It´s FAIR and generic?! A systematic analysis of FAIR Assessment Tools

Anna Lehmann-Gnirck[0000-0002-5739-4472]

[1] University of Applied Sciences Potsdam, Kiepenheuerallee 5, 14469 Potsdam, Germany
anna.lehmann-gnirck@fh-potsdam.de

Abstract: The FAIR Data Principles aim to enable interoperability and reuse of research data across disciplinary boundaries. In practice, FAIRness is often assessed using (semi)automated and manual tools that translate the principles into operational scores using assessment metrics. This paper examines whether these metrics adequately account for disciplinary contexts. A qualitative criteria catalogue, serving as the study's coding framework, was developed to systematically compare FAIR assessment tools regarding their operational logic, assessment scope, and referenced standards. A structured survey of the FAIR assessment landscape identified 34 tools, 14 of which were analyzed in detail. The results show that most tools rely on discipline-agnostic characteristics that primarily evaluate the infrastructural and machine-readable attributes of digital objects, while discipline-specific practices are rarely addressed.

Keywords: FAIR Assessment, Interdisciplinary Data Reusability, Discipline-agnostic Metrics, Criteria Catalogue, Coding Framework

1 Introduction

Data-driven academia has resulted in a variety of concepts of data, knowledge infrastructures and domain-specific research practices as Borgman [6], amongst others, already stated in the mid-2010s [15, 30]. In the same decade a diverse group of stakeholders suggested the FAIR Data Principles: a set of 15 indicators for "scholarly digital objects of all kinds to become 'first class citizens' in the scientific publication ecosystem" to turn research data into contextualized information objects [32]. Consequently, a key motivation of FAIR—and the primary motivation for this study—is the possibility of aggregating data across disciplinary boundaries. Interdisciplinary research questions require linking heterogeneous datasets at a semantic level, such as environmental and regional statistics. Interconnected data enable efficient interdisciplinary research.

However, the challenge lies in the tension between generic, discipline-agnostic infrastructures and the highly specialized *communities of practice*[1] that produce and consume the research data [8, 11, 34].

[1] For analytical purposes, disciplinary contexts are treated as institutionalized proxies for communities of practice. While communities of practice may transcend disciplinary boundaries, disciplinary metadata standards provide the most visible manifestation of community-specific information practices in FAIR assessment tools.

FAIR assessment tools are used to measure the FAIRness of scholarly digital objects. The credibility and adoption in actual usage of scientific data infrastructure are strongly intertwined with the FAIRness of digital objects stored in those services [11]. Prior works analyzed FAIR assessment tools for usability or technical implementation [3, 4, 11, 24]. Less attention has been paid to how the tools operationalize the FAIR Data Principles and whether their assessment metrics reflect discipline-specific practices. This study argues that unless assessment tools account for these discipline-specific characteristics, they will continue to certify data as FAIR in isolation, while the data remains interoperable only within its own community of practice. By making the discipline-agnostic design choices visible, this work provides a structured framework that can support researchers, data stewards, and infrastructure operators in selecting and combining tools for interdisciplinary data environments.

The approach is grounded in the idea that evaluation instruments themselves produce information [10]. Assessment tools therefore do not only measure FAIRness; they also shape how FAIR is comprehended in research infrastructures and associated communities of practice. Therefore, this study addresses two research questions:

1. Which criteria allow for a qualitative comparison of FAIR assessment tools and their metrics regarding their suitability for interdisciplinary data aggregation?
2. How do FAIR assessment tools operationalize the FAIR principles, and to what extent are their metrics linked to discipline-specific practices?

To address these questions, the study employs an analytical design: a criteria catalogue is developed to compare FAIR assessment tools qualitatively. The resulting comparison provides the empirical basis for identifying tool-inherent discipline-specific characteristics and discussing their implications for FAIR operationalization across research communities. This paper contributes by:

1. proposing a systematic qualitative criteria catalogue for FAIR assessment tools,
2. providing a reproducible comparison of 14 existing tools,
3. conceptualizing discipline awareness as a missing dimension of FAIR assessment.

Following the introduction, section 2 presents the theoretical background on the FAIR Data Principles as an information model and outlines the operational modes and target audiences of FAIR assessment tools. Section 3 describes the methodological approach, including the criteria catalogue, tool selection, and methodological limitations. Section 4 presents the results. Section 5 discusses the findings, and 6 concludes the paper.

2 Theoretical Background

The transition toward data-driven academia has necessitated a fundamental shift in how digital research objects are managed, shared, and integrated. This section establishes the conceptual framework for the study by distinguishing between (i) FAIR as an abstract information model, (ii) frameworks that operationalize the FAIR Data Principles, and (iii) FAIR assessment tools that implement these operationalizations in practice.

2.1 The FAIR Data Principles as Information Model

In this study, the FAIR Data Principles (Findable, Accessible, Interoperable, Reusable) are understood as an abstract information model that defines the conditions under which digital research objects can be discovered, accessed, combined, and reused. Rather than prescribing technical implementations, FAIR specifies what should be achieved while remaining intentionally independent of disciplinary domains and technological infrastructures [5, 32].

Accordingly, FAIR focuses on the relationship between digital objects, metadata, and infrastructures. Digital objects should be interpretable by computational systems that can automatically discover, access, and process them across heterogeneous infrastructures [26, 31, 32]. At the same time, FAIR acknowledges that meaningful reuse depends heavily on metadata, documentation, and the technical environment in which data are embedded [13, 23]. Consequently, although FAIR itself is domain-independent, its implementation inevitably reflects community-specific research practices.

Unlike previous paradigms centered on openness and accessibility, FAIR emphasizes machine-actionability alongside intellectual reusability. As Hellmann and Forberg [14] argue, FAIR extends earlier open-data concepts by recognizing that data quality cannot be reduced to binary open/closed distinctions but instead requires a multidimensional assessment of findability, accessibility, interoperability, and reusability [3].

