The role of FAIR principles in high-quality research data documentation: Looking at national election studies

Wolfgang Zenk-Möltgen[1]

Abstract

The FAIR principles as a framework for evaluating and improving open science and research data management have gained much attention over the last years. By defining a set of properties that indicates good practice for making data findable, accessible, interoperable, and reusable (FAIR), a quality measurement is created, which can be applied to diverse research outputs, including research data. There are some software tools available to help with the assessment, with the F-UJI tool being the most prominent of them. It uses a set of metrics which defines tests for each of the FAIR components, and it creates an overall assessment score.

The article examines differences between manually and automatically assessing FAIR principles, shows that there are significantly different results by using national election studies as examples. An evaluation of progress is done by comparing the automatically assessed FAIRness scores of the datasets from 2018 with those of 2024, showing that there is only a very slight yet not significant difference. Specific measures which have improved the FAIRness scores are described by the example of the Politbarometer 2022 dataset at the GESIS Data Archive. The article highlights the role of archives in securing a high level of data and metadata quality and technically sound implementation of the FAIR principles to help researchers benefit from getting the most of their valuable research data.

Keywords

FAIR principles, Data documentation, Research data management, F-UJI test, Research transparency

Introduction

Good data documentation is indispensable for working with research data. It is especially important when using data that was collected or gathered by other researchers; knowledge about the data collection procedures, applied concepts, data cleaning and transformation steps, and decisions by principal investigators along the way is necessary to fully understand the findings and implications. In recent years, the movement for more research transparency has grown considerably, partly due to discovery of errors and scientific misconduct (Christensen et al., 2019; Freese & Peterson, 2017).

Open science and requests for a transparent process of the scientific endeavor led to several improvements with the availability and re-usability of research data. Social science data archives such as the ICPSR in the U.S., GESIS – Leibniz Institute for the Social Sciences in Germany, or the UK Data Archive have existed for decades. In addition, research data centers as well as more generic repositories for research data have emerged in recent years (e.g., Zenodo[2], Figshare[3], Dataverse[4], and

the Open Science Framework OSF[5]). All these services support the documentation of research data, however in different degrees of granularity. For all of these services, it is clear that well-done data documentation and provision of research data is costly and needs considerable resources (Perry & Netscher, 2022). But cases of fraud in more than only a few disciplines (Christensen et al., 2019) and acknowledgement for the value of data have led to considerable improvements in data curation, documentation, and access. Increasingly also research funders develop guidelines and regulations that support data archiving, metadata documentation, and data sharing (American Economic Association, 2024; Deutsche Gesellschaft für Soziologie, 2019;U.S. National Science Foundation, 2018; Wissenschaftsrat, 2020).

Especially in the political science domain, there is a collective understanding and expectation for journals to have a policy that requires authors to deposit the data underlying their article findings at a trusted archive or data repository and make it available for independent scrutiny (DA-RT, 2015). It was found that data availability is much higher for articles published in journals with a data sharing policy in place (Key, 2016; Zenk-Möltgen et al., 2018). Some policies also require independent verification of the results and sometimes provide staff at the journal to conduct this work.

Literature review

The FAIR principles have been developed as 'guiding principles for findable, accessible, interoperable, and re-usable data publishing'[6] within the FORCE11 scholarly community initiative (Wilkinson *et al.*, 2016). With the FAIRsFAIR (Fostering Fair Data Practices in Europe) project (Devaraju *et al.*, 2022), the FAIR guiding principles were established in practice as an open standard for evaluating the findability, accessibility, inter-operability, and re-producibility of research data. Evaluation of datasets against these criteria have initially been done manually (Bishop & Hank, 2018; Eder & Jedinger, 2019; Guillot et al., 2023; Maxwell et al., 2021). Most evaluations in the domains of political science or sociology, but also across disciplines (Stall et al., 2019), came to the conclusion that more needs to be done to make the used research data findable, accessible, interoperable, and reusable (Betancort Cabrera et al., 2020). Efforts have also been undertaken to balance requirements of openness with privacy requirements that regularly exist in the social sciences (Borgesius, et al., 2016).

Based on these developments, automated testing tools were developed to assess the 'FAIRness' of a given dataset. For an overview of manual and automatic solutions for evaluation, see the FAIRAssist[7] list. The most prominent automatic tool is the F-UJI tool (Devaraju & Huber, 2021) which can be used as a stand-alone implementation or via the provided website[8]. It is mainly dependent on persistent identifiers (or at least URLs) for the datasets and produces detailed test metrics as well as an overall FAIRness score. A comparison of different automatic FAIR assessment tools came to the conclusion that there are significant differences in the design, implementation, and documentation of the evaluation metrics for the tools (Sun, et al, 2022).

Assessments for several domain specific research datasets have been done in previous research (Alaterä et al., 2022; Petrosyan et al., 2023; Sofi-Mahmudi & Raittio, 2022). However, the automatic assessment has also been criticized as not being able to evaluate the quality of metadata or if metadata is 'rich' and domain-specific enough to enable reusability (Musen et al., 2022). This criticism is certainly valid for a comparison between automatic assessments to a review of metadata conducted by an expert in the fields looking at a small number of datasets. But the automatic procedures allow

the evaluation of a large number of datasets in a standardized way. And given that the metrics were developed within a broad framework of stakeholders and with large support of the scientific community, it can be assumed that they represent at least some basic common understanding of documentation quality (Devaraju et al., 2021).

However, it remains unclear if progress has happened with research data transparency over the years. To evaluate this question, automatic measurements which are using clearly defined tests might be suitable, even if they cannot perform a qualitative evaluation or determine whether humans are satisfied with the level of data documentation and availability. The aim of this article is therefore to find out if a valid evaluation of FAIRness can be done by an automatic assessment, if a comparison of these assessments over time can show improvement, and how data archives and repositories can contribute to this.

Research design

Using well-known research data seems highly appropriate for a comparison between manual and automatic assessments and for evaluation of progress over time. Therefore, the national election studies used by Eder and Jedinger (2019) are employed for the current paper as an example, since they represent a selection of highly relevant and well-curated social science datasets. To illustrate improvement measures, a single dataset from the GESIS Data Archive is selected from the same area of election studies, namely the Politbarometer 2022 (Forschungsgruppe Wahlen, Mannheim, 2023).

Based on workflows for using the DDI-Codebook and DDI-Lifecycle metadata standards, GESIS has provided good quality documentation for archived datasets for many years (Akdeniz and Zenk-Möltgen, 2017; Perry et al., 2019; Zenk-Möltgen, 2012; Zenk-Möltgen, 2023). Recent work has shown that improvements can be made when using the FAIR criteria and automatic assessments like the F-UJI tool (Saldanha Bach et al., 2023). In addition, it needs to be discussed which of these improvements are simply technical and which do really contribute to higher quality of documentation and thus contribute to more transparent science.

Given these considerations, this paper will look at election data as an example for the social science domain and will focus on the following research questions:

- Q1: Are there differences between the evaluation of FAIR criteria with automated tools as compared to manual procedures?
- Q2: Has the FAIRness of research data changed considerably over the six-year period between 2018 to 2024?
- Q3: Can data archives contribute to transparent science by implementing measures to increase FAIR scores for research data?

