Community Metrics and Automated Assessment of Machine-Actionable Data Management Plans

Tomasz Miksa[1][0000-0002-4929-7875], Andres Tabima Romero[1][0009-0002-4848-5089], Maria Kontopidi[2][0000-0002-9730-5542], and Elli Papadopoulou[2][0000-0002-0893-8509]

[1] TU Wien, Austria
[2] Athena Research Center, Greece

Abstract. Data Management Plans (DMPs) are mandated by funders as instruments for promoting FAIR data practice, yet evaluating their quality remains labour-intensive and does not scale. Machine-actionable DMP (maDMPs), structured according to the RDA DMP Common Standard, open the door to automated assessment, but a community-grounded, openly published catalog of evaluation criteria has so far been missing. This paper presents three contributions made within the OSTrails project. First, we describe a participatory co-design methodology through which practitioners, tool developers, and funder representatives elicited evaluation criteria grounded in real-world DMP workflows. Second, we present the resulting DMP Evaluation Metrics Catalogue that contains 94 metrics formalised as linked-data records using the FAIR Testing Results (FTR) vocabulary, anchored to RDA DCS schema fields. Third, we demonstrate automated evaluation of maDMPs against community-specific benchmarks, showing that a substantial portion of the catalogue can be applied as deterministic, externally validated tests returning machine-readable, field-level feedback for researchers, data stewards, and funders.

Keywords: Data Management Plans · DMP evaluation · machine-actionable DMPs · FAIR · OSTrails · research data management

1 Introduction

Data Management Plans (DMPs) have become a central instrument in research data governance. Originally conceived as administrative deliverables required by funders, DMPs are increasingly recognised as living documents that guide researchers through the data lifecycle: from planning and data collection to archiving and sharing. Funding agencies across Europe and beyond now mandate DMPs not merely as compliance artefacts but as tools to actively promote good data management practice, reproducibility, and adherence to the FAIR principles [18].

Yet this growing importance has exposed a fundamental tension: the quality and completeness of DMPs varies enormously, and the burden of evaluating them falls disproportionately on reviewers, whether at funding agencies, research performing organisations, or data stewardship services. Reviewers' guideliness are often not defined and the focus of the review differs a lot. In theory, they should assess not only funder-specific compliance requirements but also the scientific soundness, the feasibility, and the FAIRness of the data management approach described. This is a labour-intensive process that does not scale well, and it provides researchers with little actionable feedback.

Machine-actionable DMPs (maDMPs) [7] offer a path forward. The RDA DMP Common Standard (DCS) [10,11] provides a JSON-based application profile enabling DMPs to be processed programmatically across platforms, and Miksa et al. [8] showed how structured plans reduce administrative overhead and improve information quality. Early work introduced quality dimensions applicable to maDMPs [9,2] and prototype metrics instantiating them. What has remained missing is a community-grounded, openly published catalogue of evaluation criteria that practitioners can inspect, extend, and apply through shared tooling.

This paper presents three contributions made within the OSTrails project, which builds interoperable infrastructure connecting DMPs, scientific workflows, and FAIR Testing Results (FTR):

1. **A community co-design methodology** for eliciting and prioritising DMP evaluation criteria from practitioners, tool developers, and funders, grounded in real-world DMP workflows across European research infrastructure.
2. **A publicly available DMP Evaluation Metrics Catalogue**[3] of 94 metrics, formalised as linked-data records using the FTR vocabulary, operationalising community-elicited criteria as deterministic, automatable checks anchored to RDA DCS schema fields.
3. **Automated evaluation of maDMPs against community benchmarks**, demonstrating that the catalogue's metrics can be applied programmatically to produce structured, FTR-compliant results for researchers, data stewards, and funders.

Section 2 reviews background and related work. Section 3 describes the community co-design process. Section 4 presents the catalogue structure and formalisation. Section 5 demonstrates automated evaluation. Section 6 concludes with directions for future work.

2 Background and Related Work

2.1 Machine-Actionable DMPs and the RDA DMP Common Standard

The concept of maDMPs emerged from the recognition that text-based plans are opaque to automated processing and difficult to integrate with research infrastructures. Miksa et al. [8] articulated ten principles for maDMPs, arguing that

[3] `https://docs.ostrails.eu/en/latest/commons/dmp/dmp-evaluation-metrics.html`

structured, interoperable plans reduce administrative burden and support automated compliance monitoring. The RDA DMP Common Standard (DCS) [10] provided the technical foundation: a JSON-based application profile with a controlled vocabulary covering datasets, distributions, licences, contributors, repositories, and funding. Its mapping to linked-data representations [4] established the basis for interoperability between DMP tools and research infrastructure services. Papadopoulou et al. showed that DCS-conformant maDMP content can be published as linked open data within the OpenAIRE Research Graph, establishing DMPs as first-class entities alongside datasets and publications [15].

