Putting FAIR Principles in the Context of Research Information: FAIRness for CRIS and CRIS for FAIRness

Otmane Azeroual[1][a], Joachim Schöpfel[2][b], Janne Pölönen[3][c] and Anastasija Nikiforova[4,5][d]

[1]*German Centre for Higher Education Research and Science Studies (DZHW), Schützenstraße 6A, 10117 Berlin, Germany*

[2]*GERiiCO Laboratory, University of Lille, 59653 Villeneuve-d'Ascq, France*

[3]*Federation of Finnish Learned Societies, Snellmaninkatu 13, 00170 Helsinki, Finland*

[4]*Institute of Computer Science, University of Tartu, Narva mnt 18, 51009 Tartu, Estonia*

[5]*European Open Science Cloud (EOSC) Task Force "FAIR metrics and data quality", 1050 Brussels, Belgium*

fi

Keywords: CRIS, FAIR, Findability, Data Management, Knowledge, Information Management, Interoperability, Research Data Management, Research Information System, Open Science.

Abstract: Digitization in the research domain refers to the increasing integration and analysis of research information in the process of research data management. However, it is not clear whether it is used and, more importantly, whether the data are of sufficient quality, and value and knowledge could be extracted from them. FAIR principles (Findability, Accessibility, Interoperability, Reusability) represent a promising asset to achieve this. Since their publication, they have rapidly proliferated and have become part of (inter-)national research funding programs. A special feature of the FAIR principles is the emphasis on the legibility, readability, and understandability of data. At the same time, they pose a prerequisite for data for their reliability, trustworthiness, and quality. In this sense, the importance of applying FAIR principles to research information and respective systems such as Current Research Information Systems (CRIS), which is an underrepresented subject for research, is the subject of the paper. Supporting the call for the need for a "one-stop-shop and register-once-use-many approach", we argue that CRIS is a key component of the research infrastructure landscape, directly targeted and enabled by operational application and the promotion of FAIR principles. We hypothesize that the improvement of FAIRness is a bidirectional process, where CRIS promotes FAIRness of data and infrastructures, and FAIR principles push further improvements to the underlying CRIS.

1 INTRODUCTION

Today, more and more data and information - both produced, collected, and available from the past - stored for decades in paper form, are being digitized, which is also the case for the research domain. However, although digitization refers to making data available in an electronic and machine-readable format for further use, making it significantly more efficient, it is not clear whether the data are of sufficient quality for further use, and they can be transformed into value and knowledge. In other words, digitization does not necessarily involve data quality management, while data require quality management as they are often affected by data quality issues of various nature ((Ferraris et al., 2018), (Ivanović et al., 2019), (Corte-Real et al., 2020), (Nikiforova, 2020), (Azeroual et al., 2022)). Here, FAIR principles become a promising asset.

The FAIR principles were originally developed as guidelines or recommendations for the effective and efficient management of research data and stewardship as part of a new open science policy framework, with a "specific emphasis on enhancing the ability of machines to automatically find and use the data" in data repositories (Wilkinson et al., 2016). Since then, FAIR principles have become central element in the debate and implementation of open science policies, and they are increasingly being applied to other "digital objects" (Wittenburg, 2019) of different levels such as institutional repositories and large infrastructures such as European Open Science Cloud (EOSC)

[a] https://orcid.org/0000-0002-5225-389X
[b] https://orcid.org/0000-0002-4000-807X
[c] https://orcid.org/0000-0003-1649-0879
[d] https://orcid.org/0000-0002-0532-3488

Azeroual, O., Schöpfel, J., Pölönen, J. and Nikiforova, A.
Putting FAIR Principles in the Context of Research Information: FAIRness for CRIS and CRIS for FAIRness.
DOI: 10.5220/0011548700003335
In *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2022) - Volume 3: KMIS*, pages 63-71
ISBN: 978-989-758-614-9; ISSN: 2184-3228

and the German NFDI (Nationale Forschungsdateninfrastruktur – the mandated organisation for Germany in the EOSC), but also to metadata, identifiers, catalogs, software, or scientific practice. The FAIR principles have been applied in the handling of electronic theses and dissertations (Ivanović et al., 2019) and as criteria for assessing technological behaviors (Mornati, 2019).