FAIRness evaluation

For an assessment of the FAIR criteria, Eder and Jedinger (2019) look at eighteen large-scale election studies from Western democracies, which cover at least two elections, sample the whole voting population and are mainly conducted for academic purposes (see Table 1). They operationalize several criteria within each of the four sections, and present tables with explanations for each of the FAIR scores, using zero if a criterion is not fulfilled, and one as fulfilled (sometimes also using 0.5 for partly

fulfilled). For all the national election studies, they provide the percent of studies fulfilling a score, but also as a summary index for each of the four criteria (summing up all operationalized scores of an area) (Eder and Jedinger, 2019, tbls. 8–11). However, they do not calculate an overall FAIRness score by combining the four indexes.

ID	Study	Abbreviation
1	American National Election Studies (ICPSR)	ANES
2	Australian Election Study	AES
3	Austrian National Election Study	AUTNES
4	Belgian National Election Study	BNES
5	British Election Study (UK Data Archive)	BES
6	Canadian Election Study	CES
7	Danish National Election Study	DNES
8	Dutch Parliamentary Election Study	DPES
9	Estonian National Election Study	ENES
10	Finnish National Election Study	FNES
12	German Longitudinal Election Study	GLES
13	Hellenic National Election Studies	ELNES
14	Italian National Election Study	ITANES
16	Icelandic National Election Study	ICENES
17	New Zealand Election Study	NZES
18	Norwegian Election Studies	NES
21	Swedish National Election Studies	SNES
22	Swiss Electoral Studies	SELECTS

Table 1: List of national election studies

The overall results from this study in 2018 show that findability of the national election studies is rather good, with an overall mean for the findability index of 5.3 (SD=1.88, min=2.0, max=8.0) (Eder & Jedinger, 2019, p. 661). Assessing the accessibility, Eder and Jedinger find: 'Most studies perform well (…). However, there seems to be room for improvement in regard to providing information on variables omitted (…) and on how to access variables in cases in which country-specific rules (…). Additionally, some studies could benefit from providing more extensive reports' (Eder & Jedinger, 2019, p. 663). They calculate a mean for the accessibility index of 5.44 (SD=1.79, min=3.0, max=9.0). For interoperability, the authors state that 'the provision of metadata is very good, with very little need for improvement', and note a mean for the interoperability index of 11.94 (SD=1.51, min=7.0, max=13.0) (Eder & Jedinger, 2019, p. 664). The authors conclude that reusability is 'quite satisfying' with a mean for the reusability index of 4.31 (SD=0.84, min=2.0, max=5.0) (Eder & Jedinger, 2019, p. 665).

Method and data

To answer the first research question 'Is the evaluation of FAIR criteria with automated tools an alternative to manual procedures?', I look at the same national election studies mentioned in the study by Eder and Jedinger (2018). The Eder and Jedinger study used national election studies (see Table 1) that were available through 2016. In a next step I searched for the persistent identifier (in this case a DOI, Digital Object Identifier) for the latest data availability and performed a F-UJI test for each of these studies. The F-UJI tool was used for the automated tests because it is among the most prominent tools, it provides detailed documentation for the test results, interprets the persistent identifiers as identifiers for the data (Sun et al., 2022), and because it has been found to correspond

well with manual assessments (Gehlen et al., 2022). A comparison was then done to show if the manually created results reported to summarize the situation in 2018 match these of the automated evaluation of the F-UJI tool for the 2018 datasets.

Using the scores provided by Eder and Jedinger (2019), a sum for an overall FAIRness score was computed in addition to the single scores for each of the four criteria. To compare the results of the manual assessments to the results of the conducted F-UJI tests, the mean percentage of the maximum possible value was calculated for each election study. The frequencies for ten quantiles of the percentages were then plotted and compared to allow an overview of the distribution of F-UJI percentages. In addition, two-sided t-tests allowed for assessing the significance of the mean differences. The Shapiro-Wilk test for normal distribution of the differences was checked as well. Because summing up four criteria with each one using a different maximum value leads to some bias towards the criteria with higher maximum values, the same procedure was applied to produce a FAIRness index using equal weight for each of the four criteria. In that way, differences between manual and automatic assessments of the same datasets were analyzed.

To answer the second research question 'Has the FAIRness of research data changed considerably over the six-year period between 2018 to 2024?', I updated the list of studies with newer rounds available until 2024. For seventeen of the eighteen election studies, a newer dataset was found, the exception being the Estonian National Election Study (ENES). For four countries (i.e., Belgium, Italy, Norway, Sweden), Module 5 of the Comparative Study for Electoral System (CSES) was used for the newer waves. The F-UJI tests were repeated for all available newer datasets (seventeen of eighteen), and scores for the single F.A.I.R. criteria (I will use the term F.A.I.R. to indicate the single scores for findable, accessible, interoperable, and re-usable) as well as for the combined FAIRness score and equally weighted FAIRness score were recorded.

To get the data for the comparison to 2024, the list of available persistent identifiers (DOIs) was researched for the newer datasets. Again, the F-UJI tests were calculated, and results were saved for analysis. A comparison of values for each of the four F.A.I.R. components, the overall FAIRness value as well as the equal weight FAIRness value was performed in the same way. This involved again plotting the distribution of frequencies for ten quantiles, this time for the automatic F-UJI tests for the 2018 datasets and the tests for the 2024 datasets. Also, t-tests were again used to find significant differences in the mean percentage values, and Shapiro-Wilk tests were conducted for assessing normal distribution.

 To answer the third research question 'Can data archives contribute to transparent science by implementing measures to increase FAIR scores for research data?' I use another example dataset, the Politbarometer 2022 study (Forschungsgruppe Wahlen, Mannheim, 2023). The Politbarometer survey series has been conducted since 1977 for the German TV network ZDF, and this dataset contains aggregated annual data from 1977 through 2022. The GESIS Data Archive creates aggregate datasets, documents the data, provides data access, and performs long-term archiving for this series, and also for many others (Schumann & Mauer, 2013; Recker et al., 2017). Data and metadata can be accessed (Hienert et al., 2019) at the GESIS Search webpage[9] as well as through the GESIS Politbarometer project webpage[10].

Since the F-UJI tool assessment relies on machine-readable links for data and metadata, results are very much dependent on the availability of standardized metadata. Since GESIS creates standardized documentation using the DDI standards[11] and controlled vocabularies[12] for archived studies (Akdeniz & Zenk-Möltgen, 2017), it was expected that the F-UJI assessment score will be high. It turned out that F-UJI scores were not as high as expected. To address this issue, GESIS has been working on several levels to improve the technical implementation of the standards that are used by the F-UJI tool for assessment (Saldanha Bach et al., 2023). By comparing the F-UJI tool assessment of this example Politbarometer dataset from an initial test on 3rd November 2023 with a later test on 29th April 2024, the effects of technical improvements, as well as effects of standardized metadata can be seen. This uses the list of metrics as applied by the F-UJI tool and describes the changes involved to improve the evaluation.