2.2 Operational Modes and Target Audiences of FAIR Assessment Tools

FAIR assessment tools assess the FAIRness score of examined units such as datasets, research software or publications. They are frequently presented as neutral instruments for evaluating data quality [25]. In practice, however, they mediate between researchers, data stewards, repositories, and research data management workflows, thereby shaping scholarly information spaces [16]. Bates [5] describes such information spaces as dynamic environments in which technological infrastructures and human interpretation influence information practices. Applied to FAIR, this perspective suggests that although infrastructures enable standardized assessment, meaningful evaluation and reuse remain dependent on community-specific knowledge and disciplinary conventions.

Because the FAIR Data Principles intentionally remain abstract, several initiatives have proposed measurable indicators to operationalize them. Early work by Wilkinson et al. [33] introduced FAIR Metrics as a first attempt to translate the principles into assessable criteria. Building on this foundation, subsequent frameworks such as the FAIRsFAIR Data Object Assessment Metrics, the RDA FAIR Data Maturity Model,

FAIRmetrics.org, and repository-specific approaches such as the DANS FAIR Metrics further refined these indicators for practical assessment. While these frameworks share the common objective of evaluating FAIRness, they differ in emphasis, granularity, and target communities. In addition to established frameworks, many tools rely on custom metrics derived from existing standards but adapted by developers or infrastructure providers. Since these adaptations are often insufficiently documented, it is not always transparent which assumptions, standards, or community-specific requirements guide the assessment logic. As previous studies have demonstrated, both of these aspects lead to divergent FAIRness scores of identical digital objects when evaluated by different assessment tools [34].

Existing tools generally follow two complementary assessment strategies. Automated assessment tools evaluate machine-readable characteristics, such as persistent identifier resolution, metadata availability, or schema compliance, thereby prioritizing scalability and reproducibility [24]. In contrast, self-assessment frameworks rely on manual questionnaires that enable contextual interpretation, particularly for Reusability (R), which requires human judgement regarding community standards and domain-specific requirements [1, 9]. FAIR acknowledges this dependency through Principle R1.3, which explicitly requires compliance with community standards [12, 32].

Rather than analyzing communities of practice directly, this study therefore examines how community-specific information practices become embedded in FAIR assessment tools. Scientific disciplines serve as an analytical proxy because these practices are commonly institutionalized through disciplinary standards and metadata conventions.

3 Methodology

This study employs a criteria catalogue for a qualitative comparative analysis of FAIR assessment tools, identifying their tool-inherent, discipline-specific characteristics through systematic evaluation. While Wilkinson et al. [32] define universal principles, this catalogue reconstructs their discipline-specific operationalization.

Following the introduction of the criteria catalogue, the tool selection is outlined.

3.1 Criteria catalogue

The criteria catalogue serves as the central data collection and structured analytical coding instrument of this study. Rather than functioning as a mere descriptive checklist, it provides a structured analytical framework that operationalizes qualitative comparison of FAIR assessment tools through predefined criteria, controlled vocabularies, and interpretation rules. Methodologically, the criteria catalogue is organized into three complementary layers:

- criteria define which characteristics are documented,
- predefined values ensure consistent coding wherever possible,
- and interpretation rules specify the analytical relevance of each criterion for identifying tool-inherent discipline-specific characteristics.

Rather than merely listing observable tool characteristics, the criteria catalogue operationalizes the comparative analysis by linking data collection (criteria), standardized coding (values), and analytical interpretation (interpretation rules). The criteria catalogue was developed iteratively. The process began with the collection of basic data on the tools, including purpose and technical characteristics. This initial step combined an analysis of tool usage with relevant publications describing their development. The next step focused on the assessment scope and metrics applied by the tools to analyze the resulting FAIRness scores qualitatively.

Lastly, the impact of FAIR assessment tools was considered as an additional dimension. However, because measuring impact presents substantial methodological challenges, it was discussed with several data experts at the Helmholtz Metadata Collaboration Conference (April 2026, Heidelberg, Germany) as part of a poster presentation [20]. The discussion highlighted that impact should be understood as a contextual and multidimensional concept rather than a directly measurable property of FAIR assessment tools. Impact measurement strongly depends on the specific dimension being considered and on the availability of reliable indicators [29]. In the context of FAIR assessment, potential signals of impact include dissemination and adoption within research practices, recommendations by (inter)national research data initiatives such as NFDI consortia, as well as references in project infrastructures or scholarly publications. However, several commonly used indicators were considered unsuitable for systematic comparison. For example, user statistics may be influenced by technical factors such as automated requests or infrastructure monitoring, which limits their reliability. More fundamentally, impact measurement in this context represents a form of value analysis, as the scientific and infrastructural value of FAIR assessment tools remains an ongoing topic of discussion within the research data community [8, 9, 11, 33, 34].

The finalized catalogue comprises four main categories to describe FAIR assessment tools: basic data, purpose & objectives, technical features, assessment scope & metrics.

Basic data captures a contextual description of the tool and functions as comparable metadata. *Purpose & objectives* summon information provided by the tool regarding its target audience and usage restrictions, drawn exclusively from the developer or publishing organization. *Technical features* describe the technical functionality for a practical overview of the tool. *Assessment scope & metrics* address the number and type of FAIR aspects assessed. All criteria, organized by those four main categories, are displayed in Figure 1.

Basic data and *technical properties* served to classify tools according to their operation mode. To reveal tool-inherent discipline-specific characteristics the main categories *purpose & objectives* and *assessment scope & metrics* were prioritized, placing special emphasis on the following criteria: tools operation methods, examined unit and granularity, degree of automation and scoring system, publishing organization, community of practices as targeted audiences, and referenced standards & schemas.

Each criterion consists of three components: (i) a criterion description defining the information to be collected, (ii) where applicable, a predefined set of values (controlled vocabulary), and (iii) an interpretation specifying the analytical relevance of the criterion for identifying tool-inherent, discipline-specific characteristics. Together, these

components form a rule-based coding scheme that guided the qualitative comparison of all analyzed tools.

