FAIR DATA: HISTORY AND PRESENT CONTEXT

Ana Carballo-Garcia[1], Juan-José Boté-Vericad[1]

[1] *Facultat d'Informació i Mitjans Audiovisuals. Universitat de Barcelona (Spain)*

Abstract

In this paper, we discuss FAIR Data, why it exists, and who it applies to. We further review the principles of FAIR data and how they are managed by research institutions. We also discuss the types of problems that researchers encounter, and what an information professional can do to assist them.

In 2016, the journal Scientific Data published the 'FAIR Guiding Principles for Scientific Data Management and Stewardship'. FAIR data are data that comply with the principles of findability, accessibility, interoperability, and reusability. The principles aim to guide data producers and publishers in the 21st-century digital environment, where data is the new gold. FAIR data tunes into the open science movement and responds to the digital revolution to maximize the added value offered by scholarly digital publications. In 2018, the European Union ratified and promoted the principles with the report 'Turning FAIR into reality', aiming to use the analytical power of machines on a large scale and ensure transparency and social utility, both of data and other digital objects produced and used for research.

Researchers generate large amounts of data, which are necessary to generate knowledge and innovation. The integration and subsequent reuse of data accompanying publications were left to the discretion of the data owner. The emergence of the FAIR principles arose because there were no guidelines or standards created for the proper integration of research data into the digital ecosystem. Therefore, the creation of the principles goes beyond collection, description, and archiving. Data management also contemplates long-term preservation, the generation of machine-processable metadata that facilitates discovery, evaluation, and reuse in further research.

At present, the vast majority of research centers subscribe to the principles. Furthermore, a Data Management Plan (DMPs) is required for the award of public funding, which must detail how the data will be managed, stored and preserved. But both centers and researchers face the arduous task of understanding the model, managing and implementing it. They must know data formats and standards. For a correct description and to facilitate data retrieval and interoperability, they must know about different types of metadata schemas. They must know about digital preservation and specific aspects of knowledge and information management. In addition, there are also ethical issues, intellectual property, and cultural differences. All these controversies translate into a huge extra workload for researchers, who only get a return in the form of citations.

For proper data management and compliance with FAIR principles, advanced knowledge of information management is needed. The creation of a repository, developing good practices that allow data management, promoting and facilitating open publishing, training in open science and data management; are subjects and disciplines that the profile of professional librarians or information managers have. Information professionals can play a key role in the proper management of research data and contribute to the achievement of the objectives described by the principles: making data findable, accessible, interoperable, and reusable.

Keywords: Data research management; FAIR principles; FAIR Data; FAIR implementation challenges; Literature review.

1 INTRODUCTION

The academic and research world, both public and private, generates large amounts of data in their research. For a long time, there was a debate on how to share and reuse data, because there were no guidelines or standards for proper correct integration of research data into the digital ecosystem.

Likewise, open science is a growing movement that facilitates access to the information generated by research, much of this research is financed with public money, and there is a general awareness that this investment must return to society itself. Thus, society increasingly can access scholarly articles in an open manner. Open access helps to improve research, which is why, like articles, data should also be accessible. Research data can be reused in other research, but it must be data that can be found, accessed, and used; to achieve these qualities, the data must have a series of appropriate features (Collins et al., 2018. p. 19).

One of these features is that the data must be findable, which means that the data must be described in a way that machines can retrieve the data being searched. Another feature is that the data must be accessed using protocols and standards.

In addition, technologies are advancing and facilitating the transmission of information, but at the same time, it is necessary to deal with the complexity that exists for machines to communicate and recognize formats. Therefore, interoperability is a feature that facilitates the correct integration of data from different sources in the same system.

The mentioned features, findable, accessible, interoperable, and reusable, are the basis of the principles FAIR Data, with the aim of using the analysis capacity of machines on a large scale and ensuring the transparency and social utility.

2 METHODOLOGY

The present study was realized by analyzing articles and informative material from relevant sources, related to the creation, emergence, and implementation of the FAIR Data guiding principles for the retrieval and use of data generated by science. The information collected has been filtered and synthesized in a way that allows us to know the most relevant points of each subtopic.

3 RESULTS

The following sections provide a historical and current overview of the context of the emergence of the principles and their adoption by stakeholders.This is followed by the implementation by the research team and ends with the role of the information professional in facilitating the management of research data.