The data and scripts for all analyses have been deposited at a trusted repository and are made available for secondary research (Zenk-Möltgen, 2024). The analysis results and figures were produced using Stata 18.

Results

Comparing manual and automatic FAIRness scores

A persistent identifier was found for the respective last wave for four of the eighteen election studies investigated, even if in the original study by Eder and Jedinger recorded no persistent identifier (Danish, Estonian, Icelandic, and Norwegian election studies). This is already an improvement and will also help with the F-UJI test that needs this persistent identifier for evaluation. Only the Italian and the Canadian Election studies do not provide a persistent identifier.

The results from the F-UJI tool assessment of the national election studies in 2018 (see Table 2, autom. 2018) show that the mean values differ quite a lot: findability is quite high (5.78 of 7), and also interoperability is high (2.44 of 4), but accessibility is quite low (1.31 of 3) as well as reusability (3.94 of 10). The overall FAIRness score mean is 11.97 of 24, showing that the election studies get, on average, only half of the scores that they could get. Using the equally weighted value for the overall FAIRness score, a mean of 13.60 of 24 is only slightly better. We can say that the datasets of election studies conducted until 2018 get only mediocre FAIRness values from the F-UJI assessment.

National Election Studies		F-score	A-score	I-score	R-score	FAIRness score	FAIRness score (equal weight)	
Manual 2018	Mean	5.33	5.44	11.94	4.31	27.03	26.92	
	Stddev	1.88	1.79	1.51	.84	4.48	4.65	
	Max	8	10	13	5	36	36	
	N	18	18	18	18	18	18	
Autom. 2018	Mean	5.78	1.31	2.44	3.94	11.97	13.60	
	Stddev	0.75	0.54	1.15	1.06	5.17	3.37	
	Max	7	3	4	10	24	24	
	N	16	16	16	16	18	16	

Table 2: Comparing manual and automatic scores for the F.A.I.R. and overall FAIRness of the national election studies (values for 'Manual 2018' calculated from Eder and Jedinger 2018)

Comparing the distribution of values between the 2018 manual and 2018 automatic assessment can best be done by using percentage values because the maximum values differ for each assessment. The results show (see Figure 1) that findability is higher for the automatic assessment. Automatic assessed scores for accessibility are rather low, and interoperability is rated at mixed levels for the national election studies. Several studies get much lower scores for interoperability with the automatic assessment than with the manual one. Reusability is also at quite a low level compared to the manual assessment.
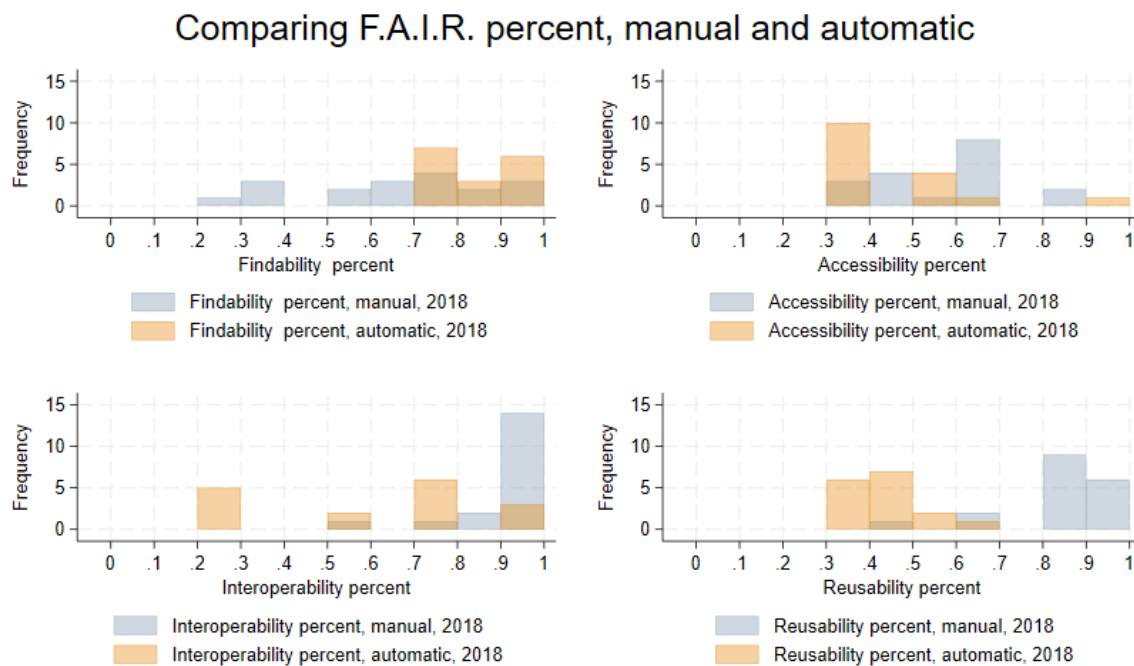

Figure 1: Comparing the F.A.I.R percent scores from manual and automatic assessment for the national election studies 2018

We can see in Figure 1 that only for findability does the automatic assessment by the F-UJI tool (lilac series) yield higher values than the manual assessments by Eder and Jedinger (red series). For accessibility, and very clearly also for interoperability and reusability, values for most election studies are lower with the automatic assessment than with the manual assessment.

The overall FAIRness scores for the automatic evaluation (see Figure 2) are medium and lower compared to the manual assessed values. Giving equal weight to the four criteria does not change this picture very much. This indicates that overall FAIRness values will yield lower scores with the automatic F-UJI tool assessment than when using a manual approach.
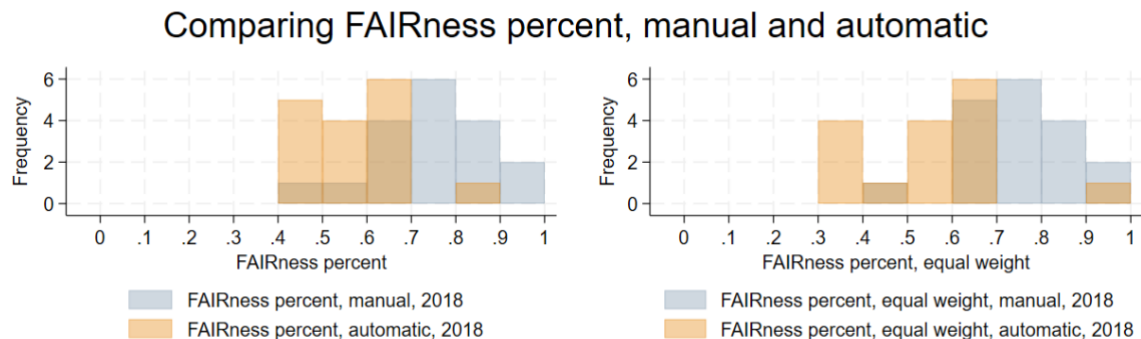

Figure 2: Comparing the overall FAIRness percent scores from manual and automatic assessment for the national election studies 2018

To see if the differences between the manual results and the automatic results of the mean values are significant, two-tailed t-tests were conducted (see Table 3, comparison A-B). Since the assumption of normal distribution is required for that, a Shapiro-Wilk test was performed for each comparison.

