A comprehensive comparison of automated FAIRness Evaluation Tools

Chang Sun[1 [0000−0001−8325−8848]], Vincent Emonet[1 [0000−0002−1501−1082]], and Michel Dumontier[1 [0000−0003−4727−9435]]

Institute of Data Science, Maastricht University, Maastricht, The Netherlands

Abstract. The FAIR Guiding Principles (Findable, Accessible, Interoperable, and Reusable) have been widely endorsed by the scientific community, funding agencies, and policymakers. However, the FAIR principles leave ample room for different implementations, and several groups have worked towards manual, semi-automatic, and automatic approaches to evaluate the FAIRness of digital objects. This study compares and contrasts three automated FAIRness evaluation tools namely F-UJI, the FAIR Evaluator, and FAIR Checker. We examine three aspects: 1) tool characteristics, 2) the evaluation metrics, and 3) metrics tests for three public datasets. We find significant differences in the evaluation results for tested resources, along with differences in the design, implementation, and documentation of the evaluation metrics and platforms. While automated tools do test a wide breadth of technical expectations of the FAIR principles, we put forward specific recommendations for their improved utility, transparency, and interpretability.

Keywords: FAIR Principles · Research Data Management · Automated Evaluation · FAIR Maturity Indicators

1 Introduction

The FAIR Guiding Principles (Findable, Accessible, Interoperable, Reusable) [1] have gained broad endorsement by funding agencies and political entities such as the European Commission, and are being implemented in research projects. However, the FAIR Principles are largely aspirational in nature and do not specify technical requirements that could be unambiguously evaluated [2,3]. A growing number of efforts have sought to evaluate the FAIRness of digital resources, albeit with different initial assumptions and challenges [4,5].

FAIRness evaluation tools range from questionnaires or checklists to automated tests based only on a provided Uniform Resource Identifier (URI) or Digital Object Identifier (DOI) [4]. The co-authors of FAIR principles published a framework for developing and implementing FAIR evaluation metrics, also called FAIR Maturity Indicators (MIs) [6,7]. These resulted in the development of an automated FAIR Evaluator [7] that evaluates the technical implementation of a resource's FAIRness against common implementation strategies. The FAIR Checker [8] is a recently developed resource that uses the reference FAIR

MIs but offers an alternate user interface and result representation. F-UJI [9] is an automated FAIR evaluation tool with its own metrics and scoring system. While these tools aim to systematically and objectively measure the FAIRness of the digital objects, they generate different FAIRness evaluation results owing to differences in strategies pertaining to information gathering, metric implementation, and scoring schemes.

We sought to compare and contrast three automated FAIRness evaluation tools (F-UJI, the FAIR Evaluator, and the FAIR checker) against their usability, evaluation metrics, and metric tests results. We generate evaluation results using three datasets from different data repositories. We discover the FAIRness evaluation tools have different coverage and emphases on the FAIR principles and apply different methods to discover and interpret the content of the digital objects. When assessing the comparable evaluation metrics, different tools may output conflicting results because of the different implementation of the metric tests. We analyze these observed differences and explore their likely bases. Our work is the first to offer a systematic evaluation of current automated FAIRness evaluators, with concrete suggestions for improving their quality and usability.

2 Materials and Methods

This study critically examines the functioning of the FAIR Evaluator, FAIR Checker, and F-UJI. These **FAIRness evaluation tools** are implemented as web applications that use web service APIs to execute a FAIRness evaluation and offer an interactive user interface through a web browser (Figure 1). These tools implement new or apply existing FAIRness **evaluation metrics**. Each metric has one or more compliance **metric tests** to determine if the digital object meets the requirements of the metric. These metric tests are the actual implementation of the evaluation metrics. Users invoke an evaluation by providing a valid URL or persistent identifier (PID) of the digital object's landing page. The tool executes a strategy to harvest relevant metadata on the URL (or its redirected URL) using a combination of content negotiation, embedded microdata, and HTTP meta rel links. The tools then test the harvested metadata, and tabulate whether and/or how they pass or fail the metric test(s). Finally, the tools present the results of the metric tests as an HTML web page that may otherwise be downloadable as a structured data file. We conducted a comprehensive comparison of the automated FAIRness evaluation tools focusing on 1) the characteristics of the evaluation tools, 2) the FAIRness evaluation metrics, and 3) the testing results using three public datasets.

