Automating Evaluation of Machine-Actionable Data Management Plans with Semantic Web Technologies

Raffael Foidl[1] [0000−0002−1748−1571], Lea Salome Brugger[1] [0000−0001−8770−4112], and Tomasz Miksa[1,2] [0000−0002−4929−7875]

[1] TU Wien, Vienna, Austria
[2] SBA Research, Vienna, Austria

Abstract. *Machine-actionable data management plans (maDMPs)* have, by their very nature, potential to bring advantages over data management plans that are written in text form. By employing maDMPs, not only researchers should be able to benefit from their merits, but also research funders receiving and assessing the DMPs. Science Europe, which is an association of major European research funders, have published an evaluation rubric that provides a common basis to support the evaluation of DMPs. By stating a set of criteria, it helps to ensure submitted DMPs cover required aspects and support FAIR data management.
In this paper, we present a semi-automatic approach to leverage the benefits of maDMPs by providing SPARQL queries that represent requirements of Science Europe. The goal is to support reviewers in the assessment of DMPs expressed as maDMPs. The results show that semantic web technologies can help in providing customised views to reviewers, but human inspection and interpretation is still needed.

Keywords: Machine-Actionable Data Management Plan · Reviewing · Funder Support · RDF · JSON-LD · SPARQL

1 Introduction

Data Management Plans (DMPs) are formal documents that describe what data are used or produced and how the data are managed with respect to FAIR principles [16] throughout a project and beyond. They are usually text documents that follow specific funder guidelines and templates and are often required by funders when researchers apply for research grants.

The benefits of DMPs are manifold. Most notably, documenting the data management in a DMP allows to preserve continuity throughout a project. Furthermore, formulating a DMP helps project members consider important, yet often neglected issues such as ethics, security and copyright right from the start.

Despite their advantages, the process of creating a DMP can be a time-consuming exercise for researchers. Reviewing a DMP can also be a cumbersome task for reviewers who are appointed by the funders to assess the quality of submitted DMPs, because despite the standardisation of templates, the information can be provided in different sections of the text document and on different level of details. Thus, making it hard for the reviewers to find the relevant information needed for the assessment of the quality and completeness of DMPs.

By capturing key information about datasets used in a project which are essential for the stakeholders involved and being machine-readable, maDMPs aim to reduce the workload associated with DMPs while still providing sufficient narrative on details on data management. A concrete example would be the facilitation of reviewing processes; maDMPs can provide views to the reviewers on the relevant information to assess specific criteria. Despite the presence of the *RDA DMP Common Standard for maDMPs* [10], there is still no tool or standard procedure that allows for (partial) automation of this step based on a given maDMP. Hence, manual assessment of maDMPs is still necessary.

In an attempt to mitigate this problem, we took a semi-automated approach to assessing maDMPs and formed the following hypothesis: By constructing SPARQL queries based on the fields defined in the RDA DMP Common Standard and applying them to actual maDMPs, one can extract the desired information without the need to manually inspect them. Depending on the query, it should also be possible to check whether information is missing or whether specific requirements are fulfilled. Thus, we should be able to facilitate the work of reviewers. As a side effect, the queries can also be used by authors to check whether their maDMP satisfies the rubric.

The *Practical Guide to the International Alignment of Research Data Management* contains an evaluation rubric [12] that serves as a guidance for reviewers of DMPs. In our work, we want to find out to what extent one can filter the information from maDMPs and map them to existing requirements imposed by the evaluation rubric. For our use case, we first declared a set of SPARQL queries based on the guidelines provided in that template. Then, we applied them to 12 different maDMPs and evaluated our mapping of criteria to queries as well as the maDMPs themselves. While the mapping proved to be useful for querying general information about the project, data management responsibilities and documentation, it lacks efficacy in the area of collection and re-use of data because there is no clear correspondence in the maDMP schema. The proposed collection of SPARQL queries aids in quickly detecting the absence of information essential for a DMP – the final judgement, however, still has to be made by the reviewer.

The remainder of this paper is structured as follows: In Section 2, we briefly present related work on the topic. Section 3 contains an explanation of the methodology we pursued. Section 4 provides an overview of our mapping from the criteria defined in the Science Europe evaluation rubric to the SPARQL queries we used and in Section 5, we discuss our use case and the results of the evaluation of our approach. Lastly, in the conclusion we summarize our findings and contributions. For a more detailed discussion and documentation of the project and our findings, the reader is directed to our GitHub repository [5] containing the input files, our SPARQL queries and the complete evaluation.