2.2 Quality Dimensions and Evaluation Frameworks

Miksa et al. [9] introduced an evaluation toolbox identifying quality dimensions for maDMPs together with prototype metrics. Arnhold et al. [2] proposed a more systematic taxonomy—completeness, openness, FAIRness, policy compliance, feasibility, adequacy, and consistency—and mapped each dimension onto RDA DCS fields. The present work adopts this taxonomy and refines it into five dimensions (merging openness into FAIRness, and adequacy into quality) to achieve a more parsimonious, mutually exclusive classification suited to automated testing. Papadopoulou et al. [16] translated these dimensions into a prototype operational evaluation rubric within the OSTrails project. The catalogue presented here extends that rubric through a full community co-design process and formalises it as openly accessible linked data.

2.3 Existing Evaluation Tooling and Its Limitations

DMP authoring tools such as DAMAP[4], Argos[5], or the Data Stewardship Wizard[6] offer template-driven guidance and basic completeness checks, but do not cross-reference declared values against external registries, validate persistent identifiers, or verify licence compliance against the funder requirements. Targeted services exist for individual checks, e.g. PID resolvers like DataCite or repository registries like re3data[7] or FAIRSharing[8], but no prior work has assembled these into a unified evaluation layer grounded in community-agreed criteria, mapped to a shared quality vocabulary, and returning machine-readable results consumable across authoring tools, submission portals, and knowledge graphs.

2.4 FAIR Testing and the OSTrails FTR Model

The FAIR Testing Results (FTR) vocabulary [14] defines classes and properties for expressing the outcome of a test applied to a digital object: the test itself, the

[4] `https://damap.org`
[5] `https://argos.openaire.eu/`
[6] `https://ds-wizard.org`
[7] `https://www.re3data.org`
[8] `https://fairsharing.org`

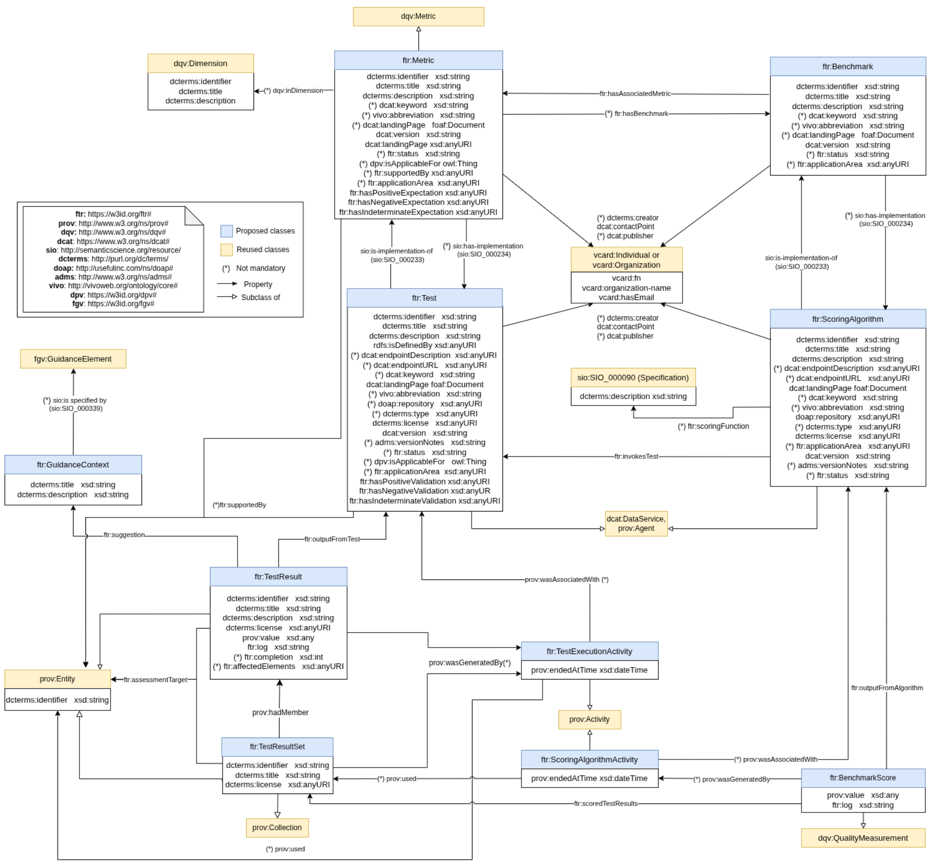

Fig. 1: FAIR Testing Results (FTR) vocabulary [14].

metric it operationalises, the subject, the outcome (pass, fail, or indeterminate), and full provenance. The vocabulary builds on DCAT [19], DQV [1], and PROV-O [12]. Within the OSTrails FAIR Reference Model [13], a *metric* is a narrative quality criterion and a *test* is the software artefact that operationalises it against a concrete digital object. This separation allows one metric to be implemented by multiple tests and allows results to be aggregated across platforms without loss of semantic precision. The catalogue reported here instantiates this model for maDMPs conforming to the RDA DCS.