3.1 The Origin of FAIR Principles

In 2014 in Liden, The Netherlands, a workshop was held under the name 'Jointly Designing a Data Fairport'. The workshop brought together both academic and private stakeholders in the fields of research infrastructure and policy, publishing, semantic web and life science research. The vision was to come together to support existing communities trying to find solutions to existing problems in scientific data retrieval and reuse (Data Fairport, 2014).

In the workshop they wanted to find answers to questions such as: if there were a suitable dataset, where could it be published and how would it be searched? With what search tools? Would it be necessary to apply filters? Is the necessary metadata for filtering integrated in the repositories? What format is the data in? Some of the obstacles to data management are a vast ecosystem of data, which is not centralized and is very diverse, making it more difficult to integrate, and compounding the problem of discovery and retrieval for both humans and machines (Wilkinson et al., 2016, p.2)

From this workshop awareness emerged that by defining a minimum set of community-agreed guiding principles and practices, it would be possible to optimally discover, access, integrate and reuse data, and to facilitate the citation of information produced by current science, which is highly data-centric. Finally, the meeting ended with a draft set of fundamental principles that were then specified in more detail. (Wilkinson et al., 2016, p.3).

Later, a FAIR working group adjusted and improved the draft. The working group was composed of several members of the FORCE11 community, which is a platform for discussion and common work to establish new paths and practices for the transformation of scholarly communication through technology and to respond to the problems that exist in maintaining a paper-based scholarly publishing system when academic journals are mostly electronic. (Martone, 2015, p.1)

In 2016, the journal Scientific Data published the 'FAIR Guiding Principles for Scientific Data Management and Stewardship'. The intention of the publication of these guiding principles is to act as a guide for those who want to improve the reuse of their data.

The FAIR Principles focus on improving the ability of machines to automatically find and use data, as well as supporting its reuse by people. (Wilkinson et al., 2016, p.3-4). The principles are as follows:

Findable - The data and the metadata are easy to find for humans and computers:

F1. (Meta)data are assigned a globally unique and persistent identifier

F2. Data are described with rich metadata (defined by R1 below)

F3. Metadata clearly and explicitly include the identifier of the data they describe

F4. (Meta)data are registered or indexed in a searchable resource

Accessible - Standard protocols are used:

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 an authentication and authorisation procedure, where necessary

A2. Metadata are accessible, even when the data are no longer available

Interoperable - Easy to combine the data with existing data:

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (Meta)data use vocabularies that follow FAIR principles

I3. (Meta)data include qualified references to other (meta)data

Reusable - The data can be used, including descriptions as clear licenses:

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

Figure 1. The FAIR Guiding Principles.

3.2 Implementation of the principles in research centers

In 2016, leaders at the G20 in Hangzhou summit supported the application of the FAIR principles in scientific research (Leaders of the G20, 2016, p.12). In 2017 a group of Australian organizations represented as The FAIR Steering Group issued a statement endorsing the principles (The FAIR Steering Group, 2017).

In 2017 too, Germany, the Netherlands and France created the International Support and Coordination Office (ISCO) to support the GO FAIR Initiative, with the aim of open the research data of scientific and academic institutions in all fields of research, as well as being a launch for the European Open Science Cloud (EOSC), project to provide an open platform for the exchange of data from all over Europe (GO FAIR, 2017).

Organizations like Research Data Alliance (RDA) and CODATA del Committee on Data of International Science Council (ISC), both international institutions working for the implementation of open science, in open research data, also endorsed the FAIR guiding principles by their communities. RDA enabled working groups to explore how to apply the FAIR principles and CODATA wrote a decanal program "Data for Planet: Making data work for cross-domain grand challenges”.

In 2018, the European Union ratifies and drives forward the principles with the report “Turning FAIR into Reality”, a report and action plan that brings together the study and analysis of what is needed to implement the FAIR principles and provides a set of recommendations and concrete actions for stakeholders (Collins et al., 2018, p.8). The *Ligue des Bibliothèques Européennes de Recherche* – Association of European Research Libraries (LIBER) was also involved in the adoption of the FAIR principles in the research data (LIBER, 2018).

As can be seen, most stakeholders strongly endorsed the FAIR principles, and many publicly funded projects demand Data Management Plans to ensure that research data are managed.