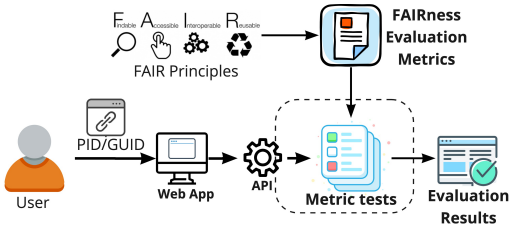

Fig. 1. A general workflow of FAIRness evaluation tools

2.1 Characteristics of the FAIRness evaluation tools

The automated evaluation tools are accessible via web applications and APIs. We extracted key features and specifications and reflected on the transparency (in terms of documentation) and extensibility of the tools. The elements such as the availability of source code, web application, the required inputs, the quality, and interpretation of the outputs are included.

2.2 FAIRness Evaluation Metrics and metric Tests

At the heart of automated FAIRness evaluation are programs that examine data resources for the presence and quality of particular characteristics. F-UJI implemented FAIRsFAIR Data Object Assessment Metrics [10], while the FAIR Evaluator implemented FAIRness Maturity Indicators (MIs) [6,7]. The FAIR Checker applies the same MIs as the FAIR Evaluator but implements a distinct web application with a different user interface. Our comparison on evaluation metrics lies between those used by F-UJI and the FAIR Evaluator/FAIR Checker. The FAIR Evaluator documented the measurements and procedures of metric tests through Nanopublication, which is readable for both machines and humans. The source code for the metric tests and evaluator application is available. F-UJI presents the names of their metric tests on the web application and published the source code of the tests. The log messages from both tools potentially indicate what properties are assessed in the (meta)data. We compare each metric/indicator from both tools and pair the metrics that are comparable to each other based on their descriptions, metric tests, and output log messages.

2.3 Tests on three public datasets

The last comparison focuses on the representation and interpretation of the evaluation results from F-UJI and the FAIR Evaluator. Three tested datasets in Table 1 are from PANGAEA [11], Kaggle [12], and Dutch Institute for Public Health and Environment (RIVM) [13]. PANGAEA assists the users to submit data following FAIR principles. All submitted data are quality checked and processed for machine readability. Kaggle recommends but is not mandatory for users to upload data with description and metadata. Unlike PANGAEA and Kaggle that are open to the general users to upload data, the RIVM data portal hosts data from governmental or authorized resources. Due to the current trend of COVID-19, CORD-19 and NL-Covid-19 were selected to evaluate their FAIRness. GeoData was included because of its descriptive metadata and quality-checked submission. The datasets are evaluated on F-UJI using its evaluation metrics v0.4 and software v1.3.5b and the FAIR Evaluator using its metric collection - "All Maturity Indicator Tests as of May 8, 2019".

Name	Host	Input for the assessment tools	Input type
GeoData	PANGAEA	10.1594/PANGAEA.908011	DOI
CORD-19	Kaggle	www.kaggle.com/allen-institute-for- ai/CORD-19-research -challenge	Metadata landing page
NL-Covid-19	RIVM	data.rivm.nl/meta/srv/eng/rdf.metadata.get?uuid=1c0fcd 57-1102-4620-9cfa-441e93ea5604&approved=true	Metadata in RDF

Table 1. Datasets for testing the automated FAIRness evaluation tools

3 Results

This section presents the results and analysis of comparing three evaluation tools. Comparison of the characteristics of the tools was performed with the FAIR Evaluator, FAIR Checker, and F-UJI, whereas a comparison of evaluation metrics was performed only with the FAIR Evaluator and F-UJI, as the FAIR Checker applies the same evaluation metrics as the FAIR Evaluator.

