euFAIR: A Digital Tool for Assessing the FAIR Principles

Matteo Lia[1,2](✉), Davide Damiano Colella[1], Antonella Longo[1,3], and Marco Zappatore[1,3]

[1] Department of Innovation Engineering, University of Salento, via Monteroni sn, 73100 Lecce, Italy
{matteo.lia,antonella.longo, marcosalvatore.zappatore}@unisalento.it, matteo.lia@parsec326.it, davidedamiano.colella@studenti.unisalento.it

[2] Parsec 3.26 Srl, via del Platano 7, 73020 Cavallino – Castromediano, Lecce, Italy

[3] Italian Research Center on High Performance Computing, Big Data and Quantum Computing (ICSC), Bologna, Italy

Abstract. Over the last decade, the importance of FAIR principles as a reference for data reusability and openness has increased constantly. Various tools for FAIR data assessment exist, including manual approaches like the FAIR Data Self-Assessment Tool (SAT) and automated tools as the FAIR Evaluation Services and Generic Automatic Tool (GAT). However, subjectivity in manual assessment and limited guidance in automated tools represent significant limitations. In such a context, the Italian "*Piano Triennale per l'Informatica nella Pubblica Amministrazione*" has laid the foundations for open data practices in the Italian Public Administration (PA) since 2017. In this work, we propose a tool called euFAIR for the automatic assessment of dataset FAIRness in the Italian PA. The tool incorporates European Data Quality Guidelines for a more comprehensive dataset evaluation and implements a set of specific ad-hoc designed metrics. With this tool, we aim at improving data quality and sharing, as well as the adherence to community standards in the Italian PA. To validate our results, we compared euFAIR with SAT and GAT, highlighting significant differences.

Keywords: FAIR · Tool · Assessment · Public Administration

1 Introduction

The Italian "*Piano Triennale per l'Informatica nella Pubblica Amministrazione*" (Three-Year Plan for Information Technology in Public Administration) [1] aligns with European and Italian regulatory developments and sets the groundwork for nationwide management of open data. Since its 2017 inception and subsequent updates, the plan enhances transparency of the Italian Public Administration (PA), facilitates citizens' access to information, and fosters new products and services by businesses. All PAs' open data is

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
T. P. Sales et al. (Eds.): ER 2023 Workshops, LNCS 14319, pp. 49–58, 2023.
https://doi.org/10.1007/978-3-031-47112-4_5

consolidated in a dedicated Web portal [2], connected to its European counterpart [3]. However, for effective and reliable data sharing and reuse, the datasets in the portal must adhere to the FAIR (Findable, Accessible, Interoperable and Reusable) principles [4] and be enriched accordingly. The FAIR principles are intended to be a guideline to enhance data reusability with specific emphasis on automation (e.g., machine readable formats). One way to support the provision of “more FAIR” data, is the development of tools to assess or even score the degree to which a dataset complies with the FAIR principles. In recent times, the scientific community has made significant strides in applying and evaluating the FAIR principles on Digital Objects (DO), encompassing publications, datasets, and research software. As a result, there are several openly available automated and manual FAIR assessment services, including FAIR Evaluation Services and FAIR Data Self-Assessment Tool (SAT). While the versatile manual FAIR assessment tools totally depend on the user, the automated FAIR assessment tools search for known structures that could be used in the metadata. This makes the automatic tools compatible with most resources, but its precision can vary depending on the resources. Automatic tools allow higher precision at the expense of less coverage, by limiting the tool’s context of use. Nevertheless, this does not represent a deficiency of the tools of this category. It is a feature deriving from the real nature of FAIRness, since *«the final FAIR Principle speaks directly to the fact that “FAIR” will have different requirements for different communities»* [5].