What is more, an investigation conducted by PwC EU Services in 2018 (European Commission and Innovation, 2019) showed that the annual cost of not having FAIR data to a minimum of €10.2 billion per year and 16.9 billion in lost innovation opportunities (Masuzzo, 2022), where the actual costs are likely to be significantly higher due to unquantifiable elements such as the value of improved research quality and other indirect positive spill-over effects of FAIR research data. They found that the impact on innovation, would account for over 60% of the likely cost of not having FAIR research data, while the minimum true cost of not having FAIR research data, encompassing indicators such as – "time spent", "license costs", "research duplication", "cost of storage" and "research retraction" – accounts for the remaining 40%. These indicators, however, represent three areas applicable to all sectors (i.e. academic, private, public, non-profit) and can be described as (1) impact on research activities, (2) impact on collaboration, and (3) impact on innovation. Therefore the topic of FAIRness becomes increasingly important.

This can also be seen as one of the reasons, why in the European Commission Open Science policy, FAIR and open data sharing is one of the eight pillars of Open Science (Commission, 2021). While in the emerging EOSC, which is represented by one of the authors of this study, the FAIRsFAIR project addresses "the development and concrete realization of an overall knowledge infrastructure on academic quality data management, procedures, standards, metrics and related matters, based on the FAIR principles", as a kind of general reference / guide to best practices in Higher Education and Research (FAIRsFAIR, 2022). As outlined in the UNESCO Recommendation on Open Science, FAIRness has become an essential feature of what has been called "open science culture" over the past couple of years. Promoting a culture of open science requires, for example, the development of rewards and incentives that "give value to all relevant research activities and scientific outputs including high-quality FAIR data and metadata" (UNESCO, 2021). At the moment, however, EOSC have begun their own research to define and introduce guidelines to application of FAIR principles to digital objects not necessarily limited to the research domain, thereby expanding the scope to the entire digital environment and all data and data-based objects (also referred to as "digital objects"), including research objects such as scientific articles and software, available on the Internet, while the main focus remains on the research domain. In this study, however, we refer to the research domain and FAIRness of research information systems (RIS).

Research information management systems or Current Research Information Systems (hereafter CRIS) are seen as "core elements of the technological solution since they provide rich additional metadata on datasets and put the datasets and their metadata into their proper context, and so significantly enhance the FAIRness of datasets" (Terheggen and Simons, 2016). In other words, CRIS are not just data repository, but a key component of the research infrastructure landscape, which is directly targeted and involved by the operational application and promotion of the FAIR principles. This involvement can be described at three levels constituting a set of propositions of our study on which we will elaborate within this paper:

1. research information management systems (CRIS) are helpful to assess the FAIRness of research data and data repositories;
2. research information management systems (CRIS) contribute to the FAIRness of other research infrastructure;
3. research information management systems (CRIS) can be improved through the application of the FAIR principles.

This study is exploratory in nature. First, we provide an overview of the challenges associated with research data and research information management, then we describe each level based on a review of relevant studies, and conclude with some perspectives for further research, thereby raising an awareness of this topic and making a call for other researchers to refer to it.

2 RESEARCH INFORMATION: WHY DOES IT MATTER?

Research information or an information about research is a crucial resource for research institutions that is dynamic in nature and growing in size. This is also due to digitization, namely the fact that the way researchers work has changed, and more and more data are being digitized and stored electronically instead of paper-based archives. But precisely because these data, i.e. research information, is not only the

basis of scientific knowledge processes, but is also related to other data, research data management (RDM) is becoming more and more important to enable and ensure comprehensive access to data stocks so that research results can not only be verified and interpreted, but it can also be understood how these results were obtained and how they can be made usable and actionable.

All in all, the "research information" term stands for "technical data, processes, methods, inventions, compositions-of-matter, and biological materials, equipment, instruments, apparatuses, devices, articles of manufacture or component parts(s) thereof developed under or resulting from the performance of Research under the Research Agreement and improvements thereof, except for improvements on new subject matter not funded by Licensee. In the event that computer software or information management systems are created or developed in the performance of Research under the Research Agreement such computer software or information management systems shall be treated as Research Information" (Insider, 2022).

In order to ensure the effectiveness, visibility and acceptance of RDM in the long term, research information should be sustainable and usable by a wide range of science and society (Flores et al., 2015). For instance, the OECD has published guidelines on access to research data for public funding (OECD, 2007), and the Alliance of German Science Organizations has adopted principles for handling research data (Society, 2010). The EU funding program Horizon 2020 has introduced FAIR data management (Commission, 2020a) guidelines / recommendations, and the new Horizon Europe program confirms that research data management cannot be "opt-out" and that projects generating research data must manage their data responsibly and in line with FAIR principles (Commission, 2020c).

One of the main goals of open science is to increase the transparency and accessibility of scientific results (Paic, 2021). The FAIR data principles set the standard for the sustainable use of research data, insofar metadata are crucial for the later interpretability, interoperability and traceability of data, including quality assurance and control measures.