The RDA website contains a section, titled, “Recommendations & Outputs”, where they provide a list of organizations and entities that have adopted the RDA recommendations on data. While it is not explicitly about the adoption of FAIR data, many of its recommendations are designed to ensure that data can be found, accessed, interoperate between systems and can be reused. For example, recommendations on “Metadata Standards Directory” provides metadata standards for documenting research data, regardless of academic discipline (RDA, 2016). But this listing of organizations adopting the recommendations does not evaluate the adoption of the FAIR principles on the data created by the research centers. Since the appearance of the principles, there have also been initiatives, such as working groups or different projects, to develop indicators to assess if the principles are applied to research data.

The RDA in 2019 created the working group “FAIR Data Maturity Model”, with the purpose of developing a set of evaluation criteria to facilitate and standardize the evaluation of the data. Finally, the working group defined some indicators derived from the FAIR principles, with the aim of formulating measurable aspects of each principle. (Bahim et al., 2020, p.1). The same group developed a study to analyze current and existing approaches related to self-assessment tools FAIR (Bahim, Dekkers, & Wyns, 2019).

The FAIRassist.org website also compiles and describes the existing resources for the evaluation of digital objects with reference to the following principles FAIR (FAIRsharing Team & University of Oxford, 2019).

The existence of several initiatives and tools for assessment highlights the difficulty in assessing FAIR compliance in data, as appropriate practical solutions are missing or not fully developed. It appears that some of the problems in assessing whether the data meet the FAIR principles are that some of them are vaguely defined (Devaraju & Huber, 2021, p.1-2).

In 2020 Jacobsen et al. published a paper with the intention of harmonizing the implementation options of the FAIR principles, which includes a discussion and a non-technical interpretation of the principles, examples of solutions and discussions on the challenges to be considered.

Some studies have been able to perform assessments, for example, a study has been conducted on 80 databases that is representative of the data streams that are critical for low carbon energy transition, where the results obtained reveal the difficulty of translating FAIR guidance principles into domain specific applications, as current FAIR data practices in the energy domain are yet to be developed, although the community has initiated efforts but still the platforms and tools are not yet ready to be integrated into the workflows of research teams (Schwanitz et al., 2022).

The study “Cultivating FAIR principles for agri-food data” concludes that many agri-food communities still lack the necessary building blocks, such as shared vocabularies, sufficient quality datasets and shared data management practices (Top, Janssen, Boogaard, Knapen, & Şimşek-Şenel, 2022, p.11),

Therefore, it is still too early to have a global picture and to know if research centers have implemented FAIR principles in data management, because metrics and evaluation methods are still under discussion or development.

3.3 Researchers and challenges with data management

Researchers have different challenges involving research data management. Jahnke & Asher (2012) published a report with five significant key findings concerning these challenges:

- Not having formal training in data management practices,
- Some early career researchers think about the long-term preservation of their data.
- Metadata and documentation are of interest only if they help a researcher work.
- Absence of collaboration tools

- Researchers are not aware of the data services that the library may provide.

In addition, there are other challenges where researchers are also involved. First, to understand that data is part of the research workflow. To this matter, it should be convenient to have document guidelines to help researchers know what to do with their data. These guidelines should recognize that access to research data needs to be carefully managed to maintain confidentiality and security at the lowest possible cost (Hermans, 2019). Good data management facilitates the verification of research results, and other researchers can base their research on reusing data (Chigwada, Chiparausha, & Kasiroori, 2017).

Second, lack of skills and knowledge concerning data management plans. A solution with research data is data management plans (DMP). Nevertheless, researchers are still not aware of DMP in their institutions. There are rubrics (Deroo, Hermans, Lammertyn, & Merten, 2019) that help researchers deal with data management plans (DMP). These rubrics need to be according to H2020 research data requirements in the case of Europe. To this matter, the European Commission states, “As open as possible, as closed as necessary” for Open Science that in practice, means to produce a Data Management Plan for the data involving the restrictions of access to the data when necessary (Lahti Schulman, Piirainen, Riihikoski, & Juslén, 2019; Tammaro & Caselli, 2020). In a study held in India analyzing 47 central universities. In their findings, it was found that 20 libraries provided enough information about how to develop a research data management (RDM) plan for grant applications or how to deposit data to data repositories. Nevertheless, they argued that researchers, project investigators, and computer divisions should collaborate closer with librarians (Tripathi, Shukla & Sonkar, 2017). Whether or not the funder requires one, many stories about data loss or breaches can be avoided through proper data management planning (Rice, 2019).