2 Related Work

"Data management plans are required by funding bodies and institutions all over the world, e.g. the National Science Foundation (NSF) in the USA, the European Commission in Europe, or the National Research Foundation (NRF) in South Africa" [9]. There are many tools for researchers that help them create a DMP, e.g. DMPonline[3], DMPTool[4], Data Stewardship Wizard[5] or RDM Organizer [4]. Typically, researchers fill out questionnaires to create a DMP that is compliant with the given funder template. An overview of tools and requirements can be found in [7].

Science Europe, which brings together European research funders, issued guidelines for the development of DMP templates [12] which contain an *evaluation rubric* that provides common guidelines for reviewers of DMPs.

The *Research Data Alliance (RDA)* identified the limitations of current DMPs and took steps towards "active" DMPs [15]. A standardized machine-actionable (meta-)data model [6,11] for DMPs is a prerequisite for DMPs to become "active" and to enable the exchange of DMP information between research systems and the automation of RDM tasks throughout the research data lifecycle. The *RDA DMP Common Standards* working group produced an official recommendation [10] that describes an application profile for machine-actionable DMPs (maDMPs). Thus, the RDA established a common way to model information that is typically described in DMPs. Information contained in this model can be re-purposed, e.g. to be used by RO-Crates [8].

Machine-actionability can be defined as *"information that is structured in a consistent way so that machines, or computers, can be programmed against the structure"* [3]. The term machine-actionability is associated with the FAIR principles to express that machines should be able to autonomously take action on digital objects [16]. Machine-actionability for DMPs shall be achieved by modelling the semantic information with the use of controlled vocabularies and standards [9], as well as, by using persistent identifiers (PIDs) to reference specific entities such as people, institutions, funders, grants, datasets, or repositories [13,14]. In [1], the authors present an ontological representation of maDMPs that is used in this paper.

3 Methodology

This section presents an overview of the methodology we followed. Details on each of the steps are provided in consecutive sections.

[3] https://dmponline.dcc.ac.uk
[4] https://dmptool.org
[5] https://ds-wizard.org

3.1 Mapping DMP Requirements to SPARQL Queries

The *Practical Guide to the International Alignment of Research Data Management*[6] contains an evaluation rubric that serves as a guidance for reviewers of DMPs. Based on the requirements stated in this document, we formulated SPARQL queries that try to capture the criteria stated in natural language in the form of triple patterns. To that end, we have mapped the requirements from the rubric to patterns requiring the presence of values for respectively suitable fields from the *RDA DMP Common Standard for machine-actionable Data Management Plans*[7] JSON schema. We developed queries that project certain subsets of the data into a customized view (`SELECT` queries) as well as ones that simply indicate whether some criteria are satisfied (`ASK` queries).

After having expressed the guidelines using SPARQL as thoroughly as possible, we also conducted a coverage evaluation. We quantified the proportion of aspects mentioned in the evaluation rubric that we were able to cover using queries. More information on the mapping itself and the coverage can be found in Section 4.

3.2 Preparing the maDMPs for Automatic Assessment

We took the maDMPs as input data for our project from the *Zenodo Community Data Stewardship 2021 – DMPs*[8]. These maDMPs were created by students in the context of the exercise part of a university lecture. They were created based on previously created "regular" DMPs which can also be found in the linked Zenodo community. We had to take some measures in order to transform the raw input maDMPs from the Zenodo community such that later on, they could be utilized to perform the (semi-)automatic assessment of the queries mentioned in Section 3.1 (see also Section 3.3). At first, we had to make slight adjustments to some maDMPs to achieve conformity with the JSON schema. This includes wrapping objects into arrays with one element when required by the schema or rectifying date(-time) formats. Along with this step, we performed normalization procedures to attain consistent formatting, indenting and JSON property sort order.

Secondly, as our ultimate goal was to assess the completeness of the maDMPs by executing SPARQL queries, we needed to transform the input maDMPs into corresponding RDF representations. In an effort to do so, we created instances of the *DMP Common Standard Ontology (DCSO)*[9] from them by employing the `dcsojson`[10] tool. It (bidirectionally) converts between schema-conform JSON maDMPs and different DCSO serializations. More specifically, we transformed the ontology instances corresponding to the normalized input maDMPs into

[6] https://doi.org/10.5281/zenodo.4915861
[7] https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard
[8] https://zenodo.org/communities/dast-2021/
[9] https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/tree/master/ontologies
[10] https://raffaelfoidl.github.io/maDMP-evaluation/0007.html

equivalent JSON-LD serializations. This was achieved with a script that runs `dcsojson` on the files residing within a given directory.