In this scenario, we developed euFAIR, an automatic tool for assessing the FAIRness level of the Italian PA datasets, thanks to ad-hoc metrics. The tool goes beyond the assessment based solely on the FAIR principles and also incorporates the European Data Quality Guidelines (DQG) [6]. By embracing this all-encompassing approach, our primary objective is to enhance overall data quality, facilitate data sharing, and encourage adherence to community standards within the Italian PAs. In order to evaluate euFAIR efficacy, we conducted a thorough comparison between euFAIR, and SAT and GAT methodologies. The outcome of this comparison exposed noteworthy disparities in the evaluation results, signifying the potential superiority and effectiveness of our proposed euFAIR framework. The paper is organized as follows: the second section will focus on the related work data FAIRness evaluation tools. The proposed euFAIR tool will be described in the third section. The comparison evaluation is described in the fourth section, while conclusions and future work will be drawn in the fifth section.

2 Related Work

Over the years, a number of tools have been developed to evaluate data FAIRness, mainly based on manual assessment, involving methods such as case studies, checklists, and templates. Other approaches leverage Web scraping and a series of tests to automatically do the assessment. In this section, a subset of those tools will be examined, as they will be compared with the proposed euFAIR tool.

The Australian Research Data Commons (ARDC) actively promotes the adoption of the FAIR principles, providing valuable resources to the global research community. Among these resources is a FAIR Data Self-Assessment Tool (SAT) [7], available on the ARDC website. The tool aims to help researchers gauge the “FAIRness” of their data and encourage discussions about improving data practices. The assessment tool is a survey with 12 questions, aligned with the four FAIR principles, to quantify compliance with each category. Some questions require single-choice responses, while others use a Yes/No format. The tool evaluates answers for consistency, linking each question to a specific FAIR principle. As researchers respond, the corresponding section's bar fills up, showing the strength of their answers in relation to the FAIR principle. The "Total across F.A.I.R" bar aggregates scores from all sections, providing an overall "FAIRness" score for the data.

FAIR Evaluation Services is an open-source framework for conducting and developing maturity indicator tests [8], but its documentation focuses primarily on creating new indicators rather than providing comprehensive guidance on metrics and results. To use the web interface for automated assessment, users must provide a GUID, title, and ORCID. The output offers an overview of passed and failed tests for each indicator, but concerns arise regarding data retention policies as all tests are stored, leading to a substantial database of evaluations [9]. The tool lacks clear guidance on dataset FAIRness improvement, and understanding the detailed output requires technical expertise. Ongoing development efforts are being made to implement further metrics, but advice mechanisms for enhancing dataset FAIRness are currently absent [10, 11].

3 The euFAIR Tool

To cope with the limitations briefly sketched in Sect. 2, the euFAIR[1] tool is proposed: it has two complementary purposes. First, to represent a suitable component for the Italian open data portal supporting the assessment *after* the datasets are harvested, to identify their weaknesses and limitations. Second, to allow PAs assessing their datasets *before* publication, so to provide stronger and more community-compliant data.

3.1 Design Methodology

The methodology followed for the definition of the requirements and the whole evaluation is the Design Science in Information Systems Research [12]. The design of the tool was inspired by the gaps found in SAT and GAT, having in mind the main user which is the open data manager of a PA. The seven research steps, according to the referenced paradigm [12] have been followed; for the lack of space, they are not described here. Here we focus only on the euFAIR functional and technical description.

3.2 Data Model

The Italian open data portal is a key component to understand the data lifecycle of Italian PA data and to analyze their FAIRness, as it provides a single point of access to datasets that can be further shared in multiple formats.

The most comprehensive path that data can follow is when: 1) they are loaded onto the proprietary portal of the Italian PA that produces them; 2) (if agreements have been

[1] The tool is available on this repository: https://github.com/datalab-unisalento/euFAIR-tool/.

made with the Italian portal) they are uploaded onto it; 3) they are (generally) loaded onto the European portal. However, not all the datasets follow these procedures, not all the PAs are yet to be harvested from the Italian portal and not all the Italian PA datasets are available in the European portal.