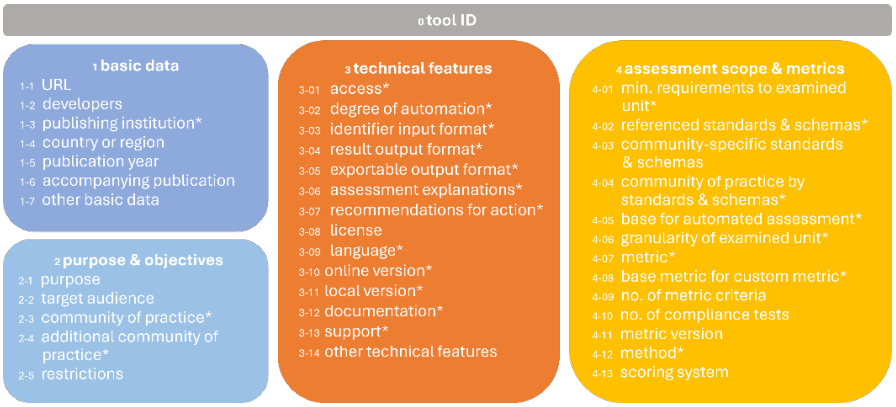

Figure 1: Criteria catalogue for qualitative comparison of FAIR assessment tools. Criteria organized by main category, * denotes controlled vocabulary.

Controlled vocabulary was used wherever possible to reduce interpretative ambiguity during coding and to ensure consistent assignment of observed tool characteristics. Free-text responses were restricted to criteria for which no meaningful standardized set of values could be defined. Overall, 59 values are either developed through iterative data collection or follow Boolean logic (yes/no). 48 values are grounded in peer-reviewed articles or in information published by research data infrastructures and funding agencies (such as re3data[2] or DFG[3]). While 26 criteria utilize controlled vocabulary values, fourteen criteria require free-text responses.

To enhance transparency and reproducibility, the complete criteria catalogue, including definitions of criteria, controlled vocabularies, interpretation rules, and raw coding data, has been published as an openly accessible research dataset[4] [21]. This dataset includes a detailed README to ensure full reproducibility and reusability of this comparative analysis.

Specifically, the classification of tools as discipline-specific, tending towards discipline-specific, discipline-agnostic, or neither was not based on subjective overall impressions. Rather, it followed the predefined interpretation logic of the criteria catalogue [21]. Classifications derived from evidence regarding discipline-specific metadata standards, supported data types, configurable assessment logic, and references to community-specific information models. Tools were assigned to the respective category only when these indicators were explicitly documented. The criteria catalogue

[2] https://www.re3data.org/browse/by-subject/

[3] German Research Funding Agency: https://www.dfg.de/resource/blob/331944/fachsystematik-2024-2028-de.pdf

[4] For quick access use: https://doi.org/10.5281/zenodo.19943297. Detailed information is provided in the reference section.

operationalizes the concept of tool-inherent discipline-specificity by translating abstract characteristics into observable coding categories and predefined interpretation rules. For analytical purposes, discipline-specificity is treated as an interpretative category derived from observable tool characteristics documented in the criteria catalogue.

The catalogue was not only used to document tool characteristics but also to transform heterogeneous tool descriptions into comparable analytical observations. By assigning standardized values to predefined criteria, qualitative information from documentation, interfaces, and publications became systematically codable across all tools. This coding process enabled subsequent interpretation of discipline-specific characteristics based on explicit evidence rather than overall impressions.

3.2 Tool selection

Through structured investigation, all available FAIR assessment tools were identified in March 2026[5]. The search began with international FAIR information infrastructures such as FAIRassist.org[6] and FAIR-IMPACT[7]. While both initiatives are related to the European Open Science Cloud (EOSC), FAIRassist.org serves as a specialized registry for community-specific FAIR tools. In contrast, FAIR-IMPACT represents a major EU-funded Horizon Europe project expanding FAIR solutions across EOSC. Both platforms provide comprehensive overviews of FAIR assessment tools. The search was supplemented by reviewing additional tool collections such as FAIR Metroline[8], Lang et al. [19][9], FAIRsFAIR Tools & Software[10] and concluded with targeted web searches using combinations of keywords such as “FAIR assessment tool”, “FAIR evaluation” and “FAIR metrics”.

In total, 34 FAIR assessment tools were identified [21]. Initial screening excluded tools that were no longer accessible, discontinued or did not fully implement or center FAIR principles into the assessment. Those “out-of-scope” tools were deemed unsuitable for qualitative comparison. For instance, KGHeartBeat evaluates Knowledge Graphs for representation, accessibility, and trust rather than FAIRness of digital objects. Similarly, the Data Stewardship Wizard/FAIR Wizard assesses Data Management Plan (DMP) FAIRness rather than digital object compliance.

The remaining 16 tools received unique IDs and entered the initial data collection. The coding followed the operationalization defined in Section 3.1. Subsequently, two additional tools (*FAIR enough* and *Fair EVA*) were excluded due to unavailable service, despite partial documentation. Both tools appear highlighted in red in figure 2. This color coding depends on the level of analysis the tools enable: Priority 1 (green: qualitative analysis + quantitative analysis are possible), Priority 2 (yellow: qualitative only), Priority 3 (red: unsuitable for analysis).

---

[5] FAIR assessment tools released after the 19th of March 2026 could not be included in analysis.
[6] https://fairassist.org/tools
[7] https://fair-impact.eu/fair-assessment-tools
[8] https://fairmetroline.org/metroline_steps/pre_fair_assessment
[9] https://zenodo.org/records/7701941
[10] https://www.fairsfair.eu/tools-software

Quantitative, FAIRness score-based comparison of these tools represents future research beyond this paper's scope. Such analysis requires automated assessment capabilities and standardized input formats. Accordingly, quantitative-capable tools must satisfy all catalogue criteria plus specific values for criterion {3-2} (degree of automation: semi-automated or automated) and criterion {4-6} (granularity of examined unit: digital object).