3.1 Comparison of characteristics of the evaluation tools

As table 2 shows, all tools are implemented as a standalone web application and API. Execution of the FAIRness evaluation is as follows: F-UJI requests a persistent identifier (PID) of the data or the URL of the dataset's landing page as input, while the FAIR Evaluator requests a global unique identifier (GUID) of the metadata. The following schemes are considered as PIDs by both tools: Handle, Persistent Uniform Resource Locator, Archival Resource Key, Permanent identifier for Web applications, and Digital Object Identifier. Both offer short descriptions about the input, while the FAIR Checker simply requests a URL or DOI without further explanation.

After the execution of the evaluation, each application presents the results differently. The FAIR Checker starts with a radar chart outlining the FAIRness scores along 5 axes (Findable, Accessible, Interoperable, Reusable, Total). The FAIR Checker does not provide detailed logs except the error messages. The FAIR evaluator presents the results of metric tests with the detailed application-level logs. The results are assigned with PIDs and stored in a persistent database where users can search, access, and download as a JSON-LD file. The F-UJI also provides application-level logs as feedback to the rationality of the test results. However, the logs are not as detailed as the FAIR Evaluator. The results from F-UJI can be downloaded as a JSON file. F-UJI and the FAIR Evaluator are both based on APIs to make their FAIRness evaluation services accessible.

	F-UJI	**FAIR Evaluator**	**FAIR Checker**
Web application	[www.f-uji.net](v1.3.5b)	[w3id.org/AmIFAIR](v0.3.1)	[fair-checker.france-bioinformatique.fr] (v0.1)
Requested input	PID,URL of dataset	GUID of the metadata	URL,DOI
Results export	JSON	JSON-LD	Not available
Output	Application-level logs	Application-level logs	Error logs
Metrics	[10]	[7]	[7]
Source code	[github.com/pangaea-data-publisher/fuji]	[github.com/FAIRMetrics/Metrics]	[github.com/IFB-ElixirFr/fair-checker]
Language	Python	Ruby	Python
Associated project/group	*FAIRisFAIR*	*FAIRSharing* *FAIR Metrics Group*	*French Institute for Bioinformatics*

Table 2. Comparison of FAIRness evaluation tools

3.2 FAIRness Evaluation Metrics

The latest evaluation metrics from F-UJI include 17 metrics to address the FAIR principles with the exception of A1.1, A1.2, and I2 (open protocol, authentication and authorization, FAIR vocabularies). The metrics are documented with

identifiers, descriptions, requirements, and other elements [10]. The FAIR Evaluator used a community-driven approach to create 15 Maturity Indicators (MIs) covering the FAIR principles except for R1.2 and R1.3 (detailed provenance, community standards). The MIs are documented in an open authoring framework (`https://github.com/FAIRMetrics/Metrics`) where the community can customize and create domain-relevant, community-specific MIs. Table 3 shows the comparison of F-UJI evaluation metrics v0.4 and the metric collection - "All Maturity Indicator Tests as of May 8, 2019" from the FAIR Evaluator corresponding to the FAIR principle. The comparable metrics are paired in the table.

F-UJI has two metric tests on data and three tests on metadata to assess the **findability**, while the FAIR Evaluator has six tests on metadata. The FAIR Evaluator requires PID for both metadata and data, while F-UJI only requires for the data. Two tools both check if the metadata is structured using JSON-LD or RDFa. However, the FAIR Evaluator requires metadata to be grounded in shared vocabularies using a resolvable namespace. F-UJI checks the predefined core elements in the metadata, such as title, description. and license.

Two tools evaluate the **accessibility** by assessing communication protocols for retrieving (meta)data, ensuring the (meta)data can be accessed through a standard protocol. The FAIR Evaluator requires authentication implementation on the data and authorizations on metadata, while F-UJI only requires metadata authorizations. The metadata persistence is discussed by both tools, but F-UJI does not implement it in their tool. The argument is that programmatic evaluation of the metadata preservation can only be tested if the object is deleted or replaced [10]. However, the FAIR Evaluator measures the metadata persistence by looking for a persistence policy key or predicate in the metadata.

