

Response to peer review, Data Science submission 706-1686

Response to open peer review <<https://datasciencehub.net/paper/packaging-research-artefacts-ro-crate>> of “Packaging research artefacts with RO-Crate” submitted to Data Science.

We thank the reviewers for their attentive and detailed comments, and we very much appreciate the time and care they have given that has helped us improve the article.

Response to Reviewer 1

Open Review was by Angelo A. <<https://orcid.org/0000-0002-4763-3943>>

We thank Angelo A for the comprehensive comments that we respond to below.

Summary of changes

The reviewer raised important points on ensuring RO-Crate targets its aim of simplicity. This prompted us to include an explicit description of our user requirement and feedback methodology, with references to specific sections.

We have reflected this discussion in the following places:

- A new [section 2.1 Development Methodology](#) that explicitly describe this requirement/feedback methodology in the article
- In section [3 RO-Crate Tooling](#), by clarifying that the role of the tools aim to make creating and managing RO-Crates simple.
- A new section [7 Future Work](#) that outlines some new directions and plans for RO-Crate 1.2 and a future user study linked to EOSC.

In addition, the reviewer has helped us improve the article with:

- Language/typography fixes in [section 2](#), in [Figure 1](#), [section 4.4](#)
- Missing [figure 4](#) restored
- Citations added in [section 3](#), [section 5.3](#)
- Clarifications added [section 2.2.3](#),
- Consistent section numbers HTML/PDF, added more cross references

Detailed response per comment

> I encourage the authors to add section numbers. Without them, it becomes hard to even scroll the document and reach the right location.

We agree - we have added section numbers to both PDF and HTML.

> **Typos:**

> - In section RO-Crate. “abuot” -> “about”

Fixed in [section 2](#).

> - *Figure 1 caption. "The RO is described within a ,". The object of the sentence is missing.*

Fixed in [Figure 1](#) - this was a Markdown error.

> - *In Data Entities are described using Contextual Entities section, you wrote "person, but secondarily also a web page and their academic work." I think the sentence is incomplete.*

The sentence was indeed incomplete. We have changed it in [section 2.2.3](#) to:

For instance <https://orcid.org/0000-0002-1825-0097> is primarily an identifier for a person, but secondarily it is also a Web page and a way to refer to their academic work.

> - *In Table 1, the tool SCHeMa has the squared parenthesis signalling a citation, but there are question marks.*

Fixed in [section 3](#).

> - *In Machine-actionable Data Management Plans section, there is a left curly bracket ({}).*

Fixed in [section 4.4](#).

> - *In Institutional data repositories – Harvard Data Commons, the reference to figure 3 is missing*

Fixed [section 4.5](#) with reference to (now) [figure 4](#).

> - *In Institutional data repositories – Harvard Data Commons (PDF version) figure 3 is missing. In the HTML version, the figure 3 is available.*

We have fixed this inconsistency (caused by graphics file extensions), Harvard Data Commons is now consistently [figure 4](#).

> - *In Packaging Workflows, there is `[cite?]' signalling a missing citation.*

We have reworded the sentence in [section 5.3](#) to further clarify, adding the missing citation:

Indeed, in the absence of a specific requirement and an editorial policy to back it up – such as eliminating the reference limit – authors are effectively discouraged from properly and comprehensively citing software [114].

> *While reading the paper, I was eager to get to the evaluation section. At some point, in the Ensuring simplicity section, you state "One aim of RO-Crate is to be conceptually simple. This simplicity has been repeatedly checked and confirmed through a community review process."*

> *I was indeed curious to see if you had **run a survey** on what users thought about RO-Crate. Perhaps, a survey gathering feedback on the user experience.*

For our early development methodology we created a pool of early adopter projects from a range of disciplines and groups that we refer to in the article in new section [2.1 Development methodology](#).

We are now ready for the next stage, which is to widen user adoption coupled with user evaluation to confirm our claims of conceptual simplicity and developer friendliness. We have added [section 7](#)

[Future Work](#) where we outline our plans for developing novice-level tutorials and training events with feedback through user studies.

> Also, the concept of simplicity is bound with the skills of the practitioner. To what extent, they found it simple?

We have expanded on this in new section [2.1 Development methodology](#).

> Also, I trust that RO-Crate is far from being complete (actually, you are in the process of releasing v1.2). What do you think needs to be improved in RO-Crate? What other new features do you plan to include? How much of this is shaped by the end-users?

We have outlined the next directions and recent new contacts in new section [7 Future Work](#).