Therefore, we modeled a dedicated database where every dataset as an ID associated with the Italian portal and some datasets have an ID for the European portal. The proposed database behind the euFAIR tool (see Fig. 1) is designed to store the (meta) data evaluations of the datasets. To do so the database is structured as follows. The *Dataset* table holds all the IDs of the dataset that we have tested or intend to test; it also keeps track of their Italian and EU IDs, as well as the dataset name and description, the timestamp of last update/check, and the accrual periodicity. The *Holder* table currently holds only the ID of the dataset holder, but it is intended to be further extended. The *Data Evaluation* and *Metadata Evaluation* tables, deriving from weak entities (see Fig. 1), strictly depend on the Dataset they refer to. Other attributes are: the metadata profile used to make the evaluation, the date, and the version of the tool. For the metadata we also keep track of the portals (origin, European, or) that produce them.

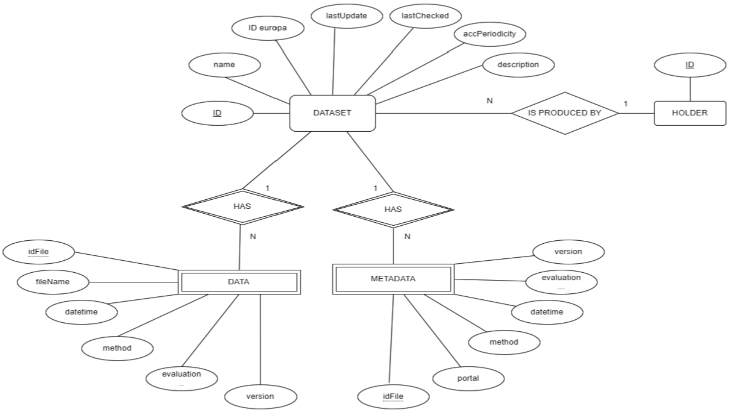


Fig. 1. Entity Relationship model of euFAIR Database

Since euFAIR uses the database to obtain the list of datasets that need to be evaluated or whose evaluation needs an update, we populated the database via the CKAN API method `current_package_list_with_resources`[2], thus retrieving all the datasets metadata from the portal, which we extract some of the dataset attributes from: ID, name, holder, last_update and accrual periodicity. To complete the data with the ID of the dataset in the European portal we need to take into consideration some details. The European ID is created based on the agreements made with the portal that offers the datasets when the harvesting has been programmed. There is no unique way to calculate the ID based on its original one. However, in our case the agreements made seem to suggest that the European ID is the name of the dataset.

[2] CKAN: API Guide, https://docs.ckan.org/en/latest/api/index.html, last accessed 2023/04/22.

In case of a double entry with the same ID a suffix is given to the dataset followed by an incremental number starting with 1. For example, if a dataset with the name '*Global_Temperature_Trends'* is being harvested from the Italian portal (Ip) but on the European portal (EUp) a dataset with that ID is already present, then the dataset is uploaded with the ID '*Global_Temperature_Trends* 1'.

To get this information we obtain a list of all the IDs in the Ip catalog present in the EU portal through the hub-search API using the list dataset method. Every ID in the list is then searched on our database and associated with the dataset with the given name through query. If an ID has the aforementioned suffix, it is removed before searching.

3.3 Evaluation Metrics

The core of the euFAIR tool is the series of metrics designed to specifically evaluate the FAIRness of the data in the context of the Italian PA and take into account the current legislation (that encourages the FAIRness of the data and mandates its openness). Every FAIR metric is evaluated through one or more tests meant to give a percentage of accomplishment. For every FAIR principle, a series of tests is then performed to check the dataset compliance with the EU Data Quality Guidelines.

In this section, we describe the test for the FAIR F2 principle: *"data are described with rich metadata"*. The best way to determine how rich our metadata are is to take into account the guidelines provided by the Ip and EUp, to refer to the metadata profiling system provided both by the European Commission (DCAT-AP [13]) and the AGID (DCAT-AP_IT [14]). These profiles give a set of attributes to be implemented in the metadata, and each of them is expressed as either mandatory, recommended, or optional. The euFAIR tool support multiple profiling systems. Similarly, the tool can be configured according to different and country-specific profiling systems. At present, both the previously mentioned profiles are implemented and a "merged" profile is automatically created (a profile that for each field takes the stricter requirement).

