally, for those not having any potential repositories in mind yet, it is recommended to visit one of the most comprehensive sources for data repositories and archives: the re3data registry[98]. It has not only search options, but also allows filtering of and browsing for data repositories along several indicators, such as content type, country, subject and many others. A further important source is FAIRsharing[99], a community-driven portal that provides repository search and additionally features a searchable database on standards and policies.

If it is still unclear what repository or data archive fits best, it is highly recommended to contact the institution's research service support or the research data management office, respectively.

Bibliography

American Economic Association (2019): Data and Code Availability Policy. July 10, 2019. https://www.aeaweb.org/journals/policies/data-code (retrieved 17.04.2023)

Baker, Karen S.; Duerr, Ruth E. (2017): Data and a Diversity of Repositories. In: Lisa R. Johnston (ed.): Curating Research Data Vol. 2: A Handbook of Current Practice. Chicago, Illinois: Association of College and Research Libraries, pp. 139-144.

Candela, Leonardo; Castelli, Donatella; Manghi Paolo; Tani, Alice (2015): Data Journals. A survey. In: Journal of the Association for Information Science & Technology 66 (9), pp. 1747-1762. https://doi.org/10.1002/asi.23358

Chavan, Vishwas; Penev, Lyubomir (2011): The Data Paper. A Mechanism to Incentivize Data Publishing in Biodiversity Science. In: BMC Bioinformatics 12 (Suppl. 15), p. 2. https://doi.org/10.1186/1471-2105-12-S15-S2

Cooper, Danielle Miriam; Springer, Rebecca (2019): Data Communities. A New Model for Supporting STEM Data Sharing. In: Ithaka S+R Issue Brief (13.05.2019). https://doi.org/10.18665/sr.311396

Corti, Louise; Van den Eynden, Veerle; Bishop, Libby; Woollard, Matthew; Haaker, Maureen; Summers, Scott (2020): Managing and Sharing Research Data. A Guide to Good Practice. 2nd edition. Los Angeles: Sage.

Cox, Andrew M.; Verbaan, Eddy (2018): Exploring Research Data Management. London: Facet Publishing.

97 Stall, S. et al. (2020), see also https://doi.org/10.5281/zenodo.3946719
98 re3data.org (2023); see also https://www.re3data.org
99 Sansone, S.-A. et al. (2019); see also https://fairsharing.org/

Deutsche Forschungsgesellschaft [DFG] (2009): Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten. https://www.dfg.de/resource/blob/169298/51f011ec1a047637243ea95a994a49e6/ua-inf-empfehlungen-200901-data.pdf (retrieved 17.04.2023)

Elsevier (2022a): Research Data Guidelines. https://www.elsevier.com/authors/tools-and-resources/research-data/data-guidelines (retrieved 17.04.2023)

Elsevier (2022b): Policies. Research Data. https://www.elsevier.com/researcher/author/tools-and-resources/research-data/data-guidelines (retrieved 17.04.2023)

Enabling FAIR Data Community; Duerr, Ruth; Kinkade, Danie; Witt, Michael; Yarmey, Lynn (2018). Data Repository Selection Decision Tree for Researchers in the Earth, Space, and Environmental Sciences. https://doi.org/10.5281/zenodo.1475430

Haaker, Maureen; Corti, Louise (2020): Publishing and Citing Research Data. In: Corti, Louise; Van den Eynden, Veerle; Bishop, Libby; Woollard, Matthew; Haaker, Maureen; Summers, Scott (eds.): Managing and Sharing Research Data. A Guide to Good Practice. 2nd edition. Los Angeles: Sage, pp. 275-307.

Hong, Neil Chue (2020): Choosing a Repository for Your Software Project. The Software Sustainability Institute. University of Edinburgh. https://www.software.ac.uk/guide/choosing-repository-your-software-project (retrieved 17.04.2023)

Hrynaszkiewicz, Iain; Simons, Natasha; Hussain, Azhar; Grant, Rebecca; Goudie, Simon (2020): Developing a Research Data Policy Framework for All Journals and Publishers. In: Data Science Journal 19 (1), p. 5. https://doi.org/10.5334/dsj-2020-005

Hrynaszkiewicz, Iain; Birukou, Aliaksandr; Astell, Mathias; Swaminathan, Sowmya; Kenall, Amye; Khodiyar, Varsha (2017): Standardising and Harmonising Research Data Policy in Scholary Publishing. In: International Journal of Digital Curation 12 (1), pp. 65-71. https://doi.org/10.2218/ijdc.v12i1.531

ICSPR [Inter-university Consortium for Political and Social Research] (2020). Guide to Social Science Data Preparation and Archiving. 6th edition. https://www.icpsr.umich.edu/files/deposit/dataprep.pdf (retrieved 17.04.2023)

John Wiley & Sons Inc. (2022). Wiley's Data Sharing Policies. https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html (retrieved 26.09.2022)

King, Garry (2007): An Introduction to the Dataverse Network as an Infrastructure for Data Sharing. In: Sociological Methods & Research 36 (2), pp. 173-199. https://doi.org/10.1177/0049124107306660

Kitchin, Rob (2021): The Data Revolution. A Critical Analysis of Big Data, Open Data and Data Infrastructures. 2nd revised edition. London: SAGE Publications.