Finally, as a post-processing step, we again established an alphabetical sorting of the JSON-LD properties. The procedure described here was necessary in order to put the mapping to the test on a concrete use case (see Section 5).

3.3 Use Case Application

The final step was to evaluate a set of maDMPs using the queries we developed. The goal was to determine the extent to which the produced queries facilitate the work of reviewers, discover strengths and shortcomings.

During our experiment, we used a local GraphDB[11] instance as triple store and SPARQL endpoint. We inserted the JSON-LD data yielded by the step described above into the triple store and executed the queries on those maDMPs.

This assessment is meant to determine the limit of what the proposed mapping is able to contribute in evaluating the completeness of given maDMPs. It is not specifically meant to gauge the usefulness of the input maDMPs from the Zenodo community, even though it may allow to draw some conclusions in that regard. The results of this evaluation are presented in Section 5.

4 Mapping

In this section, we present our mapping from the guidelines in the Science Europe evaluation rubric to SPARQL queries. First, we explain the nature of the SPARQL queries based on two examples. Next, we discuss which parts of the template we were able to cover. Finally, we describe the limitations we faced when constructing the queries and how our queries could be improved.

In general, there are two types of queries: SELECT and ASK. SELECT queries collect and display the queried information while ASK queries verify whether the maDMP fulfills a certain criterion, i.e. they return `YES` or `NO`. For each (sub)category of the Science Europe evaluation rubric, we first identified the corresponding elements in the maDMP schema and then constructed SPARQL queries based on those elements.

For example, consider the first subcategory in the Evaluation Rubric (Administrative information): *"Provide information such as name of applicant, project number, funding programme, version of DMP."* The criterion is partly verified by query `0-1` in the GitHub repository [5] which is displayed in Listing 1.1. It returns the basic information, i.e. the author, title, date of creation and language of the maDMP as well as the ID of the corresponding DMP, by querying the respective elements in the maDMP schema.

[11] https://www.ontotext.com/products/graphdb/

```
SELECT ?title ?author ?email ?created ?language ?dmpId ?
    dmpIdType WHERE {
    ?maDMP dcso:hasContact ?contact ;
           dcso:hasDMPId ?dmp ;
           dct:created ?created ;
           dcso:language ?language ;
           dct:title ?title .
    OPTIONAL { ?maDMP dcso:hasProject ?project . }

    ?dmp dct:identifier ?dmpId ;
         dcso:identifierType ?dmpIdType .

    ?contact foaf:name ?author ;
             foaf:mbox ?email .
}
```

Listing 1.1: Example SELECT Query

An example for an ASK query is provided in Listing 1.2. This query covers the third subitem of subcategory 2a in the Evaluation Rubric: *"Use community metadata standards where these are in place."* The element in the maDMP schema which describes the used metadata standard is the `metadata_standard_id` field that consists of an `identifier` and a `type` element (which can take the value `url` or `other`). Our query verifies whether the maDMP contains information about the metadata for all datasets and the metadata standard specified for each dataset is a community standard; in our case, we restrict the query to the standards DDI[12], TEI[13], EML[14], MARC[15] and DCMI[16], most of which are listed as examples in the Science Europe template. Note that this query serves as an example of how the metadata standards in place could be queried and that it would likely need to be adjusted to fit a funder's specific domain and requirements. Further, since our paper serves as a proof of concept rather than a complete solution, we followed a quite naive approach by using hard-coded strings in the query, which can be problematic. This should be kept in mind when implementing a similar query based on our example.

```
ASK WHERE {
    ?dataset dcso:hasMetadata ?metadata .

    ?metadata dcso:hasMetadataStandardId ?metadataStandard ;
              dct:description ?description ;
              dcso:language ?language .

    ?metadataStandard dct:identifier ?standardId ;
```

[12] https://ddialliance.org/
[13] https://tei-c.org/
[14] https://eml.ecoinformatics.org/
[15] https://www.loc.gov/marc/
[16] https://dublincore.org/

```
                              dcso:identifierType "url" .

    FILTER (CONTAINS(?standardId, "dublin") || CONTAINS(?
        standardId, "eml") || CONTAINS(?standardId, "marc")
        || CONTAINS(?standardId, "tei") || CONTAINS(?
        standardId, "ddi"))
}
GROUP BY ?description ?language ?standardId ?standardType
```

Listing 1.2: Example ASK Query

The complete set of queries we constructed can be found in the corresponding directory in the GitHub repository [5].