Response to Reviewer 2

By Alessia Bardi <<https://orcid.org/0000-0002-1112-1292>>

We thank Alessia Bardi for her deep dive into the article, and for thoroughly checking the RO-Crate and the formal logic statements.

Summary of changes

The reviewer found some bugs in the associated RO-Crate which we have fixed. We appreciate all the language improvements, which prompted us to revise the article readability.

- Language and typographical fixes in [section 2.1](#), [Figure 1](#), [section 2.2.1](#), [section 2.4.1](#), [section 4](#), [section 4.4](#), [section 4.5](#), [section 9.2](#)
- Language/grammar revised throughout to improve readability (thanks Alan R Williams <<https://orcid.org/0000-0003-3156-2105>>)
- Clarifications added in [section 2.2](#), [section 2.2.3](#), [Figure 2](#), [section 5.2](#), [section 7](#)
- [Listing 1](#) modified to show data entities referencing other data entities
- References added [section 2.4.1](#), [section 3](#), [section 5.3](#)
- Fixed ContextualEntity mapping in formal logic statements of [appendix A section 9.2](#)

For the benefit of readers of this response we have also below included an example of using the production rules from appendix A.

Detailed response per comment

*> The most important remark is that I had not a good experience with the RO-Crate of this publication.
> I clicked on supplementary files 1128.html and went to <https://datasciencehub.net/system/files/ds-supplementary-706-1128.html#ro-crate-metadata.json>
> - conformsTo is missing*

We apologise for this bad experience and recognise that we should have done the validation that she did on our behalf. Mea culpa - we must “eat our own dog food”!

This turned out to be a [bug in ro-crate-html-js](#) rendering, it expected an contextual entity to display the linked RO-Crate standard. We have added such an entity in order to address this issue, but

would point out this is not a formal requirement on a JSON level - see <https://www.researchobject.org/ro-crate/1.1/root-data-entity.html>

> - the link to download returns 404

This relative link was unfortunately broken by the datasciencehub.net submission website, which, out of our control, renamed the uploaded files and broke the relative links. We have checked with the IOS Press publisher, which unfortunately replied they are not able to host JSON files as supplementary material. In the accepted version we have published the corresponding RO-Crate separately using [GitHub pages](#) (with w3id.org PID) and Zenodo (DOI).

> - the RO-Crate check says it is not a valid RO-Crate

This was an unfortunate oversight in that the ro-crate-html-js rendering had a [more stringent view](#) than the specification on the requirements of license entities. We have fixed the updated RO-Crate.

> The same happens for the version deposited on Zenodo (<https://doi.org/10.5281/zenodo.5146227>) linked in the abstract.

We have updated the Zenodo deposit with the corrected RO-Crate.

> Section "Conceptual definition":

> - Change ';' with ':' in "Figure 1 shows the main conceptual elements involved in an RO-Crate;"

Fixed in [section 2.1](#).

> - Figure 1: check the caption, there are missing words at "The RO is described within a , "

Fixed [Figure 1](#), this was an unfortunate Markdown/LaTeX rendering error.

> - Conceptual definition does not clarify what's the difference between a generic Research Object (as introduced in the previous section) and a RO-Crate

We have added to [section 2.2](#):

The conceptual notion of a Research Object [11] is thus realized with the RO-Crate model and serialized using Linked Data constructs within the RO-Crate metadata file.

> Section "Linked data as a foundation"

> - Move the footnote about IRIs the first time you use the acronym, if you think the footnote is needed

Agreed, IRI footnote moved to first mention in [section 2.2.1](#).

> Section "Data Entities are described using Contextual Entities"

> The example of orcid is not very clear to me. Why would I need to include a contextual entity that describes the person? For example, as shown in listing 1, we have

> "author": {"@id": "https://orcid.org/0000-0002-1825-0097"}

> I can understand the suggestion to add

> { "@id": "https://orcid.org/0000-0002-1825-0097",

> "@type": "Person",

```
> "name": "Josiah Carberry"  
> },
```

> *for discoverability (and convenience) purposes, but following the linked data practices I could get the jsonld of the person with*

```
> curl --Location --request GET 'https://orcid.org/0000-0002-1825-0097' --header 'Accept: application/Ld+json'
```

We are aware of this capability, and we believe other readers experienced with LOD will have similar thoughts. However, this showcases an important philosophical difference between RO-Crate and traditional Linked Data practices. We've added to [section 2.2.3](#):

In RO-Crate, a referenced contextual entity (e.g. a person identified by ORCID) should always be described within the RO-Crate Metadata File with at least a *type* and *name*, even where their PID might resolve to further Linked Data. This is so that clients are not required to follow every link for presentation purposes, for instance HTML rendering. Similarly any imported [extension terms](#) would themselves also have a human-readable description in the case where their PID does not resolve to human-readable documentation.