The ever-increasing relevance of data for research requires new and efficient strategies for processing and handling research information and research data. For this purpose, the National Research Data Infrastructure (NFDI) is currently building a digital knowledge repository for the development, networking and use of research data in Germany (Wachtler et al., 2021). This includes, for example, making data collection or RDM transparent and understandable. This is important given that in universities and research institutions it is often unclear what methods are used to collect the data or how their trustworthiness is ensured (Guba and Lincoln, 1989), (Azeroual and Schöpfel, 2021). In this sense, the context of the data is of interest, so information about the collection strategy, as well as their structure, is essential.

In addition, the data must be of a certain quality in order to be further used, including but not limited for research ((Ferraris et al., 2018), (Ivanović et al., 2019), (Corte-Real et al., 2020), (Nikiforova, 2020), (Azeroual et al., 2022)). The data quality of research information, however, is in most cases ensured manually, which obviously comes with a lot of effort, where automated or at least semi-automated quality assurance processes significantly reduce these efforts. This applies in particular to automatic plausibility and completeness checks directly at the time of data entry with data profiling and data cleansing / cleaning, machine learning processes for text recognition (such as text and data mining), conspicuous data point or data anomaly detection, and statistical processes for data normalization. However, they must be applicable in practice, where some of the above can be solved by designing a data schema or defining data entry constraints or with validation checks – although even they are lacking in most cases, while some of them turn out to be less primitive in their development and implementation.

The exchange of information and knowledge is one of the pillars of science. With the increasing digitization in science and the spread of the philosophy of open science, research information is also becoming more accessible to the public. These data are in various open and proprietary formats in various data repositories. Metadata are used to describe them, the scope of which is mostly determined by the capabilities of the data portals or specialized communities. The quality of the metadata recording depends on the professional competence, including data literacy, and completeness of individual curators (Tammaro et al., 2019).

With the growing volume of data and the emerging cross-disciplinary challenges, the need for a consistent, interoperable framework for documenting data that can reliably find, filter and compare research information of different origins is becoming increasingly evident. Largely supported by Artificial Intelligence (AI), Big Data analytics and Machine Learning (ML) methods require data to be not only accessible to humans, but also be machine-readable and understandable (Nguyen et al., 2018). This requires a selection of standardized, long-life readable data formats,

as well as clear coding of the metadata descriptions, standardizing the two attribute fields and their permissible content (attribute values). The four FAIR data principles, introduced by RDM, specify goals, but do not provide solutions for this non-trivial task.

The FAIR data principles describe an ecosystem of data, metadata, scientific software, but also workflows, metrics, and the need for continued funding of infrastructures that are consistent with the implementation of the FAIR principles (e.g. repositories) (Arefolov et al., 2021). The unique identification of data (items), people, institutions, projects, as well as policies, guidelines, standards and data repositories are indispensable tools in such a system. The question implies from the above - whether and how can CRIS contribute to promotion of FAIRness?

3 CRIS AS AN ASSET IN ASSESSING FAIRness

The FAIR data principles provide a comprehensive framework and guidance on the criteria that well-preserved data must meet, as well as on the standardization of all data schemes (Mayer et al., 2021). The quality of data depends not only on the accuracy of the measurement and recording methods and, of course, on their scientific relevance, but increasingly on the quality of their processing and storage (National Academy of Sciences (US) et al., 2009). The FAIR principles can be seen as the gold standard for data quality ("fit of use"), generally recognized as a generic standard (Hasselbring et al., 2020).

Several tools have been developed over the years to assess the FAIRness of research data and data repositories, such as the Australian Research Data Commons (ARDC) FAIR Data self-assessment tool, the Dutch DANS FAIRdat tool or the EUDAT Fair Data Checklist (FOSTER, 2020). Moreover, the results obtained using different tools tend to differ significantly with a very vague understanding on what should be done to improve the result if another tool assessed the level of FAIRness as sufficient. All in all, such "ad-hoc" tools may be more or less useful for the follow-up of local FAIR programs or for training FAIR principles, thereby developing FAIR literacy. However, they create new information silos and in most cases are not linked to professional assessment systems such as CRIS.

Open data practices and the FAIRness of research data have become essential characteristics of research performance. The European Commission's Open Science Monitor includes national level information on open research data (Commission, 2020b). "Using the FAIR data principles" is one of the evaluation criteria proposed in OS-CAM, the Open Science Career Assessment Matrix (O'Carroll et al., 2017), (European Commission and Innovation, 2021). The CRIS can and even should collect, aggregate and integrate structured and carefully curated information on research data and its FAIRness to support the monitoring and assessment of this element.