3 Methodology for Metrics Elicitation

Figure 2 illustrates how the nature of assessable statements in a data management plan (DMP) evolves over the project lifecycle. In an initial DMP, statements are typically expressed in the future tense and describe planned actions or commitments, such as the intention to assign a CC-BY licence to a dataset. At

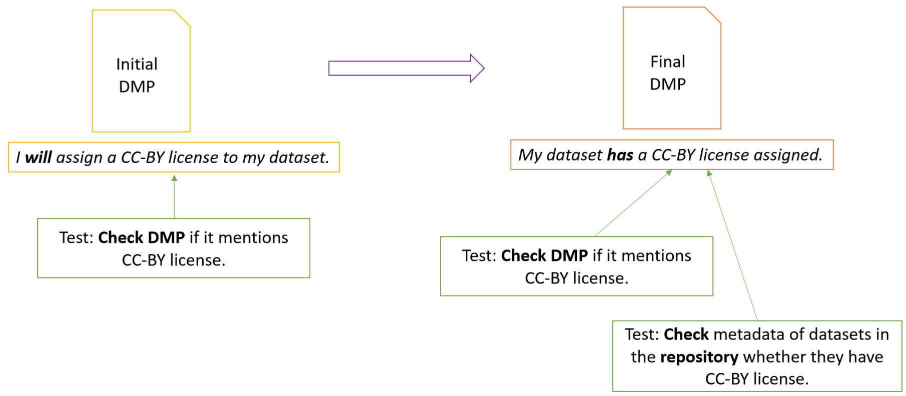

Fig. 2: Statements made and possible validations made depend on the stage within the DMP lifecycle.

this stage, assessment is necessarily limited to the DMP document itself and can, at best, verify whether such commitments are explicitly stated and whether they align with generic requirements or funder guidelines. No external evidence is available, and compliance is inferred solely from the declared intentions.

In contrast, a final DMP contains statements formulated in the past or present tense, indicating that concrete actions have been implemented. This shift enables additional forms of assessment that go beyond checking the DMP text. For example, claims made in the DMP can be verified against external sources such as repository metadata, allowing assessors to confirm whether a licence has indeed been assigned or whether a declared persistent identifier exists and resolves. In practice, these verification steps often correspond directly to established FAIR assessment procedures. For instance, checking whether a DOI resolves is both a validation of a DMP claim and an assessment of the dataset's level of findability. As a result, the boundary between DMP assessment and FAIR assessment becomes increasingly blurred and, in some cases, effectively overlaps.

We explicitly communicated this distinction and its implications to the workshop participants before they started their collaborative work. In addition, we clarified that basic technical checks—such as verifying whether mandatory fields in the RDA DMP Common Standard (DCS) are present and non-empty—are trivial and well understood from a technical perspective and can be implemented straightforwardly in automated systems. Participants were therefore asked not to focus on such syntactic or structural validations, but instead to concentrate on what they perceive as indicators of a good or poor DMP in substantive terms. This framing served to guide contributions towards metrics that capture the quality, credibility, and evaluative meaning of DMP statements, rather than their mere presence. It also explains why FAIR-related metrics are comparatively underrepresented in the catalogue. We intentionally asked participants to focus on

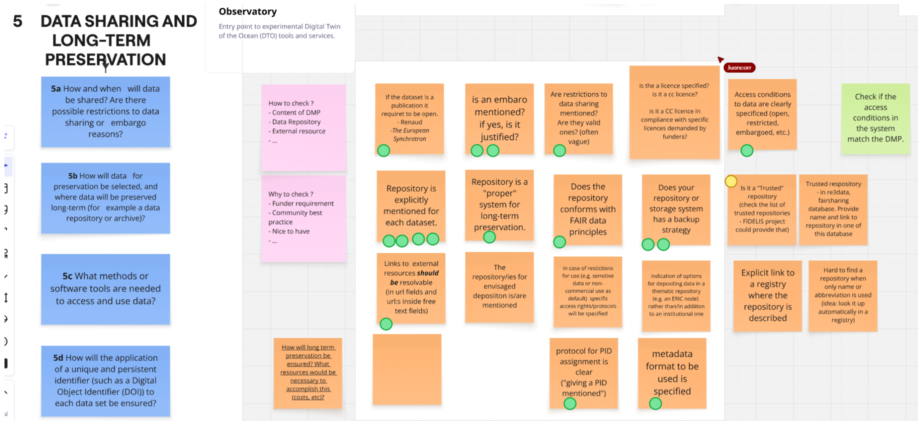

Fig. 3: A subset of sticky notes identified by workshop participants that were later turned into metrics and test.

DMP-specific assessment dimensions, assuming that FAIR metrics can be reused from existing community efforts within OSTrails, where FAIRness is assessed directly for specific digital objects and disciplinary contexts.