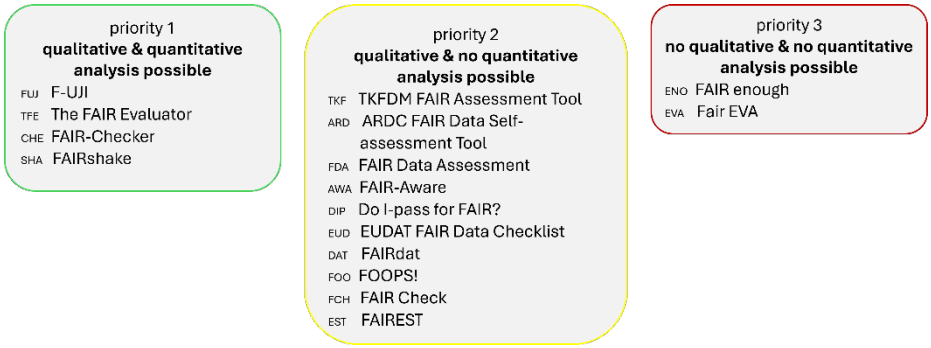

Figure 2: Analyzed FAIR assessment tools and associated tool IDs used during data collection. Priorities indicate the feasibility of qualitative and/or quantitative comparison.

The following section presents results for the remaining 14 FAIR assessment tools meeting priority 1 and 2 (listed in green and yellow boxes).

3.3 Limitations of the approach

This study does not measure the empirical FAIRness scores generated by the tools, nor does it assess their performance on specific datasets. Instead, it analyzes how tools are designed to operationalize FAIRness.

As the coding was conducted by a single researcher, explicit decision rules are provided to enhance transparency and reproducibility [21]. The resulting classifications should therefore be understood as analytically grounded rather than statistically validated. Although the criteria catalogue was iteratively refined and discussed with domain experts, interpretative bias cannot be fully excluded.

The analysis relies on publicly available documentation and observable tool behavior. Undocumented characteristics may therefore not be fully represented.

The FAIR assessment ecosystem is evolving rapidly, and the results reflect the state of the landscape at the time of the study (March 2026).

While some tools rely on manual self-assessment and may allow discipline-specific interpretations during use, this study focuses on discipline-agnostic characteristics embedded in tool design.

4 Results

This section presents findings from the qualitative analysis of 14 FAIR assessment tools conducted using a systematic criteria catalogue. The structure mirrors the criteria catalogue described in Section 3.1: basic data and purpose (4.1), technical properties (4.2), assessment scope and metrics (4.3), and communities of practice (4.4).

Key findings reveal the prevalence of custom metrics over established standards during (semi-)automated assessment, and manual questionnaires as the predominant operation mode. Comparing targeted communities of practice with referenced standards and schemas provide initial indicators of discipline-specific characteristics, even though most tools assess digital objects in a discipline-agnostic manner. Detailed variations in automation, metrics, and standards highlight divergent operationalizations of the FAIR Data Principles. Two tools (*FAIR Data Assessment, FAIR Check*) exhibit discipline-specific requirements, two tools (*F-UJI, FAIR-Checker*) show tendencies toward discipline-specific characteristics, nine tools are discipline agnostic, and one tool (*FAIRshake*) operates outside this dichotomy, meaning it is neither discipline specific nor discipline agnostic (see section 4.2, paragraph 2).

These findings address both research questions concerning criteria for tool comparison and discipline-specific characteristics.

4.1 Basic data and purpose of tool

FAIR assessment tools development appears predominantly European, with only two tools (*FAIRshake, ARDC FAIR Data Self-assessment Tool*) originating in non-European regions. No tools predate 2017, demonstrating the immediate impact of the FAIR Data Principles.

Tool purpose generally falls into four categories: monitoring, optimization, quality assurance, and data aggregation. Monitoring and optimization of datasets and software (examined unit {4-01}) predominate. Four tools emphasize quality assurance as one of the main purposes thus aligning recommendations in the initial FAIR Data Principles. Three tools (*FAIR Data Assessment, The FAIR Evaluator, FAIRshake*) target data aggregation without detailed specification of associated expectations.

Publishing institutions (criterion {1-3}) provide deeper insights. Infrastructure projects {1-3A}, funding organizations {1-3B}, universities {1-3D} and (inter)national research data initiatives {1-3E, 1-3F} typically indicate discipline-agnostic approaches. Conversely, community- and network-driven developments {1-3C} suggest potential discipline-specific needs. For non-university research institutions {1-3G} and research infrastructure/repository operators {1-3H}, closer examination of development motivations is required. Figure 3 illustrates organizational distribution among FAIR assessment tool development. Blue organizations indicate discipline-specific tool requirements, orange and yellow organizations represent discipline-agnostic approaches, and green organizations do not fall into this classification. Publishing organizations alone do not constitute sufficient evidence of tool-inherent discipline-specific characteristics but provide initial indications of target disciplines during development.

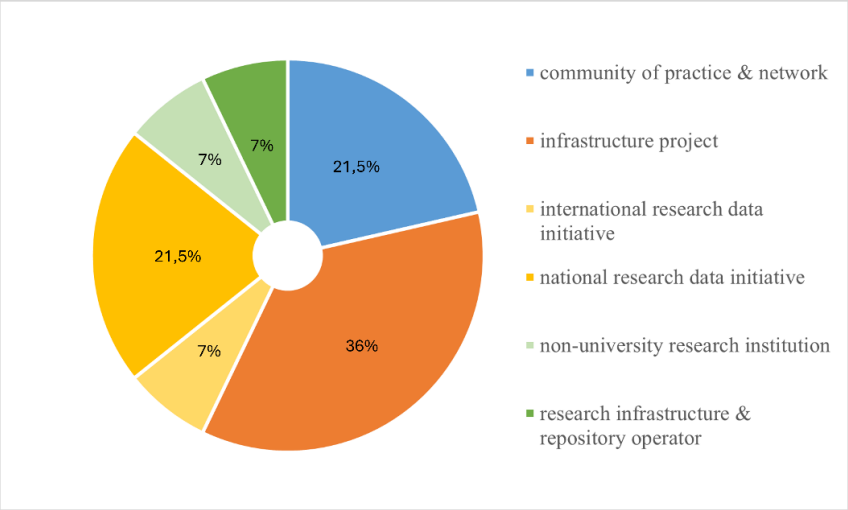

Figure 3: Distribution of publishing organizations developing FAIR assessment tools (numbers specify tools per organizational cluster).