National Election Studies		F-%	A-%	I-%	R-%	FAIRness %	FAIRness % (equal weight)	
A - Manual 2018	Mean	.695	.544	.947	.875	.769	.765	
	Stddev	.233	.167	.067	.165	.105	.113	
	N	16	16	16	16	16	16	
B - Autom. 2018	Mean	.826	.438	.609	.394	.561	.567	
	Stddev	.107	.181	.288	.106	.123	.140	
Comparison A-B	M diff	-.131	.106	.338	.481[+]	.208	.199	
	t	-1.904	1.575	4.425	8.211	4.429	3.759	
	df	15	15	15	15	15	15	
	p	0.076	0.136	0.001	<0.001	0.001	0.002	

Table 3: Comparing manual and automatic F.A.I.R. and overall FAIRness mean percentage values and t-tests ([+] Shapiro–Wilk test is significant p<0.01 and indicates that the data is not normally distributed for this comparison)

The results for the findability of the 2018 datasets with (A) manual evaluation (M=0.695, SD=0.233) and (B) automatic evaluation (M=0.826, SD=0.107) show that findability is higher for the automatic assessment, t(15)=-1.904, p=0.076, significant at the 0.10-level. Results for accessibility, interoperability, and reusability show lower values for the automatic assessment by the F-UJI tool than with the manual assessment (see Table 3, comparison A-B), however, the difference for accessibility is not significant. The higher values for mean interoperability and mean reusability with the manual assessment are significant at the 0.01-level. Also, the overall percentage scores for FAIRness differ significantly for the (A) manual assessment (M=0.769, SD=0.105) from the (B) automatic assessment

(M=0.561, SD=0.123), t(15)=4.429, p=0.001. The same is true for the equal weight FAIRness percentage.

These results suggest that the answer to research question Q1 'Are there differences between the evaluation of FAIR criteria with automated tools compared to manual procedures?' is that in this case, the automatic evaluation with the F-UJI tool gives lower scores than the manual evaluation, except for the criterium of findability, where we see somewhat higher scores for the automatic evaluation.

Comparing automatic FAIRness scores for 2018 and 2024 datasets

The results from the F-UJI tool assessment of the more recent national election studies in 2024 (see Table 4, autom. 2024) show that the mean values also differ a lot: findability is again quite high (5.91 of 7), and also interoperability is high (2.76 of 4), but accessibility is again quite low (1.47 of 3) as well as reusability (4.41 of 10). The overall FAIRness score mean is somewhat better with 13.75 of 24, showing that the newer election studies get, on average, more than half of the scores that they could get. Using the equally weighted value for the overall FAIRness score, a mean of 14.80 of 24 is again slightly better. All these values are slightly higher than the automatic assessed ones from 2018. We can say that the newer election studies datasets get moderate percentage scores, and to a certain degree better FAIRness values from the F-UJI assessment tool than the 2018 datasets.

National Election Studies		F-score	A-score	I-score	R-score	FAIRness score	FAIRness score (equal weight)	
Autom. 2018	Mean	5.78	1.31	2.44	3.94	11.97	13.60	
	Stddev	0.75	0.54	1.15	1.06	5.17	3.37	
	Max	7	3	4	10	24	24	
	N	16	16	16	16	18	16	
Autom. 2024	Mean	5.91	1.47	2.76	4.41	13.75	14.80	
	Stddev	0.85	0.57	1.15	1.62	4.66	3.42	
	Max	7	3	4	10	24	24	
	N	17	17	17	17	18	17	

Table 4: Comparing 2018 and 2024 scores from the F-UJI assessment of the national election studies

Comparing the distribution of values between the automatic assessments by the F-UJI tool for the 2018 and 2024 datasets can again be done by using percentage values. The results show (see Figure 3) that findability is nearly the same for both assessments (except that one study seems to be less findable in 2024). Assessment scores for accessibility have slightly improved, and interoperability and reusability are at higher levels for the newer election studies. We can see in Figure 3 that for all separate indices of findability, accessibility, interoperability, and reusability the assessments for the newer datasets (yellow series) yield higher values than for the older datasets (green series). For accessibility, interoperability, and reusability, this result is much clearer than for findability.
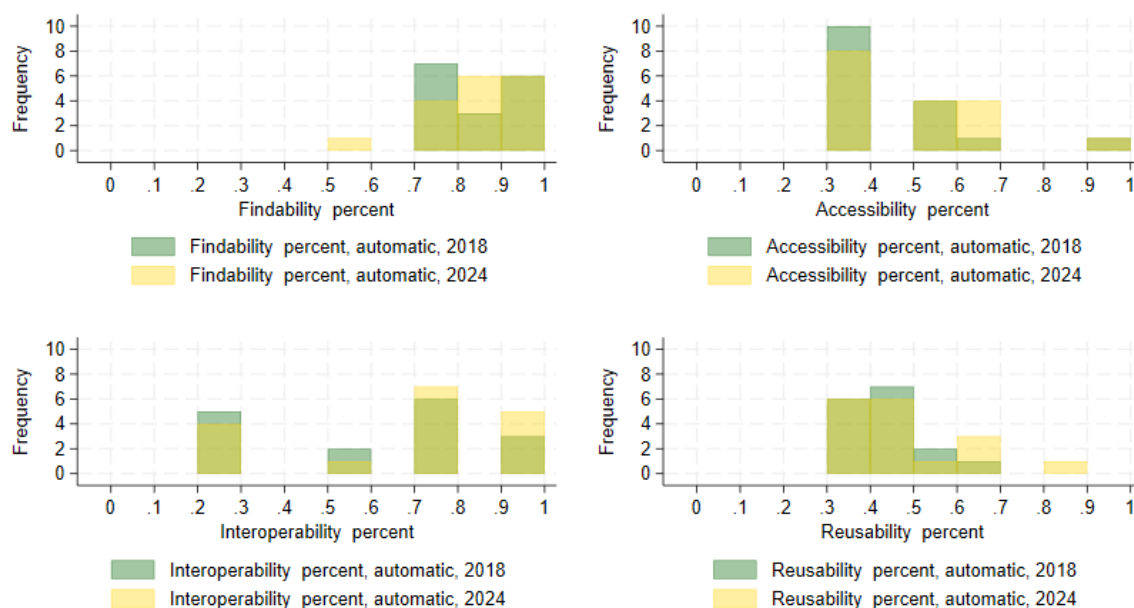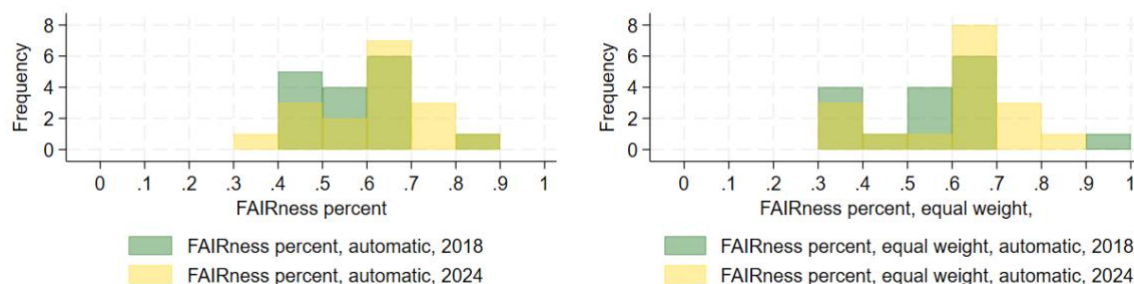
Figure 3: Comparing the F.A.I.R percent scores from automatic assessment for the national election studies 2018 and 2024

The overall FAIRness scores for the evaluation of the newer datasets from up to 2024 (see Figure 4) are still moderate but somewhat higher compared to the older assessed datasets. Giving equal weight to the four criteria of the FAIRness index does again show the same picture. This makes clear that there is a slight improvement of the overall FAIRness values for the national election studies that can be shown by conducting the automatic F-UJI tool assessment.