3.1 Community Co-design Process

The catalogue was developed through a structured community co-design process centred on a collaborative workshop with participants from the OSTrails consortium and the broader research data management community, including data stewards, tool developers, funder representatives, and domain researchers. The workshop built on a prototype evaluation framework and an initial set of quality dimensions [16] and on quality gaps identified across 24 national and thematic pilots [6], which surfaced recurring shortcomings, such as, inconsistent PID assignment, absent embargo justifications, and unresolved repository references.

Participants used a shared digital whiteboard structured around the six thematic sections of the Science Europe DMP template: data description and collection (Section 1); documentation and data quality (Section 2); storage and backup (Section 3); legal and ethical requirements (Section 4); data sharing and long-term preservation (Section 5); and data management responsibilities and resources (Section 6). For each section, participants identified questions a DMP evaluation system should address, articulated what a check should examine, and assigned a priority score reflecting perceived importance. Figure 3 shows an excerpt of the board for one of the thematic sections.

Checks were defined at two distinct evaluation stages, reflecting the lifecycle-aware view of DMPs [8]. At the *planning stage*, evaluation assesses stated intentions: are required fields present, are declared values plausible and well-formed? At the *execution stage*, evaluation validates intentions against observable evidence: does the declared PID resolve to a live record, does the repository's recorded access condition match the DMP? This distinction directly motivated the internal/external scope classification that structures the catalogue (Section 4.2).

3.2 From Ideation to a Structured Catalogue

The workshop produced a large, heterogeneous corpus of candidate checks in natural language. A consolidation phase applied grouping, deduplication, and progressive formalisation. Overlapping checks were merged into single metric records. Checks too coarse for discrete testing were decomposed into precise sub-checks or deferred. Checks requiring free-text understanding, domain-specific knowledge, or non-public infrastructure were retained as future work candidates.

Four criteria governed inclusion in the current release: (i) *implementability* as a deterministic binary; (ii) *grounding* in an identifiable RDA DCS field [10] or OSTrails Application Profile extension; (iii) *funder and policy relevance* as expressed by workshop priority scores and Horizon Europe pilot findings [6]; and (iv) *non-redundancy*, with domain-specific extensions retained in a separate layer to preserve generality of the core catalogue.

Each retained check was formalised as a metric record following the FTR vocabulary [14]: a precise narrative statement, applicable DCS fields, a quality dimension and scope, and identifiers of implementing tests. Metric records received persistent identifiers under `https://w3id.org/ostrails/dmp/metric/`. The resulting catalogue covers all six thematic DMP sections plus a transversal section on machine-actionability and DCS schema compliance.

4 The DMP Evaluation Metrics Catalogue

4.1 Structure and Quality Dimensions

The catalogue is publicly available as part of OSTrails Commons[9] and currently comprises 94 metrics.

Figure 4 depicts an example of a Metric specification. Each metric record contains: a persistent URI identifier; a human-readable name and narrative description; a quality dimension; the applicable RDA DCS and/or OSTrails Application Profile fields; and the identifiers of tests implementing the metric.

Each of the metrics belongs to one of the five quality dimensions that are based on the work of Arnhold et al. [2]: *completeness* (required fields present and non-empty); *FAIRness* (data and metadata satisfy FAIR criteria); *funder/policy compliance* (requirements of a specific benchmark or funder template are met); *quality* (values are meaningful and well-formed, e.g. a licence field contains a recognised SPDX identifier [5]); and *consistency* (values within a DMP are mutually coherent, e.g. an embargo end date falls after the project start date).