4.1 Coverage

Table 1 gives an overview of the spectrum of the Science Europe evaluation rubric that we were able to cover with SPARQL queries. We broke each category down into subitems; the table shows how many of those subitems are "largely covered" with one or multiple queries. In this case, "largely covered" means that there are one or multiple elements in the maDMP schema that can be used to describe most of the content specified in the respective subitem and there is at least one query that queries those elements.

Table 1: Ratio of Aspects From Evaluation Rubric That Are Covered by Queries

Category	Subitems	Largely Covered	Percentage
0 General Information	1	1	100 %
1 Data Description and Collection or Re-Use of Existing Data	9	3	33 %
2 Documentation and Data Quality	7	5	71 %
3 Storage and Backup During the Research Process	6	3	50 %
4 Legal and Ethical Requirements, Code of Conduct	6	3	50 %
5 Data Sharing and Long-Term Preservation	12	8	67 %
6 Data Management Responsibilities and Resources	7	5	71 %
Sum	48	28	58 %

As one can see in the table, our mapping from the criteria defined in the Science Europe evaluation rubric to SPARQL queries covers roughly half of the bullet points given in the rubric. This is mainly due to the fact that the rubric is part of a guideline for the creation of DMPs (rather than maDMPs) and for some aspects, there is no suitable field in the maDMP schema that could be queried. Furthermore, maDMPs provide a minimum set of common fields that are independent of a specific funder template. It is possible to extend maDMPs with funder specific extensions and the RDA working group is discussing this already [2].

Constructing the SPARQL queries worked especially well for the very first category in the rubric which describes general information such as the applicant and project number. Further, they proved to be quite helpful in querying information about data management responsibilities and resources (category 6 in the guidelines) and documentation and data quality (category 2). This is mainly because there is a feasible mapping in the maDMP schema which describes the majority of the contents of the respective categories (see the examples above).

On the other hand, it was rather difficult to express the criteria concerning the collection or re-use of data (category 1), mostly because they are not really covered by the maDMP schema, as described above. For instance, consider the subitem *"State any constraints on re-use of existing data if there are any."* in subcategory 1a in the evaluation rubric: While this type of information can be easily provided in textual form in a DMP, there is no corresponding element in the maDMP schema that could be used in order to fulfill the criterion. Note that this applies to quite a few subitems in the evaluation rubric because those subitems concern information that will never be machine-actionable. Please note that the maDMPs specification has a way to incorporate such non-machine-actionable fields, e.g. using the `Dataset/description` or `Distribution/description` fields, but those fields can be used to collect *any* non-machine-actionable information and thus it is hard to attribute their contents to specific questions – this can only be judged by a human at this stage.

Table 2 gives an example of how we documented the creation of our queries and the correspondence to criteria from the evaluation rubric. The whole overview is available in the GitHub repository [5].

Table 2: Excerpt From Juxtaposition of Rubric Criteria and Queries

Requirement	Covered In	Remarks
6a Who (for example, role, position, and institution) will be responsible for data management (i.e. the data steward)?		
Outline the roles and responsibilities (...). Name responsible individual(s) where possible.	`6-a-1`, `6-a-2`	The queries show all available information about the contact person and contributions.
For collaborative projects, explain the co-ordination of data management responsibilities across partners.	`6-a-2`	Depicts information about contributors defined by the maDMP.
Indicate who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised.	/	Not explicitly covered by maDMP, but 6-a-1 and 6-a-2 are good indicators of who might be responsible.

4.2 Limitations

As explained in the previous subsection, not all criteria given in the rubric can be easily queried with a SPARQL query. This issue could be (partially) solved by extending the maDMP schema such that it includes corresponding mappings for the elements in the rubric that are currently not reflected. However, this is outside of the scope of this paper.

For some queries, it was necessary to make certain assumptions. For example, a lot of queries query, among other things, the `host` element of the maDMP schema which is optional, i.e. the maDMP conforms to the schema even if this field is not defined. In those cases, we assumed that the `host` element is provided. However, this lead to some queries failing on a few of our input maDMPs due to a missing definition of the `host` element although other information that the query also aimed to collect was present (see Section 5). Similar to the issue described above, revising the maDMP schema could also potentially fix this problem: If the schema is adjusted or extended such that it enforces the declaration of certain fields that are currently optional, the respective queries would not have to rely on the assumption that the queried elements are present in the input maDMP.