For the purpose of this study, we selected euroCRIS repository - as a major source used by euroCRIS community, which participants represent practitioners dealing with the actual national, regional and international CRIS - as a source for conducting a extracting the most relevant studies with an explicit elaboration of FAIR principles and their implementation within CRIS. All in all, after a systematic literature review, three studies were found to be relevant to illustrate this approach. In other words, all relevant studies available in euroCRIS repository were selected and studied. The low number of relevant studies, however, points out the limited body of knowledge on this topic, thereby making this study unique and constituting a call for action. Let us elaborate on them. The scope of the research, however, was extended by referring to other relevant projects and initiatives found around the world, which list was identified using a snowballing approach, i.e., referring to the projects covered in the selected studies, or based on our own experience dealing with this topic at both regional, national and international levels and representing different communities.

According to (Lindelöw, 2019), in 2017, the government of Sweden gave the Swedish Research Council and the National Library of Sweden parallel assignments to propose criteria and a method for assessing how well research data and scholarly publications produced by Swedish organizations comply with the FAIR principles, based on the assumption that "the products of research must meet the FAIR principles as far as possible". Suggested criteria include (1) metadata quality (richness), (2) licensing and persistent identifiers, (3) openness, (4) accessibility and (5) standard vocabularies. The aim is to provide an "overall picture of FAIRness" of national research results, through the collected metadata.

Authors of (Miniberger and Reding, 2017) described how the implementation of a commercial CRIS at the University of Vienna and the creation of a national network of CRIS managers from all Austrian universities (FIS/CRIS Austria) contributed to the visibility, findability, accessibility and interoperability of research information, through the development of standards (including identifiers and data mod-

els) and shared strategies. More recently, this network developed a tool that enables tracking and monitoring of the transition to open access based on data stored in local CRIS which is interoperable and connected with OpenAIRE (Danowski et al., 2020).

As part of the AT2OA (Austrian Transition to Open Access) project, a sub-project deals with the development of a concept for monitoring the Open Access (OA) publication output in Austria. Authors of (Danowski et al., 2020) build their feasibility study on the analysis of international best practice models and aim to demonstrate the added value and feasibility of OA monitoring at national level. However, the idea of (Danowski et al., 2020) should also serve to (further) develop OA monitoring in other countries and also to support OA monitoring in an international context.

Based on experience with the Flemish research information system (FRIS), an application profile for research data was presented in (Vancauwenbergh, 2021), including various aspects of metadata such as description, discovery, contextualisation, coupling users, software and computing resources to data, research proposal, funding, project information, research outputs, outcomes, impact etc., which allows assessing FAIRness and compliance with open science policy. The author concluded that convergence to a common metadata model and interoperation/ interoperability across multiple metadata models are two conditions for developing such an application. This approach may also allow for the "FAIR labeling" of research infrastructures that comply with the FAIR principles.

All in all, what three initiatives have in common is that CRIS is used to assess different levels of FAIRness and FAIRness of the object under assessment at different levels, as part of the global open science assessment. This assessment is primarily based on the collection and processing / handling of metadata. Regarding researcher evaluation, for example, the Finnish and Norwegian national policies on responsible assessment of researchers indicate the national and/or local CRIS as a potential source of documentation of open data practices (of Finnish Learned Societies, 2020), (Norway, 2016).

In general, the initiatives discussed above confirm that CRIS has the potential and capacity at the institutional, regional or national level to contribute to the monitoring of open science policies and, in particular, to the follow-up of projects aimed at improving the FAIRness of research data, research repositories and other related research infrastructures. In addition, it has been clearly recognized that CRIS has the potential to support and facilitate more responsible research assessment systems to reward and incentivize researchers for open science practices, including open and FAIR data.

4 CONTRIBUTING TO THE FAIRness OF RESEARCH INFRASTRUCTURES

The current body of knowledge, i.e. scientific literature on the topic, suggests that CRIS can contribute to improving the FAIRness of other research infrastructures, such as data repositories or publication platforms.

One of the main reasons for this potential is the central position of CRIS in the research information ecosystem (Donohue et al., 2018). In short, CRIS obtains data from a wide range of external and internal sources such as scientometric databases, library catalogs, human resources, finance, project management etc., and provides the data or information in standardized formats to other infrastructures and research information tools. This potential is not a general and common characteristic of all types of CRIS - it rather depends on the degree of standardization of the CRIS data model and the CRIS format. In other words, since CRIS require and depend on exchange with (many) other infrastructures, their quality, effectiveness and performance are affected by the standardization of data and procedures. For this reason, they have a kind of standardizing impact on other infrastructures, both down-stream (output) and upstream (input).