Figure 4: Comparing the overall FAIRness percent scores from automatic assessment for the national election studies 2018 and 2024

Differences between the mean results of the older and newer national election studies were again evaluated for significance with a two-tailed t-test (see Table 5, comparison C-D), and the Shapiro-Wilk test for the assumption of normal distribution. The mean comparison results for all indicators show no significant differences. Even if means for the newer datasets are slightly higher for all the single (i.e., findability, accessibility, interoperability, and reusability) indicators, none of these differences

reach a statistically significant level (see Table 5, comparison C-D). Results for the overall indicators of FAIRness and equal weight FAIRness percentage are consistent.

National Election Studies		F-%	A-%	I-%	R-%	FAIRness %	FAIRness % (equal weight)	
C - Autom. 2018	Mean	.824	.433	.583	.393	.556	.558	
	Stddev	.111	.187	.278	.110	.126	.141	
	N	15	15	15	15	15	15	
D - Autom. 2024	Mean	.833	.478	.667	.420	.589	.599	
	Stddev	.123	.198	.294	.142	.130	.141	
Comparison C-D	M diff	-.010	-.044	-.083[+]	-.027	-.033	-.041	
	t	-0.397	-0.745	-1.581	-0.745	-1.240	-1.401	
	df	14	14	14	14	14	14	
	p	0.698	0.469	0.136	0.469	0.235	0.183	

Table 5: Comparing 2018 and 2024 F.A.I.R. and overall FAIRness mean percentage values and t-tests ([+] Shapiro–Wilk test is significant p<0.01 and indicates that the data is not normally distributed for this comparison)

These findings indicate that the answer to the second research question 'Has the FAIRness of research data changed considerably over the six-year period between 2018 to 2024?' is that six year later, there is a rather small improvement in the FAIRness of national election.

Improving FAIRness scores

Research question Q3 'Can data archives contribute to transparent science by implementing measures to increase FAIR scores for research data?' requires examining measures for improving the FAIRness scores and a discussion of the contribution of these measures to transparent science. The Politbarometer 2022 study (Forschungsgruppe Wahlen, Mannheim, 2023) is used to illustrate some of the implemented measures. This study is well suited because it has a comprehensive documentation, both in English and German, and is well known and used by many political scientists.

Table 6 lists for each section of the FAIR criteria the defined metrics and the number of technical tests performed by the F-UJI tool (see Table 6, blue section). For sixteen of the seventeen metrics, tests are included. The table also lists the scores that the F-UJI test for the Politbarometer 2022 study achieved initially on November 3rd, 2023, before implementing dedicated measures to support the FAIRness assessment (see Table 6, orange section). After several technical adaptations and changes in the metadata were implemented, a final F-UJI test for the Politbarometer 2022 study was conducted on April 29th, 2024, and those results are listed also (see Table 6, yellow section). For an easier overview, metrics that improved are indicated by an X in the column titled 'Improvement'. Overall scores for each of the F.A.I.R criteria and the FAIRness scores and percent values are shown in the 'Total' columns and rows.

Section	Metric Identifier	Metric Name	Tests	Improvement	Score	MaxScore	Total	maxTotal	Total-%	Score2	MaxScore2	Total2	maxTotal2	Total2-%
					November 3rd, 2023					April 29th, 2024				
Findability	FsF-F1-01D	Data is assigned a globally unique identifier.	2		1	1				1	1			
	FsF-F1-02D	Data is assigned a persistent identifier.	2		1	1				1	1			
	FsF-F2-01M	Metadata includes descriptive core elements (creator, title, data identifier, publisher, publication date, summary and keywords) to support data findability.	3		2	2				2	2			
	FsF-F3-01M	Metadata includes the identifier of the data it describes.	2	X	0	1				1	1			
	FsF-F4-01M	Metadata is offered in such a way that it can be retrieved by machines.	2		1	2	5	7		1	2	6	7	
Accessibility	FsF-A1-01M	Metadata contains access level and access conditions of the data.	3	X	0,5	1				1	1			
	FsF-A1-02M	Metadata is accessible through a standardized communication protocol	1		1	1				1	1			
	FsF-A1-03D	Data is accessible through a standardized communication protocol	1	X	0	1				1	1			
	FsF-A2-01M	*Metadata remains available, even if the data is no longer available.*					1,5	3				3	3	
Interoperability	FsF-I1-01M	Metadata is represented using a formal knowledge representation language.	2		1	2				1	2			
	FsF-I2-01M	Metadata uses semantic resources.	2		1	1				1	1			
	FsF-I3-01M	Metadata includes links between the data and its related entities.	2	X	0	1	2	4		1	1	3	4	
Reusability	FsF-R1-01MD	Metadata specifies the content of the data.	6	X	1	4				2	4			
	FsF-R1.1-01M	Metadata includes license information under which data can be reused.	2		1	2				1	2			
	FsF-R1.2-01M	Metadata includes provenance information about data creation or generation.	2		1	2				1	2			
	FsF-R1.3-01M	Metadata follows a standard recommended by the target research community of the data.	3		1	1				1	1			
	FsF-R1.3-02D	Data is available in a file format recommended by the target research community.	3	X	0	1	4	10		1	1	6	10	
	Total						12,5	24	52,1%			18	24	75%

Table 6: F-UJI test comparison on Nov 3rd, 2023, and Apr 29th, 2024 for Politbarometer 2022, https://doi.org/10.4232/1.14103

Altogether, FAIRness scores were improved from 12.5 to 18 points out of a maximum of 24, resulting in an increase from 52.1% to 75% (see Table 6). This was accomplished by changing data and metadata for six of the seventeen metrics. These six improvements are very generic and do not apply only to this specific study, as will become clear in the following section. The improved FAIRness score for the Politbarometer 2022 dataset moved the study from the 'moderate' score category to the 'advanced' FAIRness score category. The findability score for the Politbarometer 2022 dataset was improved from 5 to 6 with this measure, leading to the evaluation category 'advanced' instead of 'moderate'. The accessibility score for the dataset was increased from 1.5 to 3, leading to the evaluation category of 'advanced' instead of 'initial'. The interoperability score for this dataset was amended from 2 to 3, leading to an evaluation of 'advanced' instead of 'moderate'. The reusability score for the Politbarometer 2022 dataset increased from 4 to 6 with these measures, leading to the evaluation category 'moderate' instead of 'initial'.