To evaluate the **interoperability**, the FAIR Evaluator tests whether the metadata and data are structured and represented using ontology terms. F-UJI only focuses on the structure of metadata. Compared to F-UJI, the FAIR Evaluator has extensive measurements on both metadata and data to evaluate the interoperability. In the evaluation of **reusability**, F-UJI has more comprehensive measurements than the FAIR Evaluator. The FAIR Evaluator checks if license information is included in the metadata. By contrast, F-UJI setup four tests for metadata and one test for data to check the richness, licenses, and provenance of metadata and applied community-standards in metadata and data.

3.3 Compare the test results on public datasets

The evaluation results of three datasets are shown in Table 4. The full results are accessible on `https://doi.org/10.5281/zenodo.5539823`. Geodata scored perfect on all the metrics from F-UJI, but 17 out of 22 from the FAIR Evaluator. 4 out of 5 failed tests in the FAIR Evaluator assessed aspects that are not listed in F-UJI. The test on the persistence of the data identifier (F1-01D, F1-02D, MI_F1B) had different results from F-UJI and the FAIR Evaluator. Additionally, if qualified outward references in metadata (I3-01M, MI_I3A) and licenses in metadata (R1.1-01M, MI_R1.1) also had different results from two evaluators on the tested datasets. These differences are examined further in the Discussion.

Table 3: Comparison of FAIRness evaluation metrics from all tools.

FAIR Metrics	F-UJI		FAIR Evaluator/FAIR Checker	
FINDABLE				
	ID(FsF-)	Name	ID(Gen_)	Name
F1: (meta)data are assigned a globally unique and persistent identifier.	-	-	MI_F1A	(**Metadata**) Identifier uniqueness
	-	-	MI_F1B	(**Metadata**) Identifier persistence
	F1-01D	**Data** is assigned a globally unique identifier.	MI_F1A	(**Data**) Identifier uniqueness
	F1-02D	**Data** is assigned a persistent identifier.	MI_F1B	(**Data**) Identifier Persistence
F2: data are described with rich metadata.	-	-	MI_F2A	Structured **Metadata**
	F2-01M	**Metadata** includes descriptive core elements to support findability.	MI_F2B	Grounded **Metadata**
F3: metadata clearly and explicitly include the identifier of the data they describe.	-	-	MI_F3	Use of (**metadata**) GUIDs in **metadata**
	F3-01M	**Metadata** includes the identifier of the data it describes.	MI_F3	Use of (**data**) GUIDs in **metadata**
F4: (meta)data are registered or indexed in a searchable resource.	F4-01M	**Metadata** can be retrieved programmatically.	MI_F4	(**Metadata**) Searchable in major search engines
ACCESSIBLE				
A.1 (meta)data are retrievable by their identifier using a standardized communications protocol.	A1-01M	**Metadata** contains the access level and access conditions of the data.	-	-
	A1-02M	**Metadata** is accessible through a standardized communication protocol.	-	-
	A1-03D	**Data** is accessible through a standardized communication protocol.	-	-
A1.1 the protocol is open, free, and universally implementable.	-	-	MI_A1.1	Uses open free protocol for **metadata** retrieval
	-	–	MI_A1.1	Uses open free protocol for **data** retrieval
A1.2 the protocol allows for an authentication and authorization procedure.	-	-	MI_A1.2	**Metadata** authentication and authorization
	-	-	MI_A1.2	**Data** authentication and authorization
A2. metadata are accessible, even when the data are no longer available.	A2-01M	**Metadata** remains available, even if the data is no longer available. (This metric is disabled in F-UJI tool.)	MI_FA2	**Metadata** Persistence
INTEROPERABLE				
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	I1-01M	**Metadata** is represented using a formal knowledge representation language.	MI_I1A	**Metadata** Knowledge Representation Language (weak)
			MI_I1B	**Metadata** Knowledge Representation Language (strong)
	I1-02M	**Metadata** uses semantic resources	-	**-**
	-	-	MI_I1A	**Data** Knowledge Representation Language (weak)
	-	-	MI_I1B	**Data** Knowledge Representation Language (strong)
I2. (meta)data use vocabularies that follow FAIR principles.	-	-	MI_I2A	**Metadata** uses FAIR vocabularies (weak)
	-	-	MI_I2B	**Metadata** uses FAIR vocabularies (strong)
I3. (meta)data include qualified references to other (meta)data.	I3-01M	**Metadata** includes links between data and related entities.	MI_I3A	**Metadata** contains qualified outward references
REUSABLE				
R1. meta(data) are richly described with accurate and relevant attributes.	R1-01MD	**Metadata** specifies the content of the data.	-	-
R1.1. (meta)data are released with a clear and accessible data usage license.	R1.1-01M	**Metadata** includes license information.	MI_R1.1	**Metadata** Includes License (weak)
	-	-	MI_R1.1	**Metadata** Includes License (strong)
R1.2. (meta)data are associated with detailed provenance.	R1.2-01M	**Metadata** includes provenance information about data creation or generation.	-	-
R1.3. (meta)data meet domain-relevant community standards.	R1.3-01M	**Metadata** follows a standard recommended by the target research community of the data.	-	-
	R1.3-02D	**Data** is available in a file format recommended by the target research community.	-	-