Implementation. In order to achieve a valid evaluation, we have devised a scoring mechanism that assigns a maximum score of 1, 0.5, or 0.2 points to each field, depending on whether it is mandatory, recommended, or optional, respectively. The total achievable score in the test is the sum of the max scores for all the required fields. Additionally, an actual score indicator is considered, starting at 0 and increasing with each well-evaluated field. When a field is successfully implemented, the actual score grows by the value of that particular field; otherwise, no points are added. If a field contains sub-fields, the evaluation begins with them: 1) we first calculate a maximum score by summing up the max points obtainable for each sub-field while considering whether they are mandatory; 2) we then evaluate the actual score for the field as previously described; 3) dividing the actual score by the max score yields the 'percentage of completeness' for that field.

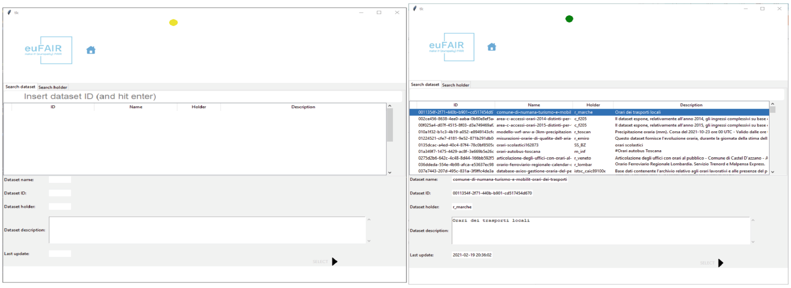
Taking into consideration both the max score of the field and its "percentage of completeness", we can determine the field point by multiplying the two values. For example: a mandatory field (max point 1.0) is expected to implement 2 mandatory fields, 1 recommended, 3 optional. The max score obtainable considering the sub fields is $maxScore = *1 + 1 * 0.5 + 3 * 0.2 = 3.1$. It follows that only one mandatory field is however implemented, 1 recommended and 2 optional. The actual score of the field will therefore be: $actualScore = 1 * 10 + 1 * 0.5 + 2 * 0.2 = 1.9$. The percentage of completeness will be: $percCompl = 61\%$. To the overall actual score will then be added 1*0.61 (i.e., max point of the field times the percentage of completeness of the field considering the subfields implemented). This procedure is repeated in case of sub-sub-fields, starting with the evaluation of the innermost section, and going up. Once all the fields are evaluated the test gives the $finalScore = actualScore/maxScore$.

Now, let us shift our attention to the European Data Quality Guidelines Test, which aligns strongly with the project's European context. Several tests are utilized to evaluate dataset FAIRness, following the EU Data Quality Guidelines (DQG). These guidelines encompass relevant metrics for assessing the FAIRness of (meta)data and offer specific formats for distributions. In this work, we focus on describing the DQG findability test as follows. Regarding findability, the DQG recommends minimizing NULL values in the metadata and emphasizes the significance of specific fields (e.g., *title*, *description*, *keywords*, *categories*, *spatial*, and t*emporal*) to ensure easy discoverability of the dataset. To implement this, the tool calculates the number of fields and empty fields, deriving a percentage of completeness. It then checks the specified fields for their implementation, assigning a maximum score of 6, with 1 point for each checked aspect. For each specific field, a point is awarded for filled-in fields, and a score between 0 and 1 (based on the completeness percentage) for the number of NULL values.