One general improvement was needed to allow the F-UJI tool to access the metadata: Because the GESIS Search webpage is a single page application that uses JavaScript for asynchronously loading content after the page has already been delivered to the client (AJAX), the F-UJI tool initially could not extract the schema.org metadata that is provided at each DOI landing page. For that reason, each landing page was modified so that it contains a FAIR Signposting[13] 'described-by'-link for retrieving the schema.org metadata in JSON-LD format, which is supported by the F-UJI tool (and other tools). The metadata of the example Politbarometer 2022 dataset is displayed in Figure 5. Subsequent improvements for the F-UJI tool assessment were dependent on this first change in the architecture of the GESIS Search webpage.

Figure 5: Metadata in JSON-LD format for Politbarometer 2022, https://doi.org/10.4232/1.14103

Findability

In the area of findability, F-UJI test scores were improved for:

- FsF-F3-01M-1 Metadata contains data content related information (file name, size, type).
- FsF-F3-01M-2 Metadata contains a PID or URL which indicates the location of the downloadable data content.

Since data downloads are already part of the GESIS Search webpage, the information about downloadable files, their name, format, size, and URL were available. This metadata was included into the JSON-LD metadata, allowing the F-UJI tool to extract this information. Having the name, size and type specified led to improvement in the score of the first metric. Providing the URL for accessing the dataset resulted in the second improvement, even when the download is possible only for registered users.

The improvement in the findability score is relevant for all studies with downloadable data files which are the majority of studies on the GESIS Data Archive.

Accessibility

In the accessibility category the F-UJI test scores were improved for:

- FsF-A1-01M-2 Data access information is machine readable.
- FsF-A1-03D-1 Metadata includes a resolvable link to data based on standardized web communication protocols.

The first improvement for data access information consists of specifying the JSON-LD fields 'conditionsOfAccess' (free text) and 'isAccessibleForFree' (controlled vocabulary). This allowed the F-UJI tool to extract the data access type. Providing this information was easy since this metadata is available for all studies at the GESIS Data Archive.

The second improvement was solved by the implementation of the findability measure of providing a URL to the downloadable dataset (see above). A URL starting with 'https' is recognized by the F-UJI tool as a standard web-protocol and therefore this improvement affects all studies with provided data download links.

Interoperability

In the interoperability category, F-UJI test scores were improved for:

- FsF-I3-01M-1 Related resources are explicitly mentioned in metadata.

To achieve this improvement RDF FAIR Signposting link headers were added into the html source of the landing page. This technique helps to identify data resources that are machine-actionable on the web, and it allows to specify metadata independently from a specific FAIR assessment tool. This change was implemented for all studies archived at the GESIS Data Archive. As a result, the F-UJI tool assessments were improved.

Reusability

For the reusability category, F-UJI test scores were enhanced for:

- FsF-R1-01MD-2 Verifiable data descriptors (file info, measured variables or observation types) are specified in metadata.
- FsF-R1.3-02D-1 The format of a data file given in the metadata is listed in the long-term file formats, open file formats or scientific file formats controlled list.

The measure for the first improvement is again due to the findability improvement (see above): Providing the data download URL also improves this score for interoperability. Again, studies with available data downloads benefit from this improvement.

The second improvement was possible because the GESIS Data Archive provides most data files as SPSS and/or Stata files. The formats for downloadable data files (provided already for the findability improvement) need to be specified as MIME types, and both 'application/x-stata-dta' (Stata) and 'application/x-spss-sav' (SPSS) formats are recognized as scientific community standards by the F-UJI tool. This improvement applies to nearly all studies with downloadable data files, given they have an SPSS and/or Stata file available.

Summary

Overall, FAIRness score for the Politbarometer 2022 data that were used as a case study, was improved from 12.5 to 18 points out of a maximum of 24, resulting in an increase from 52.1% to 75% (see Table 6). It becomes clear from the example dataset and the explanations that especially the first adaptations for findability are important for allowing the F-UJI tool to better assess the research data. For accessibility and reusability, the same measure is relevant and increases the F-UJI score of the evaluated resource also for these sections. The increase in FAIRness score was a result of several modifications: providing the downloadable research data files in standardized form with their name, format as MIME type, size, and URL, adding the access conditions in machine readable form, and including RDF FAIR Signposting link headers.

Further modifications were discussed, but not implemented, e.g., using a data license that is part of SPDX[14] would have had an improvement effect for the accessibility score FsF-A1-01M-3 ('Data access information is indicated by (not machine readable) standard terms'). Existing usage regulations[15] for the GESIS Data Archive do not allow the re-distribution of data by researchers themselves, and therefore cannot be used with any of the Creative Commons or other open licenses. However, these regulations allow the use of data for research purposes without any further restrictions (depending on access class) and should therefore still be considered as enabling transparent research practices. Other modifications are still under evaluation, e.g., implementing a formal representation of PROV-O metadata that would improve metric FsF-R1.2-01M-2 ('Metadata contains provenance information using formal provenance ontologies (PROV-O)').

Discussion and conclusions

It has been shown that the evaluation of FAIR criteria can be done with automated tools as an alternative to manual procedures. However, the resulting FAIRness levels are not comparable, and the differences for the mean assessment between both methods are statistically significant (except for accessibility). In the case of using the F-UJI tool for an automatic evaluation of the national election studies, the resulting values are higher for findability, and lower for accessibility, interoperability, and reusability than with the manual assessment.

Regarding the change that might have happened during the last six years, the F-UJI tool was used for an automated FAIRness assessment of the national election studies data. It has been demonstrated that there was not much change with the level of FAIRness from the datasets available in 2018 and those available in 2024. Slight improvements have been found, and for single studies there might have even been considerable improvements – especially in the criteria of accessibility, interoperability, and reusability. Overall, there were no significant differences in the average level of FAIRness scores of the newer datasets compared to the older ones.

Potentially, changes in the data documentation or website updates between 2018 and 2024 might have resulted in improvements for single studies that were not detected by this comparison (because the metrics were all assessed in 2024). However, this kind of improvement would have contributed to the difference between manual and automatic evaluation examined in the first step. The higher findability score identified in this comparison may explain some of these findings.

The study has several limitations that should be mentioned. Firstly, the evaluation used a selection of very prominent and widely recognized national election studies which are carefully curated and documented. Results might look different when including studies that are not as prominent, have fewer resources for data curation and documentation, or are not survey data at all. Secondly, several metrics rely on technical or metadata standards that are currently not commonly used, such as FAIR Signposting or SPDX licenses. With an uptake of those standards for web resources, research data might also benefit. Custom implementations of those standards require technical and metadata developments, and even large institutions may need years to realize this.

With the example of the Politbarometer 2022 dataset, the paper describes several modifications that led to improved FAIRness scores as measured by the F-UJI tool. This demonstrates that small technical changes can have a substantial impact. It may be worth mentioning that the implementation of these changes did not require a lot of work by the archive team. Given the significant improvement in FAIRness scores, it should be evident that the investment in providing metadata in a suitable technical format is worth the effort. But how do these technical improvements contribute to more transparent research practices, as asked in research question three?