All in all, one of the goals of CRIS is to support the monitoring and evaluation of research performance at the national, institutional or even individual levels, which usually requires and therefore indirectly facilitates the production and collection of rich, complete, structured, comprehensive, comparable, interoperable data. As an example that one of the authors came across in Finland, the need to integrate metadata from the institutional CRIS into the national system forced the development of definitions, standards and procedures at national level and their further implementation locally, thus contributing to standardization.

This positive effect has been described in various research information projects and systems such as the implementation of PURE at the University of Vienna (Miniberger and Reding, 2017), some DSpace-CRIS projects in Italy, Cyprus, Australia and Hongkong (Mornati et al., 2018), CRIS of Radboud University in The Netherlands (Jetten et al., 2018), (Jetten and Simons, 2019) and, more generally, the DSpace, Fedora

and Vivo implementations around the world (Donohue et al., 2018). The Dutch project shows that local infrastructure and support are vital for a data management policy to work, that "we need to get away from the silo- and closed vault-thinking" in order to move to a "one-stop-shop and register-once-use-many approach" based on a standard data model and format, with the argument that connecting shared (possibly, 'dark') local data repositories to CRIS will make them FAIR. In this sense, (Ivanović et al., 2019) describe how English metadata in the local CRIS improve the findability of Serbian theses and dissertations deposited in the institutional Electronic Theses and Dissertations (ETD) repository. Following (Vancauwenbergh, 2021), the need for standard data and metadata will also improve the interoperability of research infrastructures, which is another FAIR guiding principle.

The preferred standard data model and format of all these projects is the Common European Research Information Format (CERIF), flexible enough to include new elements, semantics and relations, and able to provide standard output to other infrastructures, thus increasing their FAIRness. A Swedish study explores how CERIF can improve the FAIRness of open repositories (Engelman et al., 2019). (Engelman et al., 2019) find that the "CERIF model is used to represent research information and to transfer it between repositories (...). When CERIF is employed in relevant archive processes, a FAIR compliant archive is easier to achieve", with an "archival structure based on a cfProject tree (...), archived objects (...) represented by cfResult entities and their descriptive metadata (...) given in attached Cerif entities (...)".

5 IMPROVING THE FAIRness OF CRIS

In other words, while FAIR is typically discussed at three levels - (1) digital object, which refers to dataset, videos, journals, books etc., (2) metadata about this object on elementary level, including title, creator, identifier, etc., and (3) metadata records with the reference to the body of metadata element on the object in a specific database (Engelman et al., 2019), we suggest refer not only to data and information, but also to the upper level of the data or information management systems, i.e. CRIS. Thus, our third proposition is that CRIS itself can be improved by following and / or applying FAIR principles on it. In other words, CRIS is not only suitable to improve the FAIRness of research data management, but the FAIR principles are also beneficial for the further development of sustainable and FAIR CRIS. Indeed, the Science Europe Position Statement on Research Information Systems suggests that "research information systems should foster the findability, accessibility, interoperability, and reusability of the data that they store by implementing the FAIR Guiding Principles for research activity data" (Europe, 2016).

Two levels can be distinguished here. First, the need for standard data and metadata, especially persistent identifiers, requires tools capable of producing and processing, handling them, and this is a strong argument in favor of CRIS as the central system (middleware) in the research infrastructure ecosystem. Secondly, this need also calls for more standardization of CRIS, improved data models and formats especially the long tail of less standardized research information systems (cf. the large diversity and heterogeneity of CRIS revealed in the OCLC survey (Bryant et al., 2021).

The positive impact of the FAIR principles on the CRIS infrastructures is less documented, probably because most of the research is focused on "good examples" or "best practices" and highly exemplary projects, with a high degree of standardization. However, the standard format can improve CRIS interoperability, although this is not sufficient – CRIS should (also) prefer open identifier systems "to make things findable" and link information on source data and rights information, for instance, to support access and facilitate reuse (Tatum and Brown, 2018).

The interconnection of infrastructures based on the FAIR principles is another example of an improvement in CRIS, which must fulfill certain technical requirements based on the FAIR principles. This "need for upgrading" is illustrated in (Mornati et al., 2018) that describes a list of FAIR requirements for joining the European OpenAIRE community. In other words, only CRIS meeting these criteria will be connected to OpenAIRE – another almost mandatory way to improve (and enforce) the FAIRness of CRIS.