F-UJI	FE	GeoData		CORD-19		NL-Covid-19	
		F-UJI	FE	F-UJI	FE	F-UJI	FE
F1-01D	MI_F1B	✓	✗	✓	✗	✓	✗
F1-02D		✓		✗		✗	
-	MI_F3	-	✓	-	✗	-	✗
F4-01M	MI_F4	✓	✓	✓	✓	✗	✗
A1-01M	-	✓	-	✓	-	✗	-
I1-02M	-	✓	-	✗	-	✓	-
-	MI_I2B	-	✗	-	✗	-	✓
I3-01M	MI_I3A	✓	✓	✓	✓	✗	✓
R1.1-01M	MI_R1.1	✓	✓	✓	✗	✓	✗
R1.3-01M	-	✓	-	✗	-	✓	-
R1.3-02D	-	✓	-	✗	-	✓	-
Passed/total tests:		16/16	17/22	12/16	13/22	11/16	13/22

Table 4. Selected results of evaluating datasets using F-UJI and FAIR Evaluator (FE).

CORD-19 failed 4 tests in F-UJI and 9 tests in the FAIR Evaluator mostly in the evaluation of the I and R. The poor quality of metadata of CORD-19 causes further failures in the other tests in both evaluation tools such as the persistence of the metadata identifier (F1-02D), metadata includes license (MI_R1.1). NL-Covid-19 had the lower FAIRness score from F-UJI among the three datasets (11 out of 16) and 13 out of 22 in the FAIR Evaluator. It has the same issue of the quality of metadata as the second dataset, but outperformed in the knowledge representation in data. Neither F-UJI nor the FAIR Evaluator detected the license information in the metadata of NL-Covid-19, but the metadata clearly indicates NL-Covid-19 comply with a valid license.

4 Discussion

This study compares three automated FAIRness evaluation tools on the characteristics of the tools, the evaluations metrics and metric tests, and the results of evaluating 3 datasets. The outstanding feature of the FAIR Evaluator is the community-driven framework that can be readily customized, by creating and publishing an individual or collection of Maturity Indicators (MIs) to meet the domain-related and community-defined requirements of being FAIR. The MIs and metric tests that are registered by one community are discovered and can be grouped to maximize the reusability across communities. All published MIs and conducted FAIRness evaluations are stored in a persistent database and can be browsed and accessed by the public. F-UJI visualizes the evaluation results and represents the output with better aesthetics. The source code is publicly available in Python, and well-structured for each metric test. The FAIR Checker uses the FAIR Evaluator API to perform the resource assessment, and has a more aesthetic presentation including recommendations to the failed tests, but does not allow the selection of particular metrics tests or collections, and does not offer the detailed output.