At the same time, a report from the European University Association (Morais & Borrell-Damián, 2019, p. 30) showed that almost 38% of universities lacked research data management policies, and it was not in their current agendas. However, 21% have them, and 38% are in the process of developing them. The report was based on a survey of 321 European universities. In this same report, 43% did not have an institutional guideline for Open Access to research data, but 41% indicated they had a Data Protection Officer and not other data support roles (60%).

Perrier et al. (2017) performed a scoping review concerning research data management. Their study focused on the complete research data lifecycle. This is creating, processing, analyzing, preserving, giving access, and re-using data. They found that a significant part of the researched articles was addressed by providing access to data (31.13%), a phase where researchers are involved. In this phase, distribution, data sharing, control access, establishing copyright and promoting data are activities included in this phase. They also found researchers used a minimum required approach when it was necessary to fill a metadata form on depositing data into a repository, arguing that adding metadata to a dataset is not enough. Data quality, availability, accessibility, and reuse are also crucial regarding sharing policies.

Chigwada et al. (2017) performed a study in Zimbabwe where 25 research institutions participated. They found that researchers did not share their research data in repositories. This situation was a limitation to other researchers in the access and reuse of research data. In addition, they also found that research libraries had still limited services for research data management. They also found that research data were in textual, spreadsheet, and graphical formats.

Third, researchers do not invest time in metadata and documentation regarding research data. This means that the context of information is lost, and data reuse seems impossible (Ünal, Chowdhury, Kurbanoglu, Boustany, & Walton, 2019). This generates a problem of findability which is against the FAIR principles.

Ünal et al. (2019) performed a study analyzing the behavior of researchers in France, Turkey and the UK regarding their research data. They surveyed a sample (N=1098) of researchers. They found that 73% of researchers shared research data among their research teams and 55% with researchers from other institutions. However, 86% of researchers need to spend some pre-processing time on data received. Some reasons were that researchers did not use file naming standards to their files for the research data or also never used standard metadata for their research data.

Another reason involving metadata and documentation in research data is infrastructure. Since metadata practices are independent of the discipline (Mayernik, 2016), the infrastructure involves a great complexity because it should be helpful in all fields of the institution. Then, the organization needs to adopt standards from different disciplines and fit them together. For example, in Austria, an interdisciplinary data repository was designed and later built. Researchers took part in the repository's design, especially considering that researchers lacked a suitable subject repository. These researchers were mainly this target. They adopted Eprints, open-source software with the metadata schema IST DataRep, metadata adapted for research data (Petritsch, 2017).

Adding metadata to research data allows data to be shared and later reused. This is one of the reasons why metadata is essential. Metadata provides the context and knowledge to the data. The reuse will reduce research costs and allow combining the data into a more robust data set that can be reused for other purposes (Ramstrand, Fatone, Dillon, & Hafner, 2020). However, there are some limitations. Researchers who reuse data need a great understanding of the context to interpret the data, theories, tools and other contextual knowledge (Hansson & Dahlgren, 2022). However, there might be some copyright issues to limit the reshare of the metadata. This is the case in the cultural heritage sector like archives and museum collections. Another limitation to researchers sharing data is that competitors take advantage of their research, or data can be misinterpreted (Ramstrand et al., 2020).

Fourth, another question is technical. Researchers also have to deal with format recognition. This means that data may be represented in different forms, such as text, images or videos, which are not always standardized (Boté & Térmens, 2019). Another issue is the anonymization of the qualitative data. This generates a problem not in the accuracy of the data but the reuse of the data because there are no standards to codify the data (Attard, Orlandi, Scerri, & Auer., 2015). To maximize this effort, it seems necessary to adopt FAIR (findable, accessible, interoperable, and reusable) data management principles (Travieso Rodríguez & Ferreira Araújo, 2019; Wilkinson et al., 2016).

Data privacy can be a barrier to researchers in specific cases, such as clinic data access when the researcher is non-clinical. In Health Science, this situation may happen when non-clinician researchers need to test a hypothesis or study trends with medical data (Tantoso et al., 2019). In addition, consent is required for a single clinical trial, and data cannot be collected for other purposes, which is a waste of economic resources. This can be solved with a much broader consent allowing data to be stored as anonymized in clinical databases for future research efforts (Faden et al., 2013). Quinn (2017) pointed out that research data need to be anonymized when containing personal data. Then, the risk to anonymize the data is that it may destroy the value of the data. He also argued that the secondary use of data and the reuse raised privacy concerns, especially in health. It might be difficult, while not impossible, to reconcile the information consent provided because the new research may be different from the type of research in which the consent was signed.