[9] `https://docs.ostrails.eu/en/latest/commons/dmp/dmp-evaluation-metrics.html`

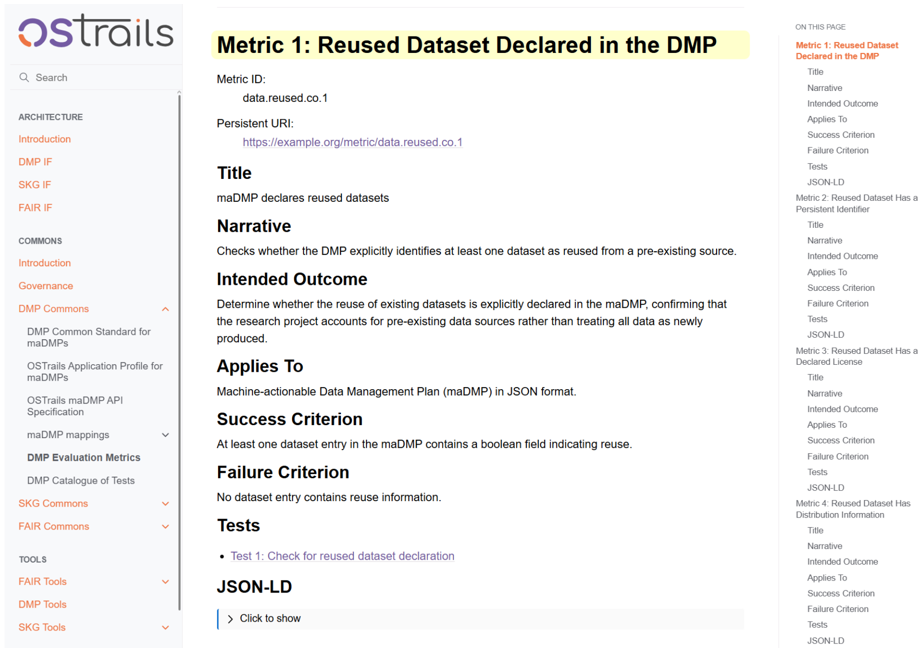

Fig. 4: Example of a Metric specification that is part of the Metrics Catalogue within OSTrails Commons.

4.2 Internal and External Scope

Internal metrics operate exclusively on the maDMP document. They check for the presence, non-emptiness, or formal correctness of declared values: for example, that every dataset entry carries a `data_access` field, or that a `distribution.license.license_ref` is a recognised SPDX identifier [5]. Internal checks require no network access, are fast and deterministic, and correspond to the planning-stage evaluation described above.

External metrics resolve or cross-reference information from outside the DMP and fall into two sub-types. *Feasibility checks* validate that stated information is consistent with observable reality: a declared DOI resolves, a stated repository API returns the referenced dataset. *Compliance checks* verify that values conform to applicable normative standards: the licence declared for a Horizon Europe-funded dataset corresponds to CC BY, or a listed organisation appears in the Research Organisation Registry [17]. These external checks build on the integration model of Papadopoulou et al. [15], which demonstrated that maDMP field values can be systematically cross-referenced against live graph records; the present work operationalises this as discrete, automatable test procedures.

4.3 Metrics, Tests, and FTR Serialisation

In the OSTrails FAIR Reference Model [13], a *metric* is a narrative quality criterion and a *test* is the software artefact that operationalises it, returning `pass`, `fail`, or `indeterminate`. A single metric may be covered by multiple tests: the PID metric, for instance, is implemented separately for DOI, Handle, and ARK schemes. Test result sets are serialised as JSON-LD following the FTR vocabulary [14], built on DCAT [19], DQV [1], and PROV-O [12], enabling evaluation outcomes to be published as linked data and exchanged across platforms without loss of provenance.

4.4 A Worked Example: Dataset PID Metric

We illustrate the metric–test–result chain with the metric *"All datasets described in the DMP carry a persistent identifier."* This metric belongs to the *FAIRness* dimension and has *internal* scope when checking for the presence of a PID field, and *external* scope when verifying that the PID resolves. The applicable field is `dataset[*].dataset_id.identifier` in the RDA DCS schema. Listing 1.1 shows the metric record as JSON-LD.

```
{
  "@context": "https://w3id.org/ftr/",
  "@type":    "ftr:Metric",
  "@id":      "https://w3id.org/ostrails/dmp/metric/DMP-M-001",
  "name":     "Datasets have persistent identifiers",
  "description": "All datasets in the DMP must carry a recognised
                   PID in the dataset_id field.",
  "qualityDimension": "FAIRness",
  "scope":    "internal",
  "appliesTo": ["dcs:dataset.dataset_id.identifier",
                "dcs:dataset.dataset_id.type"],
  "implementedBy":
    ["https://w3id.org/ostrails/dmp/test/DMP-T-001"]
}
```

Listing 1.1: JSON-LD serialisation of the Dataset PID metric record (DMP-M-001).

Test `DMP-T-001` checks that each `dataset` block carries a non-empty `dataset_id.identifier` whose `type` is a recognised PID scheme (`doi`, `handle`, `ark`, `orcid`). Listing 1.2 shows the resulting FTR record, which includes the outcome, a plain-language explanation, and the full provenance chain (subject DMP, test, metric, generating agent, timestamp).

```
{
  "@context": "https://w3id.org/ftr/",
  "@type":    "ftr:TestResult",
```

```
  "subject":  "https://example.org/dmps/dmp-42",
  "test":   "https://w3id.org/ostrails/dmp/test/DMP-T-001",
  "metric": "https://w3id.org/ostrails/dmp/metric/DMP-M-001",
  "outcome":  "pass",
  "message":  "All 2 datasets carry a recognised PID (doi).",
  "date":     "2025-11-03T10:14:22Z",
  "generatedBy": {
    "@type": "prov:SoftwareAgent",
    "@id":   "https://github.com/ostrails/dmp-eval-service",
    "label": "OSTrails DMP Evaluation Service v1.0"
  }
}
```

Listing 1.2: JSON-LD FTR result for test DMP-T-001 on a passing maDMP.