4.1 Transparency of the FAIRness evaluation tools

All the evaluation tools suffer from some aspect of clarity and transparency. F-UJI's source code is open and each evaluation metric is described in an accompanying article. However, without technical specifications of the application

functioning, it is challenging to scan the whole code repository to learn how each metric was technically implemented. It is unclear what properties are assessed and how to improve the FAIRness of the objects. F-UJI gives a FAIRness score and a maturity score to the digital objects based on the metric tests. But it is lacks of description of how these tests are scored and how the scores are operated.

The FAIR Evaluator published its MIs and metric tests in a public Git repository. The web application of the FAIR Evaluator presents detailed log messages which potentially indicate what has been tested and what caused the test failure. However, the users still suffer from the insufficient transparency of the implementation. The FAIR Checker only generates the final test results (pass or not pass) without further explanations.

4.2 Differences among the tools

In the comparison of the evaluation metrics, F-UJI has comprehensive metrics for Reusability, while the FAIR Evaluator focuses on the Interoperability. The evaluation results from three datasets reveal more significant differences between F-UJI and the FAIR Evaluator which result in conflicting results for the same metric. We summarize the following three key reasons.

1) Different understanding of certain concepts. When evaluating Geodata, F-UJI recognizes the DOI (10.1594/PANGAEA.908011) as the data identifier. F-UJI considers DOI as a persistent identifier (PID) and determines that Geodata has a valid PID for the data. However, the FAIR Evaluator defined the DOI as the identifier for the metadata instead of the data. The data download URL is recognized as the data identifier by the FAIR Evaluator. Thus, F-UJI and the FAIR Evaluator have different understanding and definitions of data and metadata identifiers, which result in differing test results.

2) Different depth of information extraction. F-UJI and the FAIR Evaluator gave conflicting results in determining whether metadata contained license information in CORD-19. F-UJI reported that license information was found, while the FAIR Evaluator did not recognize the license. From the output logs, two tools were both able to capture "Other (specified in description)" as the license information in the metadata. However, the FAIR Evaluator failed the "metadata contains licenses" test because the FAIR Evaluator requires a valid value of a license property (i.e. a URL). F-UJI passed the test but the given information for the license property is not recognized as a valid license.

When evaluating NL-Covid-19, F-UJI and the FAIR Evaluator both failed the test on "metadata contains licenses". However, the license information is clearly included in the metadata of NL-Covid-19 (RDF format) with two statements. F-UJI is unable to find the license predicate in the metadata, while the FAIR Evaluator found the license predicate but only processed the first statement - "Geen beperkingen" as an invalid license. Unfortunately, the FAIR Evaluator did not continue to process the second statement which contains the valid license information. In this case, neither F-UJI nor the FAIR Evaluator are able to find the valid licenses in the metadata of NL-Covid-19.

3) Different implementations of the metrics. F-UJI and the FAIR Evaluator both examine whether the relationships within (meta)data between local and third-party data are explicitly indicated in the metadata (I2-01M, MI_I3A). In the evaluation of NL-Covid-19, the FAIR Evaluator passed the test by discovering 26 out of 45 triples in the linked metadata pointed to resources that are hosted by a third party. F-UJI did not pass this test because it could not exact any related resources from the metadata. The conflicting test outcome results from the different implementation of recognizing the relationship between the local and third-party data. F-UJI requires the relationship properties that specify the relation between data and its related entities have to be explicit in the metadata and use pre-defined metadata schemas (e.g., "RelatedIdentifier" and "RelationType" in DataCite Metadata Schema). Compared to F-UJI, the FAIR Evaluator has a broader requirement for acceptable qualified relationship properties by including numerous ontologies which include richer relationships.

4.3 Potential limitations

This study has several limitations. The comparison of evaluation metrics between F-UJI and the FAIR Evaluator is based on the description of each metric, metric tests, and log messages. We did not conduct a detailed examination of their implementation. The FAIR Evaluator published technical specifications for each Maturity Indicator and its metric tests as well as the source code of implementation. F-UJI shares its source code and descriptions of the metrics in an article. However, metric tests and their implementation have not been sufficiently discussed. A possible solution for comparing the evaluation tools on the implementation level is to scan their entire source code. However, this will require an extensive effort by experts in both Ruby and Python to conduct this task.