3.4 Tool Architecture

The tool is divided into two components, corresponding to the dataset update and to its assessment. The *Update* component allows to check for updates in the dataset and to proceed with new evaluations for those requiring it. The user is presented with 3 functions. First, searching for new datasets by consulting the Ip to check if new datasets have been produced; and then adding the new datasets to the database. Second, checking for new properties in the dataset. Third, updating the FAIRness evaluations of the dataset. The tool also shows how many datasets need update, the current evaluation stage (i.e., the portal in which the evaluation is being made and the method). More specifically, the tool starts by retrieving the list of the datasets that need to be updated and exploits the dataset attributes in the database by looking for all the datasets: 1) for which the accrual periodicity is passed, 2) that are not yet evaluated, 3) that have no declared accrual periodicity. Ordering by accrual periodicity allows us not to give priority to those datasets that might not require updating and/or have been updated already. Once the list is retrieved the tool starts the assessment and scoring, for every available portal (native, Ip, EUp) and for every available method (DCAT_AP, DCAT_AP-IT, merged). For every dataset a call is made through CKAN API to retrieve the metadata from the Ip and metadata evaluation tests are performed. The same thing is done through hub-search API for the EUp. Finally, the webpage of the PA that shared the dataset is retrieved and the assessment is made on the native portal. The scenario is highly unpredictable due to the numerous data providers contributing to the portal and their varied data sharing methods. Initially, the tool attempts to retrieve metadata using the common CKAN API, subjecting it to the same tests as for Ip. Otherwise, Web scraping is performed, and generic tests are conducted on the scraped metadata to identify known fields and properties. However, the reliability of these tests is reduced due to the lack of knowledge about the website's structure. If data distribution is found, available files are listed, grouped to identify similar files with different extensions, and the best format is evaluated. FAIR principle tests are format-independent, while DQG-based tests are format-specific (currently supporting only CSV but with plans to expand). Lastly, the dataset or distribution's accrual periodicity is checked for the EUp, Ip, and native portal metadata.

The second component allows to interact with the database. The euFAIR UI offers the user three options. First, the *database query module* allows the user to choose between a single dataset evaluation or a Holder evaluation. When a selection is confirmed the frame with the evaluations graph is shown. The database status is also indicated (i.e., red for preliminary checks, yellow for dataset update, and green for database updated). When the user searches for a dataset ID or name, all its features are shown and the retrieved results are then used to calculate its avg, min and max scores (Fig. 2).



ID	Name	Holder	Description
0011354f-2f71-440b-b901-cd517454d6	comune-di-numana-turismo-e-mobil	r_marche	Orari dei trasporti locali
002ce456-8638-4ee0-aaba-0b60e8ef5a	area-c-accessi-orari-2014-distinti-per-	c_f205	Il dataset espone, relativamente all'anno 2014, gli ingressi complessivi su base
00f025a4-d07f-4515-8f03-d3e749469af	area-c-accessi-orari-2015-distinti-per-	c_f205	Il dataset espone, relativamente all'anno 2015, gli ingressi complessivi su base
010a1f32-b1c3-4b19-a052-e8949143cfc	modello-wrf-arw-a-3km-precipitazion	r_toscan	Precipitazione oraria (mm). Corsa del 2021-10-23 ore 00 UTC - Valido dalle ore
01224521-cfe7-4181-9e52-871b291db0	misurazioni-orarie-di-qualita-dell-aria-	r_emiro	Questo dataset fornisce l'evoluzione oraria, durante la giornata della stima dell
0135dcac-a4ed-40c4-87f4-78c0bf8505	orari-scolastici162873	SS_BZ	orari scolastici
01a349f7-1475-4429-ac8f-3e669b5e26c	orari-autobus-toscana	m_inf	#Orari autobus Toscana
0275d2b6-642c-4c48-8d44-166bb592f5	articolazione-degli-uffici-con-orari-al-	r_veneto	Articolazione degli uffici con orari al pubblico - Comune di Castel D'azzano - A
036ddeda-554e-4b98-afca-e53637ec98	orario-ferroviario-regionale-calendar-	r_lombar	Orario Ferroviario Regionale Lombardia. Servizio Trenord e Malpensa Express.
037e7443-207d-495c-831a-3f9ffc4de3a	database-axios-gestione-oraria-del-pe	istsc_caic89100x	Base dati contenente l'archivio relativo agli orari lavorativi e alle presenze del p

Dataset name: comune-di-numana-turismo-e-mobilit-orari-dei-trasporti

Dataset ID: 0011354f-2f71-440b-b901-cd517454d670

Dataset holder: r_marche

Dataset description: Orari dei trasporti locali

Last update: 2021-02-19 20:36:02

Fig. 2. Graphical user interface of the database query module

For a single dataset, the user inputs the Holder ID and selects from the results list. The assessment page displays when the selection is complete (Fig. 3). The user can choose the vocabulary profiling system and the portal to be evaluated. The assessment page summarizes results for every FAIR criterion and metric, including max, min, and avg scores. If files are linked to the dataset and have been evaluated, a clickable label provides access to those evaluations. Users can select between the evaluation results for the dataset and the holder. Additionally, users have the option to retrieve a dataset as via CKAN API by specifying the source URL or import a dataset as a CSV file.