The `outcome` field is directly consumable by a DMP platform to surface a pass/fail indicator to the researcher. When multiple tests run, result sets are aggregated into a benchmark score as described in Section 5.

5 Automated Evaluation across Community Benchmarks

5.1 Implementation

The catalogue is implemented in the OSTrails DMP Evaluation Service, a RESTful API publicly available at `https://ostrails-dmp-evaluation.arisnet.ac.at`. The service accepts a maDMP JSON document and a benchmark identifier at `POST /assess` and returns an FTR-compliant evaluation report. Test logic is encapsulated in a plugin mechanism: each plugin exports functions implementing one or more catalogue tests, covering completeness, FAIRness, funder compliance, and external validation (PID resolution; RE3data and ROR registry queries). New plugins can be registered without modifying the core engine. Two DMP authoring tools—Data Stewardship Wizard (DSW) and Argos—consume the service API to surface real-time evaluation feedback within their authoring interfaces, following the same loose-coupling pattern: serialise the current DMP as maDMP JSON, submit to `POST /assess` with a benchmark identifier, display the returned report.

5.2 Evaluation of Sample DMPs

To demonstrate the end-to-end pipeline, we evaluate two synthetic maDMPs against a benchmark derived from the FWF Criteria for Open Research Data. This benchmark is based on the FWF Open Access requirements for research data [3]. These requirements are broadly representative of those adopted by most

```
{ "dataset": [{
  "title": "Climate Observation Dataset",
  "distribution": [{
    "license": [{ "license_ref":
      "https://example.com/custom-license"
    }],
    "host": { "url":
      "https://internal-repo.uni.org" }
  }]
}]}
```

Listing 1.3: DMP-A: no PID, proprietary licence, unregistered repository.

```
{ "dataset": [{
  "dataset_id": {
    "identifier":
      "https://doi.org/10.5281/zenodo.1234567",
    "type": "doi" },
  "distribution": [{
    "license": [{ "license_ref":
      "https://creativecommons.org/licenses/by/4.0/"
    }],
    "host": { "url": "https://zenodo.org" }
  }]
}]}
```

Listing 1.4: DMP-B: CC BY, Zenodo, DOI not yet deposited.

funders that are partners in Science Europe. Moreover, interviews with funder-appointed reviewers indicate that these criteria reflect the typically minimal checks applied to DMPs during funder assessment.

DMP-A describes a climate observation dataset that violates all three FWF requirements: it lacks a persistent identifier, applies a proprietary licence, and designates a repository that is not listed in re3data.

DMP-B describes a biodiversity survey dataset with a declared DOI, a CC BY 4.0 licence, and Zenodo as the target repository. The licence and repository are compliant with FWF requirements; however, the DOI does not yet resolve (HTTP 404), which is a common situation when a DMP is authored prior to data deposition.

Both plans conform to the RDA DMP Common Standards (DCS) schema.

5.3 Results and Discussion

Table 1 summarises test outcomes for both DMPs against the three FWF metrics. Figures 5 and 6 show the corresponding service web interface.

Table 1: FWF benchmark results for two synthetic maDMPs (**F** = FAIL, **P** = PASS).

Metric	DMP-A	DMP-B
Dataset PID Resolvability	**F** — missing or invalid identifier	**F** — DOI resolves to HTTP 404
Open Licence Compliance	**F** — proprietary licence not in OKF list	**P** — CC BY 4.0 recognised
RE3data Repository Registration	**F** — not found in RE3data	**P** — Zenodo (`r3d100010468`)

The results illustrate three properties of the catalogue in practice. First, external checks are *dynamically validated*: the licence check queries the Open Knowledge Foundation API and the repository check queries RE3data at runtime, so results always reflect the current state of external registries. Second, failure explanations are *actionable at field level*: each result identifies the specific DCS field responsible, enabling targeted revision. Third, DMP-B's partial compliance profile demonstrates that the catalogue correctly *disaggregates results by metric*, giving researchers a precise view of which requirements they satisfy and which remain open—a more actionable signal than a single aggregate score.

The evaluation also exposes limitations of the current catalogue. The `indeterminate` outcome—returned when an external resource is temporarily unreachable or a field is present but unverifiable—occurs more frequently than anticipated in practice, underscoring the need for graceful degradation strategies in subsequent releases. Criteria that fall outside the catalogue's current scope, such as the appropriateness of stated embargo periods or the methodological coherence of data collection descriptions, cannot yet be checked automatically; these require either structured schema extensions or NLP-assisted evaluation components.

6 Conclusion and Future Work

This paper has presented a community-driven approach to formalising and applying evaluation criteria for machine-actionable data management plans (maDMPs), addressing a long-standing gap between funder requirements, reviewer practice, and scalable automated assessment.

The first contribution is the **community co-design methodology**: a structured process involving practitioners, tool developers, and funders that elicited, prioritised, and refined evaluation criteria grounded in real-world DMP workflows and policy requirements. By explicitly accounting for the DMP lifecycle and distinguishing between planning-stage intentions and execution-stage evidence, the methodology clarifies what can be assessed at different stages and why DMP assessment partially overlaps with FAIR assessment in later phases.