4.2 Technical properties

Technical properties of the 14 analyzed FAIR assessment tools reveal no tool-inherent discipline-specific characteristics. However, notable peculiarities emerge from data and merit discussion.

Only one tool (*FAIRshake*) requires login-restricted access {3-1}, tracing back to its operation mode: *FAIRshake* mandates user-defined "rubrics" containing arbitrary metric criteria linked to user profiles. While predefined rubrics exist, user-created rubrics render FAIR assessment fully customizable, creating a tool which is neither discipline specific nor discipline agnostic but depends on the user 's FAIR expertise.

FAIRshake uniquely offers semi-automated assessment, combining automated evaluation with manual questionnaire. Four tools provide full automation, while most tools (9 of 14) rely entirely on manual questionnaires. Self-assessment inherently depends on user interpretations of the FAIR Data Principles, introducing user-generated discipline-specific characteristics. Crucially, these characteristics are not tool-inherent even though the tool 's degree of automation {3-2} enhances that effect.

Automated tools universally require some sort of identifier input format {3-3} such as a PID or URL, while only one self-assessment tool (*FAIRdat*) mandates URLs – likely for tracking tool usage. This corresponds to additional parameters (name of reviewer, name of repository and date of assessment) collected by *FAIRdat*, despite no technical necessity.

Five tools (5 of 14) offer result export {3-4} in either JSON, PDF, CSV, or DOCX formats. Six tools present assessment explanations {3-6}; only two tools (*FAIR-Checker, FOOPS!*) provide result-tailored recommendations for action {3-7}.

All analyzed tools operate exclusively in English, underscoring dominance of Englisch in scientific communication and natural language processing infrastructure

[18, 22]. Notably, *FAIRshake* supports multilingualism based on underlying metrics and rubrics.

Ten tools (10 of 14) can be run as a local version {3-11} (e.g. private device) via GitHub-hosted open-source code and/or downloadable questionnaires.

4.3 Assessment Scope and Metrics

Ten tools (10 of 14) assess digital objects such as datasets, software or publications. Two tools evaluate organizational FAIRness (*Do I-pass for FAIR?, FAIREST*), one targets ontologies and vocabularies (*FOOPS!*) and one tool (*FAIR-Aware*) tests user's FAIR knowledge. Minimum requirements for examined unit {4-1} reveal no tool-inherent discipline-specific characteristics.

Remarkably, 12 of 14 tools implemented a custom metric {4-7} rather than adopting established frameworks. Although most custom metrics are built on established standards, the reasons for creating a custom metric, as well as the adjustments made, remain largely non-transparent. Both issues merit further investigation.

Established frameworks, which act as base metrics and underlying custom implementations, are detailed in section 2.2. Figure 4 reveals metric preferences among custom metric developers.

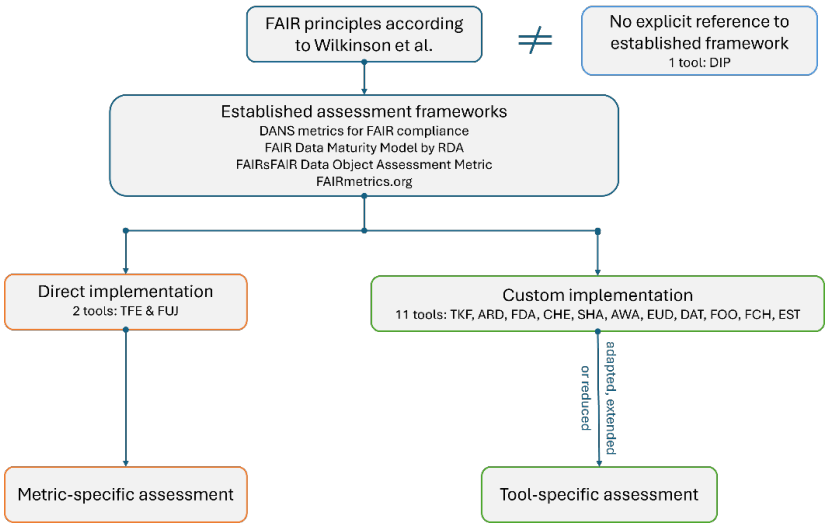

Figure 4: Implementation of FAIR assessment logic across the analyzed tools. While most tools adapt established FAIR metrics through tool-specific assessment logic, TFE and F-UJI directly implement established FAIR frameworks. One tool does not explicitly reference established FAIR assessment frameworks.

Six of 14 tools base assessments on the initial FAIR Data Principles, all utilizing self-assessment (online/offline). Only one tool (*FAIR Check*) associated itself with discipline-specific characteristics.

FAIRshake and *The FAIR Evaluator* employ FAIRmetrics.org for its assessment.

Two tools utilize RDA´s FAIR Data Maturity Model: *FAIR-Checker* employ RDA indicators exclusively, while *FOOPS!* combines them with FAIRsFAIR Data Object Assessment Metrics. Both tools use the same number of FAIR compliant tests: 24. *FAIR-Checker* is anchored in life science; *FOOPS!* appears to be discipline agnostic.

FAIR Data Assessment and *FAIRdat* both employ DANS metrics for FAIR compliance but differ in criteria and indicator count.

Do I-pass for FAIR? uniquely developed a metric from scratch, assessing organizational maturity rather than FAIR compliance of digital objects [7].

F-UJI employs FAIRsFAIR Data Object Assessment Metrics. Notably, *F-UJI* and *The FAIR Evaluator* publishing institutions not only developed the associated automated FAIR assessment tool but also the metric those tools are operated on. Although twelve of the fourteen analyzed tools rely on custom metrics, these are not independent assessment frameworks. Instead, they represent tool-specific adaptations of established FAIR metrics and indicators. Only The FAIR Evaluator (TFE) and F-UJI (FUJ) implement established assessment frameworks without adjusting them.