Kindling, Maxi; Schirmbacher, Peter (2013): „Die digitale Forschungswelt" als Gegenstand der Forschung. In: Information – Wissenschaft & Praxis 64 (2-3), S. 127-136. https://doi.org/10.1515/iwp-2013-0017

Kraft, Angelina; Razum, Matthias; Potthoff, Janet al. (2016): Archivierung und Publikation von Forschungsdaten. Die Rolle von digitalen Repositorien am Beispiel des RADAR-Projekts. In: Bibliotheksdienst 50 (7), S. 623-635. https://doi.org/10.1515/bd-2016-0077

Ludwig, Jens; Enke, Harry (Hg.) (2013): Leitfaden zum Forschungsdaten-Management – Handreichungen aus dem WissGrid-Projekt. Glückstadt, Deutschland: Verlag Werner Hülsbusch.

Marker, Hans Jørgen; Fink, Anne Sofie (2018): CESSDA – a History of Research Data Management for Social Science Data. In: Thestup, Jesper Boserup; Kruse, Fillip (eds.): Research Data Management – a European Perspective. Berlin, Boston: De Gruyter Saur, pp. 25-42. https://doi.org/10.1515/9783110365634-003

Matthews, Tristan (2022): Community Repositories. The Best Way to Share the Data Underlying Your Research. https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html (retrieved 30.08.2022)

Office of Management and Budget [OMB], Executive Office of the President (2020): Guidance for Grants and Agreements. 2 CFR parts 25, 170, 183, and 200. In: Federal Register 85 (157), p. 49551. https://www.federalregister.gov/d/2020-17468 (retrieved 17.04.2023)

OECD (2007). OECD Principles and Guidelines for Access to Research Data from Public Funding. Paris: OECD Publishing. https://doi.org/10.1787/9789264034020-en-fr

Pampel, Heinz; Vierkant, Paul; Scholze, Frank et al. (2015): The re3data.org Icon System Depicting All Possible Values for Each Icon. In: PLoS ONE. Figure. https://doi.org/10.1371/journal.pone.0078080.g002

Pampel, Heinz; Vierkant, Paul; Scholze Frank et al. (2013): Making Research Data Repositories Visible. The re3data.org Registry. In: PLoS ONE 8 (11) e78080. https://doi.org/10.1371/journal.pone.0078080

Piwowar, Heather A.; Vision, Todd J. (2013): Data Reuse and the Open Data Citation Advantage. In: PeerJ 1, e175. https://doi.org/10.7717/peerj.175

Potthoff, Jan; Wezel, Jos Van; Razum, Matthias; Walk, Marius (2014): Anforderungen eines nachhaltigen, disziplinübergreifenden Forschungsdaten-Repositoriums. In: Müller, Paul; Neumair, Bernhard; Reiser, Helmut; Rodosek, Gabi D. (Hg.): 7. DFN-Forum – Kommunikationstechnologien. Bonn: Gesellschaft für Informatik e.V., S. 44136.

re3data.org (2023). re3data.org – Registry of Research Data Repositories. https://doi.org/10.17616/R3D (retrieved 28.01.2023)

re3data.org. (2021). How Open Are Repositories in re3data? (Dec. 6th, 2021). https://coref.project.re3data.org/blog/how-open-are-repositories-in-re3data (retrieved 17.04.2023)

Resnik, David B.; Morales, Melissa; Landrum, Rachel et al. (2019): Effect of Impact Factor and Discipline on Journal Data Sharing Policies. In: Accountability in Research 26 (3), pp. 139-156. https://doi.org/10.1080/08989621.2019.1591277

Ruediger, Dylan (2022): Guest Post. The Outlook for Data Sharing in Light of the Nelson Memo. https://scholarlykitchen.sspnet.org/2022/09/06/guest-post-the-outlook-for-data-sharing-in-light-of-the-nelson-memo/ (retrieved 17.04.2023)

Sansone, Susanna-Assunta; McQuilton, Peter; Rocca-Serra, Phipippe et al. (2019): FAIRsharing as a Community Approach to Standards, Repositories and Policies. In: Nature Biotechnology 37 (4), pp. 358-367. https://doi.org/10.1038/s41587-019-0080-8

Society of American Archivists [SAA] (2022): Dictionary of Archives Terminology. SAA. Society of American Archivists. https://dictionary.archivists.org/index.html (retrieved 11.08.2022)

Spichtinger, Daniel; Blumesberger, Susanne (2020): Fair Data Management Requirements in a Comparative Perspective. Horizon 2020 and FWF policies. In: Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare 73 (2), S. 207-216. https://doi.org/10.31263/voebm.v73i2.3504