The second contribution is the **DMP Evaluation Metrics Catalogue**. The catalogue formalises community-elicited criteria as linked-data metric and test records using the FAIR Testing Results (FTR) vocabulary, anchors each

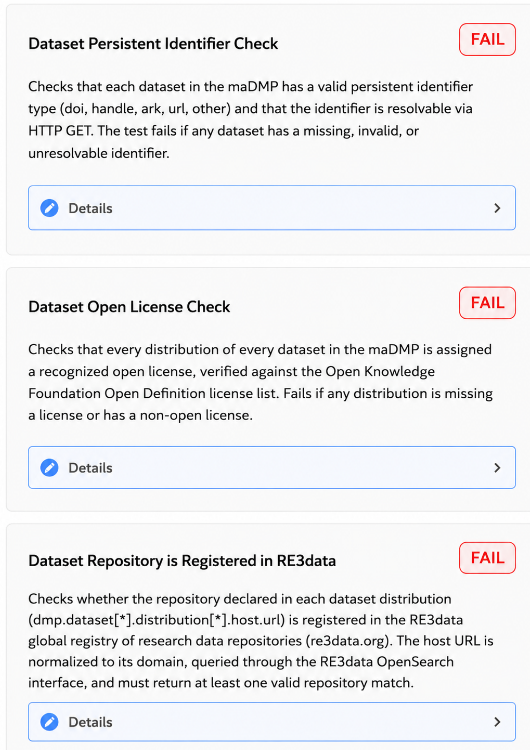

Fig. 5: DMP-A evaluation: all three tests fail with field-level explanations.

metric to specific RDA DMP Common Standard fields, and organises metrics along lifecycle-aware quality dimensions and internal/external scope. Published openly as part of OSTrails Commons, the catalogue makes evaluation criteria transparent, inspectable, and reusable across tools and contexts.

The third contribution is a **demonstration of automated evaluation** against a community-specific benchmark, showing that a substantial subset of the catalogue can be operationalised as deterministic, externally validated tests. The resulting FTR-compliant reports provide machine-readable, field-level feedback that is actionable for researchers, data stewards, and funders, and can be integrated directly into DMP authoring environments.

Building on this foundation, several directions for future work are already underway. We are currently defining *community- and policy-specific benchmarks* that represent concrete evaluation profiles for particular funders, institutions, and research communities. Each benchmark selects and combines a subset of metrics from the catalogue to reflect explicit community needs, making assessment criteria both transparent and comparable across contexts.

In parallel, we are extending the implementation coverage of the catalogue by developing additional automated tests for currently unimplemented metrics.

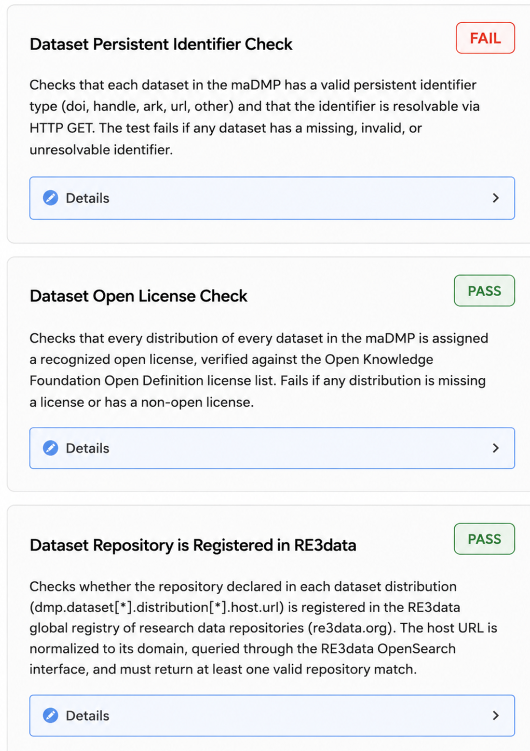

Fig. 6: DMP-B evaluation: PID resolvability fails; licence and repository tests pass.

While many metrics can already be evaluated deterministically, others require further infrastructure integrations, evolving registry interfaces, or more advanced validation logic. Closing these gaps will increase both the coverage and robustness of automated DMP assessment.

Support for benchmark-based evaluation is also being integrated into major DMP authoring tools. Data Stewardship Wizard and Argos already consume the OSTrails DMP Evaluation Service to provide real-time feedback during authoring, and DAMAP is actively adding support for benchmark-driven assessment workflows. These integrations enable consistent evaluation across tools and submission portals while reducing the burden on reviewers.