Additionally, a significant gap exists between the number of compliance tests {3-10} for (semi-)automated tools and the number of metric criteria {3-9} for self-assessment questionnaires. These numbers vary from 10 to 24 FAIR criteria for self-assessment up to over 100 indicators when it comes to (semi-)automated assessment. This disparity reflects operational differences: (semi-)automated tools efficiently process 100+ tests rapidly, whereas self-assessments become impractical at scale.

Three of six (semi-)automated tools (*F-UJI, FAIR-Checker, FOOPS!*) operate on a static, metric-based method {4-12}. Despite differences (e.g., community-specific standards & schemas, base for automated assessment, no. of metric criteria, compliance tests), all these tools share the base metrics (RDA, FAIRsFAIR) explaining their static operation. The other half of the (semi-)automated tools (*The FAIR Evaluator, FAIRshake, FAIRdat*) use the dynamic, indicator-based method. *The FAIR Evaluator* and *FAIRshake* allow users to adjust indicators and compliance tests before or during the FAIR assessment. Therefore, both tools are neither discipline specific nor discipline agnostic since they can be adapted to the discipline the assessment is needed for.

(Semi-)automated assessments use two sources: (i) the landing page a PID links to or (ii) structured metadata the examined unit provides. *F-UJI* combines the landing page information with associated repository API; *FOOPS!* uses APIs only which makes it efficient for repository FAIRness.

Even though some FAIR assessment tools share base metrics and similar test counts, they differ in operation mode and referenced standards and schemas.

4.4 Community of practices as stated by the tool developers and by standards and schemas suggested

Along with each assessment tool, most publishing institutions or developers documented purpose, metrics, target audience, etc. in accompanying publications {1-6}. These publications differ in level of detail. Three tools provide no specific paper but background information at the tool´s URL. Overall, purpose & objectives were summarized for all tools, focusing on developer-stated information.

As noted in section 2.2, all FAIR assessments incorporate standards and schemas. Some tools reference discipline-specific standards and schemas, despite appearing discipline agnostic.
Here, stated community of practices ({2-3}, {2-4}) were compared to referenced standards and schemas ({4-3}, {4-4}). Currently, no automated assessment tool tests for specific standards and schemas; instead, they refer to them in either technical or user documentation.

Notably, 7 of 14 tools self-identify as discipline agnostic, yet two of these tools referenced discipline-specific standards and/or schemas: *F-UJI* refers to TEI, which is a widespread standard for any kind of text but developed for cultural heritage texts and above all located in Humanities and Social Science. *F-UJI* refers to many more discipline-agnostic standards and schemas and does not test automatically for TEI standards.

FAIR-Checker explicitly references Life Science standards (Bioportal, Bioschemas) and is developed by an interoperability working group of the French Institute for Bioinformatics. Yet again: these are referenced to, not automatically tested.

Five tools omit community specification and therefore do not provide background information whether a tool is developed discipline specific or discipline agnostic. This could not be (dis)proven by referenced standards and schemas.

Two tools explicitly target specific disciplines. *FAIR Data Assessment* belongs to Life Sciences, more precisely: agricultural science. The tool references standards such as Agrovoc and AgriMetaMaker, while still welcoming cross-disciplinary use, without noted limitations. *FAIR Check*, developed by NFDI4Culture, optimizes Humanities and Social Science assessment, without referenced standards and schemas.

5 Discussion

This study's core finding, derived from qualitative analysis of 14 FAIR assessment tools using the systematic criteria catalogue described in Section 3.1, points to a strong tendency toward discipline agnosticism and with it a significant discrepancy between the universal aspirations of the FAIR Data Principles and their practical operationalization. 12 of 14 tools employ custom metrics rather than established frameworks {4-7}, confirming Wilkinson et al.'s observation of inconsistent assessment outcomes as technically driven rather than discipline specifics [34]. The predominance of discipline-agnostic assessment logic suggests a potential limitation for interdisciplinary data aggregation, particularly where successful reuse depends on discipline-specific metadata and semantic context. While developers create custom metrics to better align with specific project goals or perceived community needs, the lack of transparency regarding these modifications (as identified in section 4.3) creates a "black box" effect. This is further underscored by requirements regarding what constitutes "good" metadata. From an information science perspective, custom metrics function as local benchmarks, not global interoperability indicators for interdisciplinary data aggregation. This prevents the standardization and, ultimately, reusability of FAIRness scores across different assessment tools, without a deep dive into the underlying metric logic.

FAIR assessment tools favour machine-readable structures ({4-6}: digital objects dominate), meaning that tools effectively measure "infrastructural readiness" rather than "scientific reusability". This leads to a situation where data can be technically FAIR according to an automated tool yet remain practically useless for researchers due to the lack of deep semantic contextualization [17, 28]. Automated tools (4 of 14) verify 100+ indicators efficiently but check "only [...] name and type of […] standards, but not its structure or formulation unless these are 'hard-coded' into the tool" [33, 34]. This observation highlights the need for a continuous balancing between discipline-specific data practices, which may rely on community-driven and less standardized metadata structures, and the goal of enabling reuse across disciplinary boundaries.

The results indicate an ongoing trend toward discipline agnosticism. This pattern has also been observed in previous studies [8] and can interpreted as "lowest common denominator" approach. While this facilitates the movement of data into large-scale infrastructures like the EOSC, it risks stripping data of the granular metadata required for complex, cross-domain queries. As Appe [2] states, aggregation often leads to trade-offs regarding context-specific phenomena. If discipline-specific metadata requirements remain excluded from automated assessment, FAIRness scores may insufficiently reflect the semantic context required for interdisciplinary reuse.

These findings indicate that future FAIR assessment approaches could benefit from complementing metadata presence checks with mechanisms that evaluate the compatibility of underlying information models.

One possible implication of these findings is that funding agencies and infrastructure operators could consider complementing universal FAIR assessments with configurable, community-oriented assessment components. This could lead to a two-step assessment strategy:

- First Agnostic Step: Use automated tools (e.g., F-UJI) for baseline infrastructural compliance (PIDs, Licenses, standard APIs).
- Second Disciplinary Step: Use community-validated, indicator-based tools (e.g., FAIRshake) that can configured with discipline-specific "rubrics".