However, we must keep in mind that the FAIRness of research information management infrastructures has its specific limitations due to the nature of the data and the potential impact of their reuse. Some of the data can be personal data and protected by privacy laws such as General Data Protection Regulation (GDPR), other information may be confidential for other reasons, e.g. being highly financial and of interest for competitors etc. Thus, for ethical and legal reasons, the accessibility of CRIS data must be controlled and respect the above, i.e. it cannot be a guideline and require an openness of all data, following the H2020 Program Guidelines on FAIR Data of "as open as possible, as restricted as necessary" (Landi et al.,

2020) and distinguishing FAIR and open data, while the preferred option is a combination of both.

6 DISCUSSION AND CONCLUSIONS

This paper focuses on FAIR data principles and their potential and real contribution to the quality and the reusability of research information with the further possibility of creating knowledge from it and an efficient knowledge management, which is only possible if the list of the discussed prerequisites is met. It is crucial that the research information is available in such a way that it can be found, accessed, linked and reused as easily as possible (for authorized users) thereby being as FAIR as possible. Essentially, it is a quality requirement for data management as a prerequisite for being able to use the data, which are often obtained with great effort, which quality is not preserved when the data are stored and further processed thereby losing their value.

Research information is not just research data, and research information management systems such as CRIS are not just repositories for research data. They are much more complex, alive, dynamic, interactive and multi-stakeholder objects. However, in the real-world they are not directly subject to the FAIR research data management guiding principles. But as described above, CRIS are part of the research infrastructure ecosystem and are linked to data repositories, where the idea of CRIS partly overlaps with the main goal of FAIR principles, where the original scope of CRIS is more limited. At the same time, the scope of the FAIR principles has been extended recently to include a broader variety of "digital objects", infrastructures and content. For both reasons, CRIS can (and already does) improve the FAIRness of research infrastructures and data through the evaluation (monitoring) and standardization of data and metadata.

In this paper we have raised a discussion on this topic showing that the improvement of FAIRness is a dual or bidirectional process, where CRIS promotes and contributes to the FAIRness of data and infrastructures, and FAIR principles push for further improvement in the underlying CRIS data model and format, positively affecting the sustainability of these systems and underlying artifacts. CRIS are beneficial for FAIR, and FAIR is beneficial for CRIS. Nevertheless, as pointed out by (Tatum and Brown, 2018), the impact of CRIS on FAIRness is mainly focused on the (1) findability ("F" in FAIR) through the use of persistent identifiers and (2) interoperability ("I" in FAIR) through standard metadata, while the impact on the other two principles, namely accessibility and reusability ("A" and "R" in FAIR) seems to be more indirect, related to and conditioned by metadata on licensing and access. Paraphrasing the statement that "FAIRness is necessary, but not sufficient for 'open'" (Tatum and Brown, 2018), our conclusion is that "CRIS are necessary but not sufficient for FAIRness".

In terms of rewards and incentives, the FAIRness of CRIS data, as recommended by European Commission, is critical to ensure the "independence and transparency of the data, infrastructure and criteria necessary for research assessment and for determining research impacts" (European Commission and Innovation, 2021). While CRIS has great potential to support more responsible assessments with reliable, comprehensive, well-structured and comparable qualitative and quantitative data and metrics on research, they still offer only limited support for the evaluation of a broad range of open science practices other than publications (Mustajoki et al., 2021).

For further research, more case studies are needed to explore the potential of research information management to monitor FAIR projects and infrastructures at the local, regional, national and international levels. More empirical evidence needs to be presented on the real and specific impact of CRIS on the development of FAIR data repositories and other research infrastructures, with a particular focus on standardization. Finally, further development of CRIS data models and formats should focus on FAIR principles, especially findability and interoperability, in an explicit way. At the same time, the ethical and legal aspects of accessibility of CRIS data require further investigation to get a full picture of what it really means to apply the FAIR principles to research information management. While this is a research currently conducted by the European Open Science Cloud Task Force by means of both surveys, interviews, case studies and other activities, it can and should be supplemented with other independent and use-case based studies.

REFERENCES

Arefolov, A., Adam, L., Brown, S., Budovskaya, Y., Chen, C., Das, D., Farhy, C., Ferguson, R., Huang, H., Kanigel, K., Lu, C., Polesskaya, O., Staton, T., Tajhya, R., Whitley, M., Wong, J., Zeng, X., and McCreary, M. (2021). Implementation of the FAIR Data Principles for Exploratory Biomarker Data from Clinical Trials. *Data Intelligence*, 3(4):631–662.

Azeroual, O. and Schöpfel, J. (2021). Trustworthy or not? Research data on COVID-19 in data repositories. In Baker, D. and Ellis, L., editors, *Libraries, Digital In-*

formation, and COVID, Chandos Digital Information Review, pages 169–182. Chandos Publishing.