In general, we want to stress that some queries could be formulated less strict, i.e. `OPTIONAL` blocks could be inserted for triple patterns that join elements from the maDMP schema that are, by the schema definition, optional. Nonetheless, as this paper is more of a proof-of-concept-kind, this could easily be done when extending or building upon the work at hand – in order not to "lose" any results/information about the corresponding maDMP.

Lastly, using SPARQL queries only facilitates information collection and filtering; for example, automatic assessment is not possible with SPARQL queries alone, i.e. they cannot replace a reviewer.

5 Use Case

In this section, we present an aggregated assessment of the 12 maDMPs submitted to the aforementioned Zenodo community.

We used a satisfaction value (SV), a numeric value on a scale from zero to five, to assess the maDMPs. A value of five is equivalent to a holistic fulfilment of the respective criterion, a value of zero denotes that the criterion is "not satisfied". In this case, "not satisfied" means that the maDMP provides insufficient information with regard to the respective criterion.

We evaluated the maDMPs as follows: For each category, we first applied the SPARQL queries to the input file. The completeness of the maDMPs was determined based on whether the result set of the queries was empty or not. If it was empty, we inspected the respective file manually. In the case that we found answers to the respective criteria in the correct maDMP fields, we used that information to iteratively improve our queries, as this indicated that we had made incorrect assumptions when designing them. After having collected all information provided in the input maDMPs for the respective category, we determined the satisfaction value depending on the completeness of the results. Criteria from the evaluation rubric that do not have a corresponding mapping in the maDMP schema – e.g. information that will never be machine-actionable, such as general considerations which only fit in the free text `description` fields of the schema – were ignored in the evaluation since they cannot be fulfilled.

Note that an maDMP is not necessarily of high quality if the SPARQL queries produce results. This is the intended approach to reviewing an maDMP with the help of SPARQL queries: The queries are used to gather information without having to manually search for them. However, the reviewer is still responsible for interpreting the results and assessing their informational value.

Table 3 shows an aggregated overview of the evaluation results by depicting the average satisfaction value for each category defined in the rubric as well as the sum of averages. For the individual evaluation results, the reader is directed to the GitHub repository [5].

Table 3: Aggregated Overview of the Outcome of the maDMP Evaluation

Category	**Average SV**
0 General Information	3.9
1 Data Description and Collection or Re-Use of Existing Data	4.0
2 Documentation and Data Quality	1.6
3 Storage and Backup During the Research Process	2.3
4 Legal and Ethical Requirements, Code of Conduct	3.2
5 Data Sharing and Long-Term Preservation	3.6
6 Data Management Responsibilities and Resources	3.5
Sum	22/35

As one can see in Table 3 and the individual evaluations in the GitHub repository [5], the main issues in the input maDMPs are insufficient documentation of the metadata accompanying the used and produced data as well as lacking information about the storage and backup of data (categories 2 and 3). For instance, frequently missing information was detected concerning the description of security measures and how sensitive data was stored. Another example of missing information we were able to find with our queries concern data collection and re-use (category 3). Missing definitions of file formats, sizes and general data descriptions were discovered. Regarding the other categories, most of the maDMPs provided a decent amount of information, with occasional shortcomings. One aspect worth mentioning here is the missing definition of the `host` element which is an issue that appeared in quite a few maDMPs and led to incomplete information in multiple categories (3, 4, and 5). Further, the costs were neglected in all files and the required resources were only specified in one maDMP.

Table 4 gives an example of how we documented the application of queries to our input maDMPs for a specific instance. The complete tables for all 12 maDMPs are available in the GitHub repository [5].

Table 4: Excerpt From Evaluation of Example maDMPs

Category	SV	Justification
0 General Information	2	Sufficient information about DMP. Information about project not included.
1 Data Descirption and Collection or Re-Use of Data	4	There is a clear description for each distribution. The file formats and data size are specified (except for the code).
2 Documentation and Data Quality	0	No keywords specified, information about metadata and versioning is missing. Minimal information regarding data quality assurance provided.

6 Conclusion

In this paper, we proposed a semi-automated approach to the assessment of maDMPs by utilizing semantic web technologies. Based on the evaluation rubric for DMPs provided by Science Europe, we defined SPARQL queries and applied them to 12 maDMPs obtained from a university course. We then evaluated the maDMPs with respect to completeness based on the results of the queries.

Although it is clear that manual evaluation by reviewers will never be fully replaceable, SPARQL queries can certainly facilitate their work. They enable filtering of relevant information in order to create custom views to answer specific questions. Further, they help validating the fulfillment of certain requirements. Nonetheless, not everything can be easily covered by SPARQL queries.