In general, FAIR assessment is most successful if the metric used is defined or at least selected by users to tailor down FAIR assessment to community of practice in question. Furthermore, following Poveda et al. [27], a strategic portion of research funding should be dedicated not just to data production, but to the maintenance of the semantic infrastructures (standard registries and disciplinary FAIR metrics) that make these assessment tools meaningful.

From the outset, it must be acknowledged that the FAIR Data Principles were designed primarily to improve machine-readability, not necessarily to resolve the complexities of semantic aggregation. This leads to what may be described as a potential "Checklist Paradox": a digital object may achieve a perfect FAIRness score through technical compliance (syntactic interoperability) while remaining semantically opaque to interdisciplinary or even disciplinary researchers. The present comparative study provides empirical evidence of such tensions.

9 of 14 tools rely on manual self-assessments dependent on user interpretation of FAIR Data Principles, introducing non-tool-inherent discipline-specific characteristics

enhanced by tool design [8]. As identified in the results section, the burden of discipline-specifics is carried by the user rather than being tool-inherent. This highlights the need for assessment workflows that combine the efficiency of automated metadata checking (as seen in *F-UJI* or *FOOPS!*) with the nuanced, context-aware validation that only a human-in-the-loop or a discipline-specific knowledge graph can provide [3]. A (semi-)automated assessment tool like *FAIRshake* can counteract this phenomenon.

Result-driven heterogeneity reflects Bates' [5] dynamic information spaces: Automated tools appear domain-neutral through machine-readable focus; manual assessments carry "the burden of discipline-specifics created by the user" – not tool-inherent, though design-amplified. Custom metrics (12 of 14) create implicit discipline bias: *F-UJI* references TEI (Humanities-associated) without automated testing; *FAIR-Checker* cites Bioportal/Bioschemas (Life Sciences) only in its documentation.

The prevalence of discipline-agnostic tools suggests that in striving for universal applicability, the current FAIR ecosystem has sacrificed the granular, domain-specific metadata required for true knowledge integration. Therefore, this comparison is not just a tool evaluation; it is a fundamental critique of a "checklist culture" that mistakes structural metadata presence for meaningful scientific communication. Without addressing this gap, FAIR remains a measure of archival readiness rather than a facilitator of cross-disciplinary discovery.

6 Conclusion

By developing a qualitative criteria catalogue, this study provides a systematic reconstruction of the disciplinary suitability of FAIR assessment tools and their potential for interdisciplinary data aggregation. Developed iteratively, the catalogue enables a structured analysis of how FAIR assessment tools operationalize the FAIR Data Principles and which requirements about research data practices they embed.

Applying this catalogue to the analysis of 14 tools reveals that discipline-agnosticism appears to reflect a deliberate design choice. Their metrics primarily evaluate infrastructural and machine-readable characteristics of digital objects, while discipline-specific metadata practices and semantic contexts are rarely addressed explicitly. The technical foundations for FAIR assessment are maturing; the semantic layer remains underdeveloped. The widespread use of custom metrics (86% of the analyzed tools) further contributes to heterogeneous interpretations of FAIRness across tools.

By uncovering the characteristics embedded in FAIR assessment tools, this study contributes a structured analytical framework. The proposed criteria catalogue helps researchers, data stewards, and infrastructure operators better understand how different tools operationalize FAIR.

The results indicate that effective FAIR evaluation in interdisciplinary environments may require combining automated infrastructure checks with discipline-aware assessment approaches, ensuring that research data becomes a true "first-class citizen" in the global scientific publication ecosystem.

Acknowledgement
I am grateful to Fachhochschule Potsdam (FHP) for supporting this doctoral research. My sincere thanks go to my supervisors, Prof. Dr. rer. nat. Heike Neuroth (FHP) and Prof. Vivien Petras, PhD (HU Berlin), for their insightful discussions that substantially shaped this work.

Disclosure of Interest
The author has no competing interests to declare that are relevant to the content of this article.

AI Disclaimer
AI used exclusively for linguistic subtleties, grammar and spelling correction. Used model was Perplexity, during the timespan 1st-15th April 2026 available under www.perplexity.ai.

References all links last accessed on 29th April 2026

1. Ammar, A., Bonaretti, S., Winckers, L., Quik, J., Bakker, M., Maier, D., Lynch, I., Van Rijn, J., Willighagen, E., 2020. A Semi-Automated Workflow for FAIR Maturity Indicators in the Life Sciences. Nanomaterials 10, 2068. https://doi.org/10.3390/nano10102068
2. Appe, S., 2025. Commentary on Part IV - Data Aggregation. In: Bassi, A., Aquino Alves, M., Cordery, C. (Eds.), The Future of Third Sector Research, Nonprofit and Civil Society Studies. Springer Nature Switzerland, Cham, pp. 179–185. https://doi.org/10.1007/978-3-031-67896-7_15
3. Assmann, C., Gerlach, R., Lang, K., Neute, N., Rex, J., 2023. FAIR Assessment Tools: An evaluation of assessment tools of data sets according to the FAIR principles. Bauhaus Universität Weimar, Universität Erfurt, Technische Universität Ilmenau. https://doi.org/10.11588/HEIDOK.00033145
4. Bahim, C., Dekkers, M., Wyns, B., 2019. Result of an Analysis of Existing FAIR Assessment Tools. https://doi.org/10.15497/rda00035
5. Bates, M.J., 2002. Toward an Integrated Model of Information Seeking and Searching, in: New Review of Information Behaviour Research. Presented at the Fourth international Conference on information Needs, Seeking and Use in Different Contexts, pp. 1–15. URL https://pages.gseis.ucla.edu/faculty/bates/articles/info_SeekSearch-i-030329.html
6. Borgman, C.L., 2015. Big data, little data, no data: scholarship in the networked world. The MIT Press, Cambridge, Massachusetts.
7. Bruin, T.D., Coombs, S., Jong, J.D., Haslinger, I., Hoogen, H.V.D., Huigen, F., Jetten, M., Koster, J., Miedema, M., Öllers, S., Slouwerhof, I., Verheul, I., Ringersma, J., 2020. Do I-PASS for FAIR. A self assessment tool to measure the FAIR-ness of an organization. https://doi.org/10.5281/ZENODO.4080867
8. Candela, L., Mangione, D., Pavone, G., 2024. The FAIR Assessment Conundrum: Reflections on Tools and Metrics. CODATA 23, 33. https://doi.org/10.5334/dsj-2024-033
9. Devaraju, A., Huber, R., 2021. An automated solution for measuring the progress toward FAIR research data. Patterns 2, 100370. https://doi.org/10.1016/j.patter.2021.100370
10. Floridi, L., 2011. The philosophy of information. Oxford university press, Oxford.