> *Figure 2 does not clearly show how the data entities can be linked to each other. Neither does listing 1.*

The UML is a simplification of a Linked Data model with an Open World Assumption. We have modified the [Figure 2](#) title to "Simplified UML class diagram of RO-Crate" and expanded "contextual entities (which may include aggregated and non-aggregated data entities)."

We have modified [Listing 1](#) to showcase this possibility of data entities linking each others, and added "In this example about of the CSV data entity reference the ImageObject, which then take the roles of both a data entity and contextual entity"

> *Section "People"*

> *- add links to all mentioned projects.*

In [section 2.4.1](#) we have added links and citations for the projects.

> *- Change ';' with ':' in "A key set of stakeholders are developers;"*

Fixed in [section 2.4.1](#).

> *- Table 1: check missing link for SCHeMa*

Fixed in [section 3](#).

> *Section "Profiles of RO-Crate in use"*

> *- Cultural Heritages: remove the final 's'*

Removed in [section 4](#).

> *Section "Machine-actionable Data Management Plans"*

> *Pending brace at " RO-Crate {RDA maDMP Mapper"*

Fixed in [section 4.4](#).

- > *Section "Institutional data repositories – Harvard Data Commons"*
- > - *missing figure number at "in the metadata to allow their future reuse (Figure~)"*

Fixed in [section 4.5](#)

- > *Section "FAIR Digital Objects"*
- > - *the link behind "As an "abstract protocol"" returns 404*

The link was accidentally broken by figure handling code in the manuscript HTML build process. Rephrased text in [section 5.2](#) to:

The Digital Object Interface Protocol [[@doip2.0](#)] can be considered an "abstract protocol" of requirements, DOs could be implemented in multiple ways.

- > *Section "packaging workflows"*
- > - *Missing citation at "authors might be penalised for doing so [cite?]"*

We rephrased this in [section 5.3](#) to clarify, and added the missing citation:

Indeed, in the absence of a specific requirement and an editorial policy to back it up – such as eliminating the reference limit – authors are effectively discouraged from properly and comprehensively citing software [[116](#)].

- > *Section "Minimal RO-Crate"*
- > - *Align terms in the definition of minimal RO-crate (published vs datePublished)*
- > - *Check footnote 8*

Fixed to datePublished() in [section 9.2](#) to be consistent [with footnote](#).

- > *Section "Mapping to RDF with schema.org"*
- > - *The mapping for ContextualEntity(e) is missing*

Thanks for spotting this. For any interested readers, below we explore the mapping following the production rules.

The old mapping in [appendix A](#) assumed there was an explicit type(e, c) for the contextual entity, e.g. if we have:

```
E ≡ <http://example.com/>
L ≡ <https://spdx.org/licenses/CC0-1.0>
license(E, L)
```

Then following the **Lrdf** production rules:

```
license(E, L)
⇒ Relation(E, <http://schema.org/license>, L) ∧ CreativeWork(L)
```

CreativeWork(L)

\Rightarrow ContextualEntity(L) \wedge

type(L, <http://schema.org/CreativeWork>)

type(L, <http://schema.org/CreativeWork>)

\Rightarrow Relation(L, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>, <http://schema.org/CreativeWork>)

\wedge Class(<http://schema.org/CreativeWork>)

Relation(E, <http://schema.org/license>, L)

\Rightarrow Triple(E, <http://schema.org/license>, L)

\wedge IRI(E) \wedge IRI(L)

Relation(L, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>, <http://schema.org/CreativeWork>)

\Rightarrow Triple(L, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>, <http://schema.org/CreativeWork>)

\wedge IRI(L) \wedge IRI(<http://schema.org/CreativeWork>)

In this example, ContextualEntity(L) did not need to get expanded to $\mathbb{L}rd\mathbb{f}$, as we already got an implied type CreativeWork(L) from license(E, L).

Expanding back E and L we get the corresponding triples:

Triple(<http://example.com/>, <http://schema.org/license>, <https://spdx.org/licenses/CC0-1.0>)

Triple(<https://spdx.org/licenses/CC0-1.0>, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>, <http://schema.org/CreativeWork>)

However this expectation of an explicit type(e, l) was indeed not enforced in our original