3.5 Tool Settings

The euFAIR tool settings are provided via a dedicated JSON file. The file allows the user to specify: 1) *properties_settings*, to store the metadata properties as for F2 ad their weight, as well as new metadata profiles; 2) *Vocabularies*, to indicate, as for I2, the vocabulary format and properties; 3) *"format"*, to specify the best dataset format when multiple distribution formats are expected (every format is given a score based on its machine-readability and whether it is proprietary or not, according to the data quality guidelines of the European Commission); 4) *"permanent_link"*, which provides the structure of the tool's permalink.

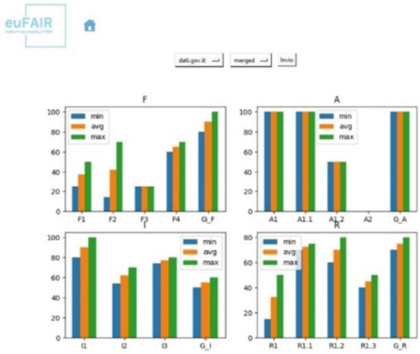


Fig. 3. Vocabulary profiling system interface.

4 Comparative Evaluation

To better understand euFAIR's potential, we assessed the same dataset with different tools, through the EUp, under a DCAT_AP profile. To carry out the comparison of tools we're taking into account the metadata of one of the datasets available in the Ip. The sample dataset contains (in JSON and CSV formats) the railway lines and stations of the city of Porto Recanati[3], located in the Italian administrative region of Marche.

The comparison between euFAIR and the other tools (SAT and GAT), highlighted several differences. The proposed euFAIR tool present additional advantages if compared to SAT, and GAT as it euFAIR offers additional features such as recognizing the validity of the EUp permalink, quantifying data richness, considering metadata persistence through data nature and portal, assessing the quantity of FAIR vocabulary usage, and evaluating license, provenance, and community standard adherence.

The Self-Assessment Tool (SAT) shows immediately its weakness as a deep knowledge of the dataset is required to the user in order to give the information required. Furthermore, answering these questions is rather difficult since questions such as "Did you provide sufficient metadata (information) about your data for others to find, understand and reuse your data?" are liable to non-objective answers.

The generic automatic tool (GAT) shows great potential, correctly scraping much of the data but analyzing it through generic metrics.

Table 1 shows the differences in the assessment score, which are mainly due to the different metrics and methods used by the tools, while Table 2 and Table 3 examine the differences between SAT and euFAIR and between GAT and euFAIR, respectively.

[3] The sample dataset is avaiable at the following link: https://dati.gov.it/view-dataset/dataset?id=1653a9ef-529f-4b65-91d7-894d19dea2e5, last accessed 2023/04/22.

Table 1. Assessment of the same dataset with different tools

	Self-Assessment Tool	Generic Automatic Tool	**euFAIR**
F	70	50	**57**
A	80	80	**80**
I	100	71	**29**
R	100	0	**29**

Table 2. Main differences between SAT and euFAIR

	Self-Assessment Tool (SAT) vs **euFAIR**
F	SAT awards full points for a DOI permalink, and euFAIR additionally includes the EUp portal permalink, with more options for datasets. SAT assesses comprehensive metadata descriptions and structured language usage, whereas euFAIR quantifies the extent of metadata descriptions
A	No relevant differences as both the tools have similar behaviors
I	SAT only takes into consideration the data format and the vocabulary type used in the metadata; euFAIR also focuses on how much the vocabularies are used
R	SAT only considers whether a license and a provenance are given; euFAIR also focus on whether the license uses its vocabulary and how the provenance is made explicit

Table 3. Main differences between GAT and euFAIR

	Generic Automatic Tool (GAT) vs **euFAIR**
F	GAT does not consider the EUp permalink as valid and does not quantify data richness
A	GAT does not consider the metadata persistent because it cannot find a persistence Policy key; euFAIR considers the nature of the data and the portal itself
I	GAT evaluates if the resource uses FAIR vocabs.; euFAIR also assesses how much
R	GAT only looks for license; euFAIR looks also for provenance and community standard

5 Discussion and Conclusions

In the open science arena, a big effort is required to provide and assess FAIR data. Existing SAT and GAT tools show some drawbacks which have been analyzed and overcome in this paper by proposing euFAIR, a novel tool to assess data FAIRness.