11. 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
12. GO-FAIR, n.d. R1.3: (Meta)data meet domain-relevant community standards. go-fair.org. URL https://www.go-fair.org/fair-principles/r1-3-metadata-meet-domain-relevant-community-standards/
13. Greenberg, J., 2017. Big Metadata, Smart Metadata, and Metadata Capital: Toward Greater Synergy Between Data Science and Metadata. Journal of Data and Information Science 2, 19–36. https://doi.org/10.1515/jdis-2017-0012
14. Hellmann, S., Forberg, J., 2026. FAIR Reality Check: Cost–Benefit and Tooling for Effective Research Data Infrastructure. URL https://helmholtz-metadaten.de/events/fair-reality-check
15. Ingwersen, P., Järvelin, K., 2005. The Turn, The Information Retrieval Series. Springer-Verlag, Berlin/Heidelberg. https://doi.org/10.1007/1-4020-3851-8
16. Krans, N.A., Ammar, A., Nymark, P., Willighagen, E.L., Bakker, M.I., Quik, J.T.K., 2022. FAIR assessment tools: evaluating use and performance. NanoImpact 27, 100402. https://doi.org/10.1016/j.impact.2022.100402
17. Krefeld, T., 2020. FAIR – Forschungsethik im Web. Lehre in den Digital Humanities. URL https://www.dh-lehre.gwi.uni-muenchen.de/?p=181800&v=2
18. Labra Gayo, J.E., Kontokostas, D., Auer, S., 2015. Multilingual linked data patterns. SW 6, 319–337. https://doi.org/10.3233/SW-140136
19. Lang, K., Assmann, C., Neute, N., Gerlach, R., Rex, J., 2023. FAIR Assessment Tools Overview. https://doi.org/10.5281/ZENODO.7701941
20. Lehmann-Gnirck, A., 2026. FAIR Assessment Tools as Supporters for Interdisciplinary Reseach. Helmholtz Metadata Collaboration Conference 2026, Heidelberg. Book of Abstracts p. 39. https://events.geomar.de/event/884/book-of-abstracts.pdf.
21. Lehmann-Gnirck, A., 2026. Criteria catalogue and tool selection for systematic analysis of FAIR Assessment Tools (V1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.19943297
22. Li, K., Zheng, X., Ni, C., 2024. The anglicization of science in China. Research Evaluation 33, rvae036. https://doi.org/10.1093/reseval/rvae036
23. Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L.O.B., Wilkinson, M.D., 2017. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. Information Services and Use 37, 49–56. https://doi.org/10.3233/ISU-170824
24. Patra, P., Pompeo, D.D., Marco, A.D., 2025. An Evaluation Framework for the FAIR Assessment tools in Open Science. https://doi.org/10.48550/arXiv.2503.15929
25. Pellegrino, M. A. & Tuozzo, G, 2025. How Fair is FAIR? Understanding LOD Cloud FAIRness Through Correlation Patterns. In: CIKM '25: Proceedings of the 34th ACM International Conference on Information and Knowledge Management, 2315-2325. https://doi.org/10.1145/3746252.3761092
26. Petras, V., 2024. The identity of information science. JD 80, 579–596. https://doi.org/10.1108/JD-04-2023-0074
27. Poveda, L., Farrell, G., Tosatto, S.C.E., Zahn-Zabal, M., Ruch, P., Gobeill, J., Waterhouse, R.M., Dessimoz, C., 2026. The missing link in FAIR data policy: biodata resources in life sciences. Sci Data 13, 373. https://doi.org/10.1038/s41597-026-06690-w

28. Rümmler, A., Figgemeier, H., Henzen, C., 2022. Lösungsansätze zur automatisierten Erfassung und Weiterverarbeitung von strukturierten Provenance-Informationen in Forschungsdateninfrastrukturen am Beispiel von Analyse-Workflows in R. Bausteine Forschungsdatenmanagement. https://doi.org/10.17192/BFDM.2022.1.8367
29. Teparek, M.; Metilli, D. 2026. Making FAIR Happen: Culture Change Across the Research Ecosystem. Helmholtz Metadata Collaboration Conference 2026, Heidelberg. Book of Abstracts p. 17. https://events.geomar.de/event/884/book-of-abstracts.pdf.
30. Tolle, K.M., Tansley, D.S.W., Hey, A.J.G., 2011. The Fourth Paradigm: Data-Intensive Scientific Discovery [Point of View]. Proc. IEEE 99, 1334–1337. https://doi.org/10.1109/JPROC.2011.2155130
31. Vogt, L., Strömert, P., Matentzoglu, N., Karam, N., Konrad, M., Prinz, M., Baum, R., 2025. Suggestions for extending the FAIR Principles based on a linguistic perspective on semantic interoperability. Sci Data 12, 688. https://doi.org/10.1038/s41597-025-05011-x
32. 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., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18
33. Wilkinson, M.D., Dumontier, M., Sansone, S.-A., Da Silva Santos, L.O.B., Prieto, M., McQuilton, P., Gautier, J., Murphy, D., Crosas, M., Schultes, E., 2018. Evaluating FAIR-Compliance Through an Objective, Automated, Community-Governed Framework. https://doi.org/10.1101/418376
34. Wilkinson, M.D., Sansone, S.-A., Marjan, G., Nordling, J., Dennis, R., Hecker, D., 2022. FAIR Assessment Tools: Towards an "Apples to Apples" Comparisons. https://doi.org/10.5281/ZENODO.7463421