Springer Nature (n.d.). Research Data Policy. Legacy Data Policy Types. https://www.springernature.com/gp/authors/research-data-policy/research-data-policy-types (retrieved 06.09.23)

Stall, Shelley; Martone, Maryann E.; Chandramouliswaran, Ishwaret al. (2020): Generalist Repository Comparison Chart. https://doi.org/10.5281/zenodo.3946720

Stigler, Johannes H.; Steiner, Elisabeth (2018): GAMS – Eine Infrastruktur zur Langzeitarchivierung und Publikation geisteswissenschaftlicher Forschungsdaten. In: Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 71 (1), S. 207-216. https://doi.org/10.31263/voebm.v71i1.1992

Strecker, Dorothea; Weisweiler, Nina Leonie (2021): Wie wird das Registry of Research Data Repositories re3data verwendet und referenziert? – Eine kurze Literaturanalyse. Blog der DINI AGs FIS & EPUB. https://doi.org/10.57689/DINI-BLOG.20210527

Swiss National Science Foundation (2022). General Implementation Regulations for the Funding Regulations, Version: 1.1.2023 (76 p.). Bern: Swiss National Science Foundation. https://www.snf.ch/media/en/B0SWnPsrDCRTaiCx/snsf-general-implementation-regulations-for-the-funding-regulations-e.pdf (retrieved 17.04.2023)

Taylor & Francis (2018): Data Sharing Policies. https://authorservices.taylorandfrancis.com/wp-content/uploads/2019/04/Author-Services-Data-sharing-policies.pdf (retrieved 17.04.2023)

The Software Sustainability Institute. (2018): Software Deposit Guidance for Researchers. The Software Sustainability Institute, University of Edinburgh. https://softwaresaved.github.io/software-deposit-guidance/ (retrieved 30.01.2022)

Ulrich, Eldon, L.; Akutsu, Hideo; Doreleijers, Jurgen F. et al. (2008): BioMagResBank. In: Nucleic Acids Research 36 (Issue suppl., 1 January 2008), pp. D402–D408. https://doi.org/10.1093/nar/gkm957

U.S. Geological Survey [USGS] (2022): Archive vs. Repository. Is There a Difference? USGS (U.S. Department of the Interior). https://www.usgs.gov/data-management/archive-vs-repository-there-difference (retrieved 11.08.2022)

Uzwyshyn, Ray (2016): Research Data Repositories. The What, When, Why, and How. In: Computers in Libraries 36 (3), pp. 18-21. https://www.proquest.com/docview/1792217242?accountid=29104&sourcetype=Trade%20Journals (retrieved 17.04.2023)

Vierkant, Paul; Pampel, Heinz; Elger, Kirsten; Kindling, Maxi; Ulrich, Robert; Witt, Michael; Fenner, Martin (2018): Status Quo and Perspective of re3data. Paper presented at

the MERIL-2 Interoperability Workshop, Athens, Greece. https://doi.org/10.5281/zenodo.1297432

Vierkant, Paul (2022). How to find a repository for your research outputs. In: DataCite Blog. DataCite. https://doi.org/10.5438/mz0y-7j88

Vines, Timothy H.; Albert, Arianne (2020): The effect of a strong data archiving policy on journal submissions (Part II). In: The Scholarly Kitchen (Aug 26, 2020). https://scholarlykitchen.sspnet.org/2020/08/26/__trashed/ (retrieved 17.04.2023)

Vines, Timothy H; Andrew, Rose L. et al. (2013): Mandated data archiving greatly improves access to research data. In: The FASEB Journal 27 (4), 1304-1308. https://doi.org/10.1096/fj.12-218164

Wilkins, R. Bert (2021): Do You Know Me?: The Subtle Distinction Between "Anonymous" and "De-identified" Data in Clinical Research (White Paper, Western-Copernicus Group Institutional Review Board). https://www.wcgclinical.com/insights/do-you-know-me-the-subtle-distinction-between-anonymous-and-de-identified-data-in-clinical-research/ (retrieved 11.10.2023)

Wilkinson, Mark; Dumontier, Michel; Aalbersberg, IJsbrand et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. In: Scientific Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

Whyte, Angus (2015): Where to keep research data. DCC checklist for evaluating data repositories (v.1.1). Edinburgh: Digital Curation Centre. https://www.dcc.ac.uk/guidance/how-guides/where-keep-research-data (retrieved 17.04.2023)

wwPDB consortium (2018): Protein Data Bank: the single global archive for 3D macromolecular structure data. In: Nucleic Acids Research 47 (D1), D520-D528. https://doi.org/10.1093/nar/gky949

Thomas Seyffertitz ist seit 2013 an der Universitätsbibliothek der WU Wien beschäftigt. Er ist Fachreferent für Wirtschaftswissenschaften sowie Mathematik, Statistik und Finanzierung und zuständig für Forschungsdatenmanagement. Darüber hinaus ist er Mitglied der Kommission für Forschung an der WU Wien.