As of now, euFAIR is still in its alpha version, more testing being needed in order to validate its usability and efficacy and we hope a discussion will open to further enrich the tool's metrics and possibilities. The tool can target different recipients and different versions could be created depending on the final user and more extensive validation procedures will also be performed accordingly. Forthcoming improvements will target better data analysis and visualization to support Ip's assessment and a reporting system to give feedback to the PA so to improve data quality prior to dataset publication. In addition, we will evaluate whether decoupling the dataset evaluation update feature from other tool versions, such as one designed for the PA, as background evaluations could impact performances and a PA may not be interested in the update of all its datasets. By adopting euFAIR as a tool for PAs to evaluate the FAIRness of their datasets, it becomes feasible to use euFAIR metrics tests and to develop a collection to share on an already existent FAIR evaluator tool. Finally, we plan to widen the number of accepted formats to provide a more complete assessment on the data side and to consider more Holders' attributes such as geographical data to provide a deeper data analysis.

Acknowledgements. This work has been partially developed with the financial support of the Puglia region under the RIPARTI projects - POC PUGLIA FESRTFSE 2014/2020 – CUP: F87G22000270002, with the support of CETMA DISHME project (EU Grant Agreement n. 101083699), and with the support from Italian Research Center on High Performance Computing, Big Data and Quantum Computing (ICSC) funded by EU – NextGenerationEU (PNRR-HPC, CUP: C83C22000560007).

References

1. AGID, Dipartimento per la Trasformazione Digitale: Piano Triennale per l'Informatica (2022)
2. AGID: I dati aperti della PA. https://www.dati.gov.it/. Accessed 02 May 2023
3. European Commission: The official portal for European data. https://data.europa.eu/en. Accessed 01 May 2023
4. Wilkinson, M.D., et al.: The FAIR guiding principles for scientific data management and stewardship. Sci. Data. **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
5. Fairsharing: Philosophy of FAIR testing. https://fairsharing.github.io/FAIR-Evaluator-FrontEnd. Accessed 27 Apr 2023
6. Publications Office of the European Union.: Data.europa.eu data quality guidelines (2021)
7. Australian Research Data Commons (ARDC): Fair Self-Assessment Tool. https://ardc.edu.au/resource/fair-data-self-assessment-tool/. Accessed 25 Apr 2023
8. FAIRmetrics, FAIRsharing: The FAIR Maturity Evaluation Service. https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/. Accessed 26 Apr 2023
9. FAIRMetrics: FAIR Maturity Indicators and Tools. https://github.com/FAIRMetrics/Metrics. Accessed 26 Mar 2023
10. Krans, N.A., Ammar, A., Nymark, P., Willighagen, E.L., Bakker, M.I., Quik, J.T.K.: FAIR assessment tools: evaluating use and performance. NanoImpact. **27**, 100402 (2022). https://doi.org/10.1016/j.impact.2022.100402
11. Gehlen, K.P., Höck, H., Fast, A., Heydebreck, D., Lammert, A., Thiemann, H.: Recommendations for discipline-specific fairness evaluation derived from applying an ensemble of evaluation tools. Data Sci J. **21** (2022). https://doi.org/10.5334/dsj-2022-007
12. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research (2004)
13. Pavlina Fragkou (SEMIC EU): DCAT-AP 3.0. https://semiceu.github.io/DCAT-AP/releases/3.0.0/. Accessed 20 Mar 2023
14. AGID, Team Digitale: Linee guida per i cataloghi dati (2020)