Finally, the DMP Evaluation Metrics Catalogue itself will continue to evolve. Future extensions will incorporate additional community contributions, domain-specific layers, and richer consistency and lifecycle-transition metrics. Where appropriate, these extensions will be aligned with ongoing OSTrails efforts to define FAIR assessment metrics for specific digital objects, ensuring complementarity rather than duplication between DMP evaluation and FAIRness assessment.

Taken together, these developments position the catalogue not as a static checklist, but as a living, community-maintained evaluation infrastructure supporting scalable, transparent, and context-aware assessment of data management plans.

Acknowledgments

This work has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101130187 (OSTrails – Open Science Plan-Track-Assess Pathways).

SBA Research is a COMET Centre within the framework of COMET — Competence Centers for Excellent Technologies.

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Albertoni, R., Isaac, A.: Data Quality Vocabulary (DQV) (2016), https://www.w3.org/TR/vocab-dqv/
2. Arnhold, L., Miksa, T., Staudinger, M.: Quality dimensions and evaluation framework for machine-actionable data management plans. ACM Journal of Data and Information Quality **17**(4), Article 29 (2025). https://doi.org/10.1145/3776555
3. Austrian Science Fund (FWF): Open-access policy for research data (2019), https://www.fwf.ac.at/en/about-us/what-we-do/open-science/open-access-policy/open-access-policy-for-research-data
4. Cardoso, J., Castro, L.J., Ekaputra, F.J., Jacquemot, M.C., Suchanek, M., Miksa, T., Borbinha, J.: DCSO: towards an ontology for machine-actionable data management plans. Journal of Biomedical Semantics **13**(21) (2022). https://doi.org/https://doi.org/10.1186/s13326-022-00274-4
5. Linux Foundation: Software Package Data Exchange (SPDX) license list (2010), https://spdx.org/licenses/
6. Manola, N., Papadopoulou, E., Kontopidi, M., Grypari, I., Spichtinger, D., Kakaletris, G., Tziotzios, D.: D4.2: Horizon europe report. Tech. rep., OSTrails Project (2025). https://doi.org/10.5281/zenodo.16753141
7. Miksa, T., Neish, P., Walk, P., Rauber, A.: Defining Requirements for Machine-actionable Data Management Plans. In: McGovern, N., Whiteside, A. (eds.) Proceedings of the 15th International Conference on Digital Preservation, iPRES 2018, Boston, MA, USA, September 24-28, 2018 (2018), https://hdl.handle.net/11353/10.923628
8. Miksa, T., Simms, S., Mietchen, D., Jones, S.: Ten principles for machine-actionable data management plans. PLOS Computational Biology **15**(3), e1006750 (2019). https://doi.org/10.1371/journal.pcbi.1006750

9. Miksa, T., Suchanek, M., Slifka, J., Knaisl, V., Ekaputra, F.J., Kovacevic, F., Ningtyas, A.M., El-Ebshihy, A., Pergl, R.: Towards a toolbox for automated assessment of machine-actionable data management plans. Data Science Journal **22**, 28 (2023). https://doi.org/10.5334/dsj-2023-028
10. Miksa, T., Walk, P., Neish, P.: RDA DMP common standard for machine-actionable data management plans (2020). https://doi.org/10.15497/rda00039
11. Miksa, T., Walk, P., Neish, P., Oblasser, S., Murray, H., Renner, T., Jacquemot-Perbal, M.C., Cardoso, J., Kvamme, T., Praetzellis, M., Suchánek, M., Hooft, R., Faure, B., Moa, H., Hasan, A., Jones, S.: Application profile for machine-actionable data management plans. vol. 20, p. 32. Ubiquity Press, Ltd. (Oct 2021). https://doi.org/10.5334/dsj-2021-032
12. Moreau, L., Missier, P.: PROV-O: The PROV ontology (2013), https://www.w3.org/TR/prov-o/
13. OSTrails Consortium: OSTrails FAIR reference model and infrastructure (2024), https://ostrails.eu
14. OSTrails Consortium: FAIR testing results (FTR) vocabulary (2025), https://w3id.org/ftr/
15. Papadopoulou, E., Bardi, A., Kakaletris, G., Tziotzios, D., Manghi, P., Manola, N.: Data management plans as linked open data: Exploiting ARGOS FAIR and machine-actionable outputs in the OpenAIRE research graph. Journal of Biomedical Semantics **14**, 17 (2023). https://doi.org/10.1186/s13326-023-00297-5
16. Papadopoulou, E., Kontopidi, M., Miksa, T., Romero, A.: D3.1: DMP evaluation rubric and service specifications. Tech. rep., OSTrails Project (2025). https://doi.org/10.5281/zenodo.14795222
17. Research Organization Registry Community: Research organization registry (ROR) (2019), https://ror.org
18. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., et al.: The FAIR guiding principles for scientific data management and stewardship. Scientific Data **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
19. World Wide Web Consortium: Data Catalog Vocabulary (DCAT) — version 3 (2024), https://www.w3.org/TR/vocab-dcat-3/