Finally, funding agencies and other parties claim for research data sharing and researchers still do not have recognition in their careers to keep their data available. Consequently, researchers do not make an effort to liberate them, add the corresponding metadata, and upload them to a repository (Peset, Aleixandre-Benavent, Blasco-Gil, & Ferrer-Sapena, 2017).

3.4 The role of Information professionals and what can they do to help researchers and challenges with data management

The role of information professionals can be essential in supporting researchers when managing research data. One of the activities information professionals offer in higher education institutions is research data services that have emerged as a critical service. When an information professional is involved in the research data life cycle, their job's title reflects mainly science librarians that serve science. For example, data librarian, data science librarian, or research data management librarian (Bishop, Collier, Orehek, & Ihli, 2021). These information professionals cover various disciplines such as social sciences, engineering, physics, and life sciences. Nevertheless, an information professional can help researchers write DMP or implement it. It seems reasonable that the information professional can help standardize the research data cycle process. Then we stay in this section in four different aspects: training collaborated with researchers, stewardship services.

3.4.1. Training. Self-training and other training

Information professionals who want to deal with research data need the training to be specialized data librarians. Librarians with no prior training in research data management need to upgrade their skills in digitization, electronic resources management, data standardization skills, and literature gap (Barfi & Sackey, 2021). In addition, it is possible that in some countries, the role of data librarianship does not exist and also needs this training (Barfi & Sackey, 2021; Tammaro & Caselli, 2020).

In institutions where the budget is limited, or there is not enough workforce, they may train researchers team members or different researchers from different teams on how to deal with data management plans and data. This training can be offered through workshops, consultations, relevant courses for faculties or research teams or by creating online resources through collaborations between the library and other units of the research center. In the United States, 142 universities participated into a study concerning research data services staff. They found that the most perceived critical data services were "assistance with data archiving", "assistance with data preservation", or "assistance with data documentation", among others. They also found that the primary skills useful for research data service were "developing and teaching instructional content related to data services", "data management planning", and "data ethics" (Joo & Schmidt, 2020).

There are experiences where librarians and researchers are joined together in research data management training. Librarians and researchers can review data from others' research, and librarians have an opportunity to expand their data services around a specialized community (Muellenbach, 2021). Then, the training concerning research data that information professionals could offer should cover the complete research cycle. This is research planning, project management, data preservation or research dissemination (Barfi & Sackey, 2021; Bishop et al., 2021).

3.4.2. Collaboration with researchers: the embedded librarian and stewardship services

There are different approaches from the information professionals' point of view to help researchers, such as collaborating from the library as a liaison librarian, embedding an information professional into research teams, or data stewardship. One example of this situation was in Malawi, where a study was performed comparing two universities. The three most relevant research activities where librarians help researchers were research areas, data collection and data cleaning. Moreover, the three most critical competencies in managing research data for an information professional were identified as new standards and practices for curation, curating digital objects using curation lifecycle and long-term digital data preservation strategies (Chawinga & Zinn, 2021).

Information professionals can assist researchers' teams with information literacy topics such as disseminating scientific literature and generating knowledge. This is the case of the term "Embedded Librarian", where the information professional is integrated into strategic organizational points, supporting researchers in a research team in retrieving scientific literature (Wu & Mi, 2013). This same term could be renamed "Embedded Information Professional". In Italy, concerning new knowledge generation, there was an experience related to health literacy where five librarians were part of an Editorial Committee to build a new website for the National Institute of Health. Their former skills as a liaison librarian, such as cataloguing, indexing and information literacy, were helpful in the project. They were in charge of selecting content for the website in tasks for librarian curator, line editing, finding the right words online, taking profit from their indexing skills and search engine optimization (SEO), and taking advantage of working with metadata. Finally, they organized courses to combat misinformation in health. Thanks to their traditional role as a librarian, they could get other new skills by taking advantage of the skills they already had (Barbaro, Amicarella, Ferrari, Sorcini, & Zedda, 2021).