Christensen, Freese and Miguel (2019, p. 12) link the question 'What is ethical research?' back to the foundations laid by Robert K. Merton in 1942 (Merton, 1942). They write that 'Openness, integrity, and transparency are at the very heart of Merton's influential articulation of scientific research norms' (Christensen et al., 2019, p. 20). However, they also describe the gap between current research practices and the scientific ideal and formulate recommendations for research practices in four areas: Reporting standards, replication, data sharing, and reproducible workflow (Christensen et al., 2019, p. 141ff). Except for reporting standards, those areas are also relevant when performing FAIRness tests: Replication is only possible with data sharing, which in turn means that the research data is FAIR: findable, accessible, interoperable, and reusable. Reproducible workflows become possible when FAIR data is the reality. Thus, improving the FAIRness of research data is an essential foundation for ethical research practices.

A certain degree of FAIRness can also be achieved by using unstandardized methods. Making research data available upon request or on a website, describing it in a report, and using a commonly used data format may already be a first step for some researchers. But given the developments in data science and big data, the use of social media data to analyze social phenomena, and given the increasing variety, volume, and velocity of change, known as the 'three V' (Kockum & Dacre, 2021), a machine-actionable approach is needed (Jensen et al., 2019; Weller & Kinder-Kurlanda, 2021; Weller & Strohmaier, 2014). Further on are developments of artificial intelligence in the social sciences that rely largely on machine-accessible data and are increasingly being used (Grossmann et al., 2023). Supporting researchers working in such environments requires a research infrastructure that delivers FAIR data services. FAIRness indicators provide a standardized way to assess the quality of documentation in this respect.

The F-UJI tool for assessing the FAIRness of research data makes it possible to document the FAIRness of research data in a transparent way. As shown in this article, the basis for FAIRness evaluations in an automated procedure is usage of open standards that are applied according to principles of open science. The openness can be seen when the F-UJI tool performs the assessment based on data and

metadata that is freely available on the web. Anyone can access the basis for evaluation. The integrity is represented by several checks of the F-UJI tool which are validations of the provided metadata, e.g., FsF-R1-01MD-3 ('Data content matches file type and size specified in metadata'), FsF-R1-01MD-4 ('Data content matches measured variables or observation types specified in metadata') or FsF-F2-01M-3 ('Core descriptive metadata is available' checks for defined metadata fields). The transparency of the process is enabled by the availability of the F-UJI tool itself as an open source on GitHub under the MIT open-source license[16] , making criteria for evaluation visible, and allowing users to scrutinize each single metric and test included. Data archives and other research data centers can contribute considerably to the level of FAIRness of the research data they curate and disseminate, as has been shown by the example dataset.

However, there is more to ethical research practices than implementing measures to achieve the highest score of research data FAIRness. Additional things to consider include for example reporting standards and transparency in methods for creating scientific results. Archives and other data curating institutions can provide scientific data with the highest possible FAIRness levels, and thus can provide one piece in the mosaic of the scientific ideal.

Acknowledgements

References

Akdeniz, E. and Zenk-Möltgen, W. (2017). 'DDI-Lifecycle at the Data Archive: the Metadata Schema for Documentation in Different Software Tools', *GESIS Papers*, *2017*(18). https://doi.org/10.21241/SSOAR.52487

Alaterä, T., Kleemola, M., Ala-Lahti, H. and Jerlehag, B. (2022). D4.5 Report on completed FAIR data standard adoption and certifications of data repositories in the region. https://doi.org/10.5281/ZENODO.7303538

American Economic Association (2024). Data and code availability policy. https://www.aeaweb.org/journals/data/data-code-policy

Betancort Cabrera, N., Bongartz, E.C., Dörrenbächer, N., Goebel, J., Kaluza, H. and Siegers, P. (2020). White paper on implementing the FAIR principles for data in the social, behavioural, and economic eciences', *RatSWD Working Paper Series*. https://doi.org/10.17620/02671.60

Bishop, B. W., & Hank, C. (2018). Measuring fair principles to inform fitness for use. *International Journal of Digital Curation*, *13*(1), 35–46. https://doi.org/10.2218/ijdc.v13i1.630

Borgesius, F.Z., Gray, J. and Van Eechoud, M. (2016). Open data, privacy, and fair information principles: Towards a balancing framework', *Berkeley Technology Law Journal*, *30*(3), 2073-2131. https://doi.org/10.15779/Z389S18 .

Christensen, G.S., Freese, J. and Miguel, E. (2019). *Transparent and reproducible social science research: How to do open science*. University of California Press.

DA-RT (2015). 'Data Access & Research Transparency'. https://www.dartstatement.org/about

Deutsche Gesellschaft für Soziologie (2019). 'Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie'. https://soziologie.de/aktuell/stellungnahmen/news/bereitstellung-und-nachnutzung-von-forschungsdaten-in-der-soziologie

Devaraju, A., & Huber, R. (2021). An automated solution for measuring the progress toward FAIR research data. *Patterns*, *2*(11), 100370. https://doi.org/10.1016/j.patter.2021.100370

Devaraju, A., Huber, R., Mokrane, M., Herterich, P., Cepinskas, L., de Vries, J., L'Hours, H., Davidson, J. and White, A. (2022). *FAIRsFAIR Data Object Assessment Metrics*. https://doi.org/10.5281/zenodo.6461229

Devaraju, A., Mokrane, M., Cepinskas, L., Huber, R., Herterich, P., De Vries, J., Akerman, V., L'Hours, H., Davidson, J., & Diepenbroek, M. (2021). From conceptualization to implementation: Fair assessment of research data objects. *Data Science Journal*, *20*, 4. https://doi.org/10.5334/dsj-2021-004

Eder, C. and Jedinger, A. (2018). FAIR national election studies: How well are we doing? (GESIS SDN-10.7802-1761) [Data set]. GESIS. https://doi.org/10.7802/1761

Eder, C., & Jedinger, A. (2019). FAIR national election studies: How well are we doing? *European Political Science*, *18*(4), 651–668. https://doi.org/10.1057/s41304-018-0194-3

Forschungsgruppe Wahlen, Mannheim (2023). Politbarometer 2022 (Cumulated Data Set). (GESIS, ZA7970; Version 1.0.0) [Data set]. GESIS. https://doi.org/10.4232/1.14103

Freese, J., & Peterson, D. (2017). Replication in social science. *Annual Review of Sociology*, *43*(1), 147–165. https://doi.org/10.1146/annurev-soc-060116-053450

Gehlen, K. P., Höck, H., Fast, A., Heydebreck, D., Lammert, A., & Thiemann, H. (2022). Recommendations for discipline-specific fairness evaluation derived from applying an ensemble of evaluation tools. *Data Science Journal*, *21*, 7. https://doi.org/10.5334/dsj-2022-007

Grossmann, I., Feinberg, M., Parker, D. C., Christakis, N. A., Tetlock, P. E., & Cunningham, W. A. (2023). AI and the transformation of social science research. *Science*, *380*(6650), 1108–1109. https://doi.org/10.1126/science.adi1778

Guillot, P., Bøgsted, M., & Vesteghem, C. (2023). FAIR sharing of health data: A systematic review of applicable solutions. *Health and Technology*, *13*(6), 869–882. https://doi.org/10.1007/s12553-023-00789-5

Hienert, D., Kern, D., Boland, K., Zapilko, B., & Mutschke, P. (2019). A digital library for research data and related information in the social sciences. *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 148–157. https://doi.org/10.1109/JCDL.2019.00030

Jensen, U., Netscher, S. and Weller, K. (eds) (2019). *Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten: Grundlagen und praktische Lösungen für den Umgang mit quantitativen Forschungsdaten*. Verlag Barbara Budrich.