Azeroual, O., Schöpfel, J., Ivanovic, D., and Nikiforova, A. (2022). Combining Data Lake and Data Wrangling for Ensuring Data Quality in CRIS. In *In CRIS2022: 15th International Conference on Current Research Information Systems*, In CRIS2022: 15th International Conference on Current Research Information Systems.

Bryant, R., Jan Fransen, P. d. C., Helmstutler, B., and Scherer, D. (2021). Research Information Management in the United States: Part 1—Findings and Recommendations.

Commission, E. (2020a). Exploitation & Open Science in Horizon Europe.

Commission, E. (2020b). Facts and Figures for open research data.

Commission, E. (2020c). Research and innovation.

Commission, E. (2021). Open Science.

Corte-Real, N., Ruivo, P., and Oliveira, T. (2020). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, 57:103141.

Danowski, P., Ferus, A., Hikl, A.-L., McNeill, G., Miniberger, C., Reding, S., Zarka, T., and Zojer, M. (2020). "Empfehlung" für die weitere Vorgangsweise für das Open-Access-Monitoring. Deliverable des AT2OA-Teilprojekts TP1-B. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*, 73:278–284.

Donohue, T., Mennielli, M., and Wilcox, D. (2018). DuraSpace for FAIRness and Data Protection: DSpace and Fedora. In *CRIS2018: 14th International Conference on Current Research Information Systems (Umeå, June 13-16, 2018)*.

Engelman, A., Enkvist, C., and Pettersson, K. (2019). A FAIR archive based on the CERIF model. *Procedia Computer Science*, 146:190–200.

Europe, S. (2016). Science Europe Position Statement on Research Information Systems.

European Commission, D.-G. f. R. and Innovation (2019). Cost-benefit analysis for FAIR research data : cost of not having FAIR research data.

European Commission, D.-G. f. R. and Innovation (2021). *Towards a reform of the research assessment system : scoping report*. Publications Office.

FAIRsFAIR (2022). Practical solutions for the use of FAIR principles throughout the data life cycle.

Ferraris, A., Mazzoleni, A., Devalle, A., and Couturier, J. (2018). Big data analytics capabilities and knowledge management: impact on firm performance. Management Decision. *Management Decision*, 57:1923–1936.

Flores, J. R., Brodeur, J. J., Daniels, M. G., Nicholls, N., and Turnator, E. (2015). Libraries and the research data management landscape. *The process of discovery: The CLIR postdoctoral fellowship program and the future of the academy*, 2010:82–102.

FOSTER (2020). Assessing the FAIRness of data .

Guba, E. and Lincoln, Y. (1989). *Fourth generation evaluation*. SAGE Publications, Inc.

Hasselbring, W., Carr, L., Hettrick, S., Packer, H., and Tiropanis, T. (2020). From FAIR research data toward FAIR and open research software. *it - Information Technology*, 62(1):39–47.

Insider, L. (2022). Research Information definition.

Ivanović, D., Ivanović, L., and Layfield, C. (2019). FAIRness at University of Novi Sad - Discoverability of PhD research results for Non-Serbian scientific community–. *Procedia Computer Science*, 146:3–10.

Jetten, M. and Simons, E. (2019). Research data management incorporated in a Research Information Management system. A case study on archiving data sets and writing Data Management Plans at Radboud University, the Netherlands. In *EUNIS19: 25th EUNIS Annual Congress (June 5-7, 2019, NTNU, Trondheim, Norway)*.

Jetten, M., Simons, E., and Rijnders, J. (2018). The role of CRIS's in the research life cycle. A case study on implementing a FAIR RDM policy at Radboud University, the Netherlands. In Sicilia, M., Simons, E., Clements, A., de Castro, P., and Bergström, J., editors, *14th International Conference on Current Research Information Systems, CRIS 2018, FAIRness of Research Information, Umeå, Sweden, June 14-16, 2018*, volume 146 of *Procedia Computer Science*, pages 156–165. Elsevier.

Landi, A., Thompson, M., Giannuzzi, V., Bonifazi, F., Labastida, I., da Silva Santos, L. O. B., and Roos, M. (2020). The "A" of FAIR – As Open as Possible, as Closed as Necessary. *Data Intelligence*, 2(1-2):47–55.

Lindelöw, C. (2019). FAIR already? Principles of reusability and research output – evaluation at a national level. In *Spring 2019 euroCRIS Membership Meeting (CSC, Espoo-Helsinki, Finland, May 27-29, 2019)*.

Masuzzo, P. C. (2022). Open Science & FAIR data: infinite possibilities.