In South Korea, a study was performed studying the contributions of authors-contribution in scientific papers in many disciplines such as Medicine or Computer Science. They reported that the librarian's role as author-contribution was varied—for instance, search strategy, data curation or reading and approving the final manuscripts. Then, embedded librarians increased over the years, transforming their role as supportive in research projects (Shin, 2021).

Another approach from the information professionals' point of view to help researchers can also be data stewardship. Data stewardship is a very specialized position where the data steward has skills in a concrete discipline and concrete duties. Concerning the data steward title, Tammaro & Caselli (2020) pointed out that data steward has different names depending on the institutional context and the international border. For instance, "data librarian" or "data curator" are some of the characters. Then, the role of the data steward is to support data reuse, storage, and later access of data. This last aspect is important because research data has to be digitally preserved.

In Austrian universities, Gruber, Schranzhofer, Knopper, Stryeck, and Hasani-Mavriqi (2021) exposed three models of the existing data steward: service point, data steward center and data steward network. The data steward as a service point would offer support, but they do not take any action in the institutes they worked.

In the second role, the data steward center, stewards work together in the same office, and their competencies are distributed. The third role consists of a person who works in a faculty or specializes in a concrete field plus one coordinator. They can offer support to a discipline—for instance, a data steward in librarianship. Indeed, data stewardship (Awada, Phillips, & Bogdan, 2022) can ensure that data policies and standards are effectively applied. Orrù (2020) pointed out that data stewards' skills respond to Data Management Plans and support researchers on how data is generated or retrieved, how information is protected or how the data is described and documented.

4 CONCLUSIONS

The emergence of the FAIR principles responds to a need in the research community to provide guidance for sharing and reusing research data. Many stakeholders in the digital ecosystem of research data endorsed the principles.

There are difficulties in being able to evaluate the implementation of the principles on research data, metrics and evaluation methods are still being developed and debated. Researchers, the creators of the data, find it difficult to make the data FAIR, from the application of metadata, to recognizing formats, to the application of standards.

Information professionals can play an important role in the implementation of the principles, offering training in research centers, or collaborating in the research itself and managing the data; in any of the facets the information professional must complete their academic training with specific training on everything surrounding the ecosystem of the FAIR principles.

REFERENCES

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, *32*(4), 399–418. doi:10.1016/j.giq.2015.07.006

Awada, L., Phillips, P. W. B., & Bogdan, A. M. (2022). Governance and stewardship for research data and information sharing: Issues and prospective solutions in the transdisciplinary plant phenotyping and imaging research center network. *Plants People Planet, 4*(1), 84–95. Scopus. doi:10.1002/ppp3.10238

Bahim, C., Casorrán-Amilburu, C., Dekkers, M., Herczog, E., Loozen, N., Repanas, K., ... & Stall, S. (2020). The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments. *Data Science Journal, 19*(1), p.41. doi:10.5334/dsj-2020-041

Bahim C., Dekkers M., Wyns B. (2019). Results of an Analysis of Existing FAIR Assessment Tools. doi:10.15497/rda00035

Barbaro, A., Amicarella, S. E., Ferrari, P., Sorcini, I., & Zedda, M. (2021). Embedded librarians: An innovative experience in health and wellness communication. Health Information and Libraries Journal. doi:10.1111/hir.12410

Barfi, F. K., & Sackey, E. K.-A. (2021). The Role of the Technical Universities' Librarians in the Generation and Management of Technical Research Data (TRD) to Advance Inventions, Innovation and Commercialization in Ghana. *Library Philosophy and Practice, 2021*, 1–20.

Boté, J.-J., & Térmens, M. (2019). Reusing data: Technical and ethical challenges. *DESIDOC Journal of Library and Information Technology*, *39*(6), 329–337. doi:10.14429/djlit.39.6.14807

Chawinga, W. D., & Zinn, S. (2021). Research Data Management in Universities: A Comparative Study from the Perspectives of Librarians and Management. *International Information & Library Review, 53*(2), 97–111. doi:10.1080/10572317.2020.1793448

Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research Data Management in Research Institutions in Zimbabwe. *Data Science Journal*, *16*(0), 31. doi:10.5334/dsj-2017-031

Data Fairport. (2014). Data FAIRport. Retrieved from https://www.datafairport.org/index.html

Deroo, K., Hermans, E., Lammertyn, J., & Mertens, M. (2019). Generic DMP Evaluation Rubric. DMP Belgium. https://github.com/DMPbelgium/Guidance/blob/6dc18570ab4f82173bd51911c25028c4c30bbd34/DMPReviewRubrics/UGentGenericDMP_Rubric.pdf (Original work published 2017)