Key, E. M. (2016). How are we doing? Data access and replication in political science. *Political Science and Politics*, *49*(2), 268–272. https://doi.org/10.1017/S1049096516000184

Kockum, F., & Dacre, N. (2021). Project management volume, velocity, variety: A big data dynamics approach. *Advanced Project Management, 21*(1). https://doi.org/10.2139/ssrn.3813838

Maxwell, L., Shreedhar, P., Dauga, D., McQuilton, P., Terry, R., Denisiuk, A., Molnar-Gabor, F., Saxena, A. and Sansone, S.-A. (2021). *FAIR, ethical, and coordinated data sharing for COVID-19 response: a review of COVID-19 data sharing platforms and registries*. https://doi.org/10.21203/rs.3.rs-1045632/v1

Merton, R.K. (1942). A note on science and democracy. *Journal of Legal and Political Sociology*, 1(1), 115–126.

Musen, M. A., O'Connor, M. J., Schultes, E., Martínez-Romero, M., Hardi, J., & Graybeal, J. (2022). Modeling community standards for metadata as templates makes data FAIR. *Scientific Data*, *9*(1), 696. https://doi.org/10.1038/s41597-022-01815-3

Perry, A., & Netscher, S. (2022). Measuring the time spent on data curation. *Journal of Documentation*, *78*(7), 282–304. https://doi.org/10.1108/JD-08-2021-0167

Perry, A., Watteler, O., Zenk-Möltgen, W. and Gregory, A. (2019, May 27-31). *How can research projects benefit from standardized metadata Like DDI?* [Conference presentation]. IASSIST 2019, Sydney, Australia. https://doi.org/10.5281/ZENODO.3612730

Petrosyan, L., Aleixandre-Benavent, R., Peset, F., Valderrama-Zurián, J. C., Ferrer-Sapena, A., & Sixto-Costoya, A. (2023). FAIR degree assessment in agriculture datasets using the F-UJI tool. *Ecological Informatics*, *76*, 102126. https://doi.org/10.1016/j.ecoinf.2023.102126

Recker, J., Zenk-Möltgen, W., & Mauer, R. (2017). Applications of research data management at GESIS Data Archive for the Social Sciences. In J. B. Thestrup & F. Kruse (Eds.), *Research Data Management—A European Perspective* (pp. 119–146). De Gruyter. https://doi.org/10.1515/9783110365634-008

Saldanha Bach, J., Klas, C.-P., Mathiak, B., Yudong Zhang and Mutschke, P. (2023). *FAIRness assessment: A comparison of the RDA model and the F-UJI Automated tool report*. https://doi.org/10.5281/ZENODO.8308902

Schumann, N. and Mauer, R. (2013). The GESIS Data Archive for the Social Sciences: A widely recognised data archive on its way. *International Journal of Digital Curation*, *8*(2), 215–222 https://doi.org/10.2218/ijdc.v8i2.285

Sofi-Mahmudi, A. and Raittio, E. (2022). Transparency of COVID-19 related research in dental journals. *Frontiers in Oral Health*, *3*, p. 871033. Available at: https://doi.org/10.3389/froh.2022.871033

Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., Parsons, M., Robinson, E. and Wyborn, L. (2019). Make scientific data FAIR. *Nature*, *570*(7759), 27–29. https://doi.org/10.1038/d41586-019-01720-7

Sun, C., Emonet, V. and Dumontier, M. (2022). A comprehensive comparison of automated FAIRness evaluation tools. In K. Wolstencroft, A. Splendiani, M.S. Marshall, C. Baker, A. Waagmeester, M. Roos, R. Vos, R. Fijten, & L.J. Castro (Eds.) *13th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences*, 44-53. https://ceur-ws.org/Vol-3127/paper-6.pdf.

U.S. National Science Foundation (2018). Data management guidance for SBE directorate proposals and awards. https://new.nsf.gov/sbe/data-management

Weller, K. and Kinder-Kurlanda, K. (2021). Uncovering the challenges in collection, sharing and documentation: The hidden data of social media research?, *Proceedings of the International AAAI Conference on Web and Social Media*, 9(4), 28–37. https://doi.org/10.1609/icwsm.v9i4.14687

Weller, K. and Strohmaier, M. (2014). Social media in academia: How the Social Web is changing academic practice and becoming a new source for research data, *it - Information Technology*, 56(5), 203–206. https://doi.org/10.1515/itit-2014-9002

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wissenschaftsrat (2020). 'Zum Wandel in den Wissenschaften durch datenintensive Forschung'. https://www.wissenschaftsrat.de/download/2020/8667-20.pdf?__blob=publicationFile&v=5

Zenk-Möltgen, W. (2012). 'Metadaten und die Data Documentation Initiative (DDI)', in R. Altenhöner and C. Oellers (Eds.) *Langzeitarchivierung von Forschungsdaten: Standards und disziplinspezifische Lösungen* (pp. 111–126). Scivero Verl

Zenk-Möltgen, W. (2023, November 27-29) *Implementing Colectica at the GESIS data archive* [Conference presentation].EDDI2023, Ljubljana, Slovenia. https://doi.org/10.5281/ZENODO.10257202

Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V. and Balaban, E. (2018.) Factors influencing the data sharing behavior of researchers in sociology and political science, *Journal of Documentation*, *74*(5), 1053–1073. https://doi.org/10.1108/JD-09-2017-0126

Zenk-Möltgen, W. (2024). Replication data and code for: The role of FAIR principles in high-quality research data documentation: Looking at national election studies, (GESIS SDN-10.7802-2798; Version 1.0.0) [Data set]. GESIS. https://doi.org/10.7802/2798

Endnotes

[1] Wolfgang Zenk-Möltgen works at GESIS – Leibniz Institute for the Social Sciences, Germany. He can be reached by email: wolfgang.zenk-moeltgen@gesis.org. https://orcid.org/0000-0002-2158-3941

[2] https://zenodo.org/
[3] https://figshare.com/
[4] https://dataverse.harvard.edu/
[5] https://osf.io/
[6] https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/
[7] https://fairassist.org/#!/
[8] https://www.f-uji.net/
[9] https://search.gesis.org/
[10] https://www.gesis.org/en/elections/politbarometer
[11] https://ddialliance.org/
[12] https://ddialliance.org/controlled-vocabularies
[13] https://signposting.org/FAIR/
[14] https://spdx.dev/about/overview/
[15] https://www.gesis.org/fileadmin/user_upload/Usage_regulations.pdf
[16] https://github.com/pangaea-data-publisher/fuji