Mayer, G., Müller, W., Schork, K., Uszkoreit, J., Weidemann, A., Wittig, U., Rey, M., Quast, C., Felden, J., Glöckner, F. O., Lange, M., Arend, D., Beier, S., Junker, A., Scholz, U., Schüler, D., Kestler, H. A., Wibberg, D., Pühler, A., Twardziok, S., Eils, J., Eils, R., Hoffmann, S., Eisenacher, M., and Turewicz, M. (2021). Implementing FAIR data management within the German Network for Bioinformatics Infrastructure (de.NBI) exemplified by selected use cases. *Briefings in Bioinformatics*, 22(5).

Miniberger, C. and Reding, S. (2017). From data collection to FAIR use in CRIS. The case of University of Vienna. In *euroCRIS Strategic Membership Meeting Autumn 2017 (CVTI SR, Bratislava, Slovakia, Nov 20-22, 2017)*.

Mornati, S. (2019). Enhancing interoperability: the implementation of OpenAIRE Guidelines and COAR NGR Recommendations in CRIS/RIMS. In *OR19 Workshop on Repository/CRIS Interoperability*.

Mornati, S., Bollini, A., and Pascarelli, L. (2018). How to make research information FAIR: DSpace-CRIS

and best practices in open research information. In *CRIS2018: 14th International Conference on Current Research Information Systems (Umeå, June 13-16, 2018)*.

Mustajoki, H., Pölönen, J., Gregory, K., Ivanović, D., Brasse, V., Kesäniemi, J., Koivisto, E., and Pylvänäinen, E. (2021). Making FAIReR assessments possible. Final report of EOSC Co-Creation projects: "European overview of career merit systems" and "Vision for research data in research careers".

National Academy of Sciences (US), N. A. o. E. U., of Medicine (US) Committee on Ensuring the Utility, I., and of Research Data in a Digital Age, I. (2009). Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age.

Nguyen, G. T., Dlugolinsky, S., Bobák, M., Tran, V. D., García, Á. L., Heredia, I., Malík, P., and Hluchý, L. (2018). Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*, 52:77–124.

Nikiforova, A. (2020). Definition and evaluation of data quality: User-oriented data object-driven approach to data quality assessment. *Baltic Journal of Modern Computing*, 8:391–432.

Norway, U. (2016). NOR-CAM – A toolbox for recognition and rewards in academic careers.

O'Carroll, C., Vandevelde, K., McAllister, D., Metcalfe, J., Maas, K., Rentier, B., Kaunismaa, E., Esposito, F., and Cabello Valdes, C. (2017). *Evaluation of research careers fully acknowledging Open Science practices : rewards, incentives and/or recognition for researchers practicing Open Science*. Publication Office of the European Union.

OECD (2007). OECD Principles and Guidelines for Access to Research Data from Public Funding.

of Finnish Learned Societies, F. (2020). Good practice in researcher evaluation. Recommendation for the responsible evaluation of a researcher in Finland.

Paic, A. (2021). Open Science - Enabling Discovery in the Digital Age.

Society, M. P. (2010). Principles for the Handling of Research Data.

Tammaro, A. M., Matusiak, K. K., Sposito, F. A., and Casarosa, V. (2019). Data Curator's Roles and Responsibilities: An International Perspective. *Libri*, 69(2):89–104.

Tatum, C. and Brown, J. (2018). Principles and Pragmatics of "as open as possible": persistent identifiers as the interface between research information commons and closed systems. In *CRIS2018: 14th International Conference on Current Research Information Systems (Umeå, June 13-16, 2018)*.

Terheggen, C. and Simons, E. (2016). Towards a Model for FAIR Data Information Infrastructures. In *euroCRIS Strategic Membership Meeting Autumn 2016 (EKT, Athens, Greece, Nov 8-10, 2016)*.

UNESCO (2021). UNESCO Recommendation on Open Science.

Vancauwenbergh, S. (2021). Research Information Systems as Leverage for Open Science. In *EARMA Digital Conference 2021*.

Wachtler, T., Bauer, P., Denker, M., Grün, S., Hanke, M., Klein, J., Oeltze-Jafra, S., Ritter, P., Rotter, S., Scherberger, H., Stein, A., and Witte, O. W. (2021). NFDI-Neuro: building a community for neuroscience research data management in Germany. *Neuroforum*, 27(1):3–15.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9.

Wittenburg, P. (2019). Large Research Infrastructure Building using FAIR Digital Objects. In *Autumn 2019 euroCRIS Strategic Membership Meeting (WWU Münster, Germany, Nov 18-20, 2019)*.