Devaraju, A., & Huber, R. (2021). An automated solution for measuring the progress toward FAIR research data. *Patterns*, *2*(11). doi:10.1016/j.patter.2021.100370

FAIRsharing Team, & University of Oxford (2019). FAIRassist.org. Retrieved from https://fairassist.org

Faden, R. R., Kass, N. E., Goodman, S. N., Pronovost, P., Tunis, S., & Beauchamp, T. L. (2013). An Ethics Framework for a Learning Health Care System: A Departure from Traditional Research Ethics and Clinical Ethics. *Hastings Center Report*, *43*(s1), S16–S27. doi:10.1002/hast.134

G20 leaders (2016). G20 Leaders' Communique Hangzhou Summit. *europa.eu*. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_16_2967

GO FAIR. (2017). Progress towards the European Open Science Cloud: GO FAIR Office established. Retrieved from https://www.go-fair.org/2017/12/07/progress-towards-european-open-science-cloud-go-fair-office-established/

Gruber, A., Schranzhofer, H., Knopper, S., Stryeck, S., & Hasani-Mavriqi, I. (2021). Kompetenzen von Data Stewards an österreichischen Universitäten. *VOEB-Mitteilungen, 74*(1), 12–32. doi:10.31263/voebm.v74i1.6255

Hansson, K., & Dahlgren, A. (2022). Open research data repositories: Practices, norms, and metadata for sharing images. *Journal of the Association for Information Science and Technology*, *73*(2), 303–316. doi:10.1002/asi.24571

Hermans, E. (2019). *4 Challenges for institutional research data management support*. OpenAIRE. Retrived from https://www.openaire.eu/blogs/4-challenges-for-institutional-research-data-management-support

Collins, S., Genova, F., Harrower, N., Hodson, S., Jones, S., Laaksonen, L., ... & Wittenburg, P.(2018). Turning FAIR into reality: Final report and action plan from the European Commission expert group on FAIR data. Retrieved from https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_0.pdf

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., ... & Schultes, E. (2020). FAIR principles: interpretations and implementation considerations. *Data Intelligence, 2*(1-2), 10-29. doi:10.1162/dint_r_00024

Jahnke, L., & Asher, A. (2012). The Problem of Data: Data Management and Curation Practices Among University Researchers*. CLIR*. Retrieved from https://www.clir.org/pubs/reports/pub154/problem-of-data/

Joo, S., & Schmidt, G. M. (2020). Research data services from the perspective of academic librarians. *Digital Library Perspectives, 37*(3), 242–256. doi:10.1108/DLP-10-2020-0106

Lahti, K., Schulman, L., Piirainen, E., Riihikoski, V.-M., & Juslén, A. (2019, June 26). *'As Open as* Possible*, as Closed as Necessary' – Managing legal and owner-defined restrictions to openness of biodiversity data: Biodiversity_Next*. doi:10.3897/biss.3.37395

LIBER. (2018). Open Consultation on FAIR Data Action Plan. Retrieved from https://libereurope.eu/article/fairdataconsultation/

Martone M. E. (2015) FORCE11: Building the Future for Research Communications and e-Scholarship. *Bioscience 65,* 635. doi:10.1093/biosci/biv095?login=true

Mayernik, M. S. (2016). Research data and metadata curation as institutional issues. *Journal of the Association for Information Science and Technology*, *67*(4), 973–993. doi:10.1002/asi.23425

Morais, R., & Borrell-Damián, L. (2019). *2017-2018 open access survey results.pdf*. https://eua.eu/downloads/publications/2017-2018%20open%20access%20survey%20results.pdf

Muellenbach, J. M. (2021). A Pilot to Initiate Research Data Management Services Within Academic Libraries Helps Librarians to Learn About, Engage With, and Enhance Skills Within Their Research Communities. *Evidence Based Library and Information Practice, 16*(1), 104–106. doi:10.18438/eblip29879

Orrù, D. (2020). Open data steward: Librarians and open data literacy [Open data steward: Bibliotecari e alfabetizzazione ai dati aperti]. *AIB Studi, 60*(2), 311–323. doi:10.2426/aibstudi-12123