Our queries proved especially useful in retrieving general information (such as project number and applicant), information regarding data management responsibilities and resources as well as documentation and data quality. In contrast, it was, for instance, difficult to gather information concerning the collection or re-use of data with a query. The main reason why some (sub)categories cannot be expressed as SPARQL queries is that they concern information that is not machine-actionable (and will never be), such as general considerations that only fit in the free text `description` fields of the schema.

Future work will focus on further evaluation of the degree to which we can automate and facilitate the assessment of maDMPs using semantic web technologies. We plan to include a larger sample of maDMPs that also cover funder-specific extensions, so that more funder-specific requirements are reflected. Apart from that, we will seek collaboration with reviewers of DMPs and funders to identify specific use cases in which the SPARQL queries can support them most. Furthermore, as validation languages such as SHACL and ShEx are becoming more mature, we intend to explore how they can aid in the process of evaluating maDMPs.

References

1. Cardoso, J., Garcia Castro, L.J., Ekaputra, F., Jacquemot-Perbal, M.C., Miksa, T., Borbinha, J.: Towards semantic representation of machine-actionable data management plans. In: DaMaLOS – First Workshop on Data and Research Objects

Management for Linked Open Science: Co-located at the International Semantic Web Conference ISWC 2020. PUBLISSO (2020), https://doi.org/10.4126/FRL01-006423289

2. Cardoso, J., Jones, S., Miksa, T., Hasan, A., Praetzellis, M., Lieby, P., Papadopoulou, E., Gierend, K.: Mapping of madmps to funder templates (Jul 2020), https://doi.org/10.5281/zenodo.3944458, this report is a summary of work done by the TigTag team at the RDA Hackathon on machine- actionable Data Management Plans that took place between 27th and 29th May 2020.
3. DDI Alliance: Machine-actionable (2021), https://ddialliance.org/taxonomy/term/198, accessed 06 August 2021
4. Engelhardt, C., Enke, H., Klar, J., Ludwig, J., Neuroth, H.: Research data management organiser. In: Proceedings of the 14th International Conference on Digital Preservation, iPRES 2017, Kyoto, Japan, September 25 – 29, 2017 (2017)
5. Foidl, R., Brugger, L.: Evaluation of maDMPs using SPARQL (Jul 2021), https://doi.org/10.5281/zenodo.4997671
6. Gregory, A.: Data Management Planning and the Data Documentation Initiative (DDI). Tech. Rep. Version 1.0, Data Documentation Initiative (2013), https://ddialliance.org/publication/data-management-planning-and-the-data-documentation-initiative-ddi
7. Jones, S., Pergl, R., Hooft, R., Miksa, T., Samors, R., Ungvari, J., Davis, R.I., Lee, T.: Data management planning: How requirements and solutions are beginning to converge. Data Intelligence **2**(1-2), 208–219 (2020), https://doi.org/10.1162/dint_a_00043
8. Miksa, T., Jaoua, M., Arfaoui, G.: Research Object Crates and Machine-actionable Data Management Plans. In: DaMaLOS – First Workshop on Data and Research Objects Management for Linked Open Science: Co-located at the International Semantic Web Conference ISWC 2020. PUBLISSO (Nov 2020), https://dx.doi.org/10.4126/FRL01-006423291
9. 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
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. RDA: WG DMP Common Standards Case Statement. Tech. rep., Research Data Alliance (2017), https://www.rd-alliance.org/groups/dmp-common-standards-wg
12. Science Europe: Practical Guide to the International Alignment of Research Data Management – Extended Edition (Jan 2021), https://doi.org/10.5281/zenodo.4915862
13. Simms, S., Jones, S.: Next-Generation Data Management Plans: Global, Machine-Actionable, FAIR. International Journal of Digital Curation (IJDC) **12**(1), 36–45 (Sep 2017), https://doi.org/10.2218/ijdc.v12i1.513
14. Simms, S., Jones, S., Mietchen, D., Miksa, T.: Machine-actionable data management plans (maDMPs). Research Ideas and Outcomes **3**, e13086 (Apr 2017), https://doi.org/10.3897/rio.3.e13086
15. Simms, S., Jones, S., Miksa, T., Mietchen, D., Simons, N., Unsworth, K.: A Landscape Survey of ActiveDMPs. International Journal of Digital Curation (IJDC) **13**(1), 204–214 (Dec 2018), https://doi.org/10.2218/ijdc.v13i1.629

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