The discovery of the evaluation results from the three tools is possibly limited by our selection of the datasets. To increase the objectiveness of the evaluation, more representative datasets from various data repositories are required to test the different evaluation tools. A potential solution could be to construct a framework that evaluates and compares the FAIRness evaluation tools in an automatic and systematic manner. The framework executes the evaluation tools on a set of standard benchmarking datasets, examines what properties are being tested, and generates evaluation results automatically. This automated evaluation framework will overcome the qualitative nature of the current study and the shortcomings of requiring substantial manual effort and proning to the errors. Finally, the evaluation tools in this study are all under active development. The evaluation metrics and implementations of metric tests in these tools can probably be changed over time.

5 Conclusion

This study conducted a comprehensive comparison among three automated FAIRness evaluation tools (F-UJI, the FAIR Evaluator, and the FAIR checker)

covering the tool characteristics, evaluation metrics and metric tests, and evaluation results of three public datasets. Our work revealed differences among the tools and offers insights into how these may lead to different evaluation results. Finally, we presented the common issues shared by all FAIRness evaluation tools and discussed the advantages and limitations of each tool. We note the tools are under active development and are subject to change. Future work could focus on standardized benchmarks to critically evaluate the functioning of these and future FAIRness evaluation tools.

References

1. M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.*, "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.
2. B. Mons, C. Neylon, J. Velterop, M. Dumontier, L. O. B. da Silva Santos, and M. D. Wilkinson, "Cloudy, increasingly fair; revisiting the fair data guiding principles for the european open science cloud," *Information Services & Use*, vol. 37, no. 1, pp. 49–56, 2017.
3. A. Ammar, S. Bonaretti, L. Winckers, J. Quik, M. Bakker, D. Maier, I. Lynch, J. van Rijn, and E. Willighagen, "A semi-automated workflow for fair maturity indicators in the life sciences," *Nanomaterials*, vol. 10, no. 10, p. 2068, 2020.
4. R. de Miranda Azevedo and M. Dumontier, "Considerations for the conduction and interpretation of fairness evaluations," *Data Intelligence*, vol. 2, no. 1-2, pp. 285–292, 2020.
5. "Fairassist - discover resources to measure and improve fairness." `https://fairassist.org/`, 2021. Accessed: 2021-09-28.
6. M. D. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. B. da Silva Santos, and M. Dumontier, "A design framework and exemplar metrics for fairness," *Scientific data*, vol. 5, no. 1, pp. 1–4, 2018.
7. M. D. Wilkinson, M. Dumontier, S.-A. Sansone, L. O. B. da Silva Santos, M. Prieto, D. Batista, P. McQuilton, T. Kuhn, P. Rocca-Serra, M. Crosas, *et al.*, "Evaluating fair maturity through a scalable, automated, community-governed framework," *Scientific data*, vol. 6, no. 1, pp. 1–12, 2019.
8. "Fair-checker." `https://fair-checker.france-bioinformatique.fr/base_metrics`, 2021. Accessed: 2021-09-28.
9. A. Devaraju, M. Mokrane, L. Cepinskas, R. Huber, P. Herterich, J. de Vries, V. Akerman, H. L'Hours, J. Davidson, and M. Diepenbroek, "From conceptualization to implementation: Fair assessment of research data objects," *Data Science Journal*, vol. 20, no. 1, 2021.
10. A. Devaraju, R. Huber, M. Mokrane, P. Herterich, L. Cepinskas, J. de Vries, H. L'Hours, J. Davidson, and A. White, "Fairsfair data object assessment metrics," *Zenodo, Jul*, vol. 10, 2020.
11. "Pangaea - data publisher for earth & environmental science." `https://www.pangaea.de/`, 2021. Accessed: 2021-09-22.
12. "Kaggle: Your machine learning and data science community." `https://www.kaggle.com/`, 2021. Accessed: 2021-09-22.
13. "Dutch institute for public health and environment data portal." `https://data.rivm.nl/meta/srv/eng/catalog.search#/home`, 2021. Accessed: 2021-09-22.