Perrier, L., Blondal, E., Ayala, A. P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M., Trimble, L., & MacDonald, H. (2017). Research data management in academic institutions: A scoping review. *PLOS ONE*, *12*(5), e0178261. doi:10.1371/journal.pone.0178261

Peset, F., Aleixandre-Benavent, R., Blasco-Gil, Y., & Ferrer-Sapena, A. (2017). Datos abiertos de investigación. Camino recorrido y cuestiones pendientes. *Anales de Documentación*, *20*(1), Article 1. doi:10.6018/analesdoc.20.1.272101

Petritsch, B. (2017). Metadata for Research Data in Practice. *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare*, *70*(2), 200–207. doi:10.31263/voebm.v70i2.1678

Quinn, P. (2017). The Anonymisation of Research Data—A Pyric Victory for Privacy that Should Not Be Pushed Too Hard by the eu Data Protection Framework? *European Journal of Health Law*, *24*(4), 347–367. doi:10.1163/15718093-12341416

Ramstrand, N., Fatone, S., Dillon, M. P., & Hafner, B. J. (2020). Sharing research data. *Prosthetics and Orthotics International*, *44*(2), 49–51. doi:10.1177/0309364620915020

Rice, R. (2019). Supporting Research Data Management and Open Science in Academic Libraries: A Data Librarian's View. *VOEB-Mitteilungen*, *72*(2), 263–273. doi:10.31263/voebm.v72i2.3303

RDA. (2016). Adoption of Recommendations. Retrieved from https://www.rd-alliance.org/recommendations-outputs/adoption-recommendations

Shin, E.-J. (2021). Embedded librarians as research partners in South Korea. *Journal of Librarianship and Information Science, 53*(3), 466–474. doi:10.1177/0961000620962550

Schwanitz, V. J., Wierling, A., Biresselioglu, M. E., Celino, M., Demir, M. H., Bałazińska, M., ... & Suna, D. (2022). Current state and call for action to accomplish findability, accessibility, interoperability, and reusability of low carbon energy data. Scientific Reports, 12(1), 1-9

Tammaro, A. M., & Caselli, S. (2020). Training Data Stewards in Italy: Reflection on the FAIR RDM Summer School. In *Italian Research Conference on Digital Libraries*, 163-172. Springer, Cham. doi:10.1007/978-3-030-39905-4_16

Tantoso, E., Wong, W.-C., Tay, W. H., Lee, J., Sinha, S., Eisenhaber, B., & Eisenhaber, F. (2019). Hypocrisy Around Medical Patient Data: Issues of Access for Biomedical Research, Data Quality, Usefulness for the Purpose and Omics Data as Game Changer. *Asian Bioethics Review*, *11*(2), 189–207. doi:/10.1007/s41649-019-00085-3

The FAIR Steering Group. (2017). Policy statement on FAIR access to Australia's research outputs. Retrieved from https://www.fair-access.net.au/fair-statement

Top, J., Janssen, S., Boogaard, H., Knapen, R., & Şimşek-Şenel, G. (2022). Cultivating FAIR principles for agri-food data. *Computers and Electronics in Agriculture, 196*. doi:10.1016/j.compag.2022.106909.

Travieso Rodríguez, C., & Ferreira Araújo, R. (2019). Methodological issues of open research data: Analysis of the datasets from SciELO included in Figshare [Aspectos metodológicos de los datos abiertos de investigación: Análisis de los conjuntos de datos de la colección SciELO incluidos en Figshare]. *Revista Espanola de Documentacion Cientifica*, *42*(3). doi:10.3989/redc.2019.3.1597

Tripathi, M., Shukla, A., & Sonkar, S. K. (2017). Research Data Management Practices in University libraries: A study. *DESIDOC Journal of Library & Information Technology*, *37*(6), 417. doi:10.14429/djlit.37.11336

Ünal, Y., Chowdhury, G., Kurbanoglu, S., Boustany, J., & Walton, G. (2019). *Research Data Management And Data Sharing Behaviour Of University Researchers*. http://openaccess.hacettepe.edu.tr:8080/xmlui/handle/11655/23736

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci Data, 3*. doi:10.1038/sdata.2016.18

Wu, L., & Mi, M. (2013). Sustaining Librarian Vitality: Embedded Librarianship Model for Health Sciences Libraries. *Medical Reference Services Quarterly, 32*(3), 257–265. doi:10.1080/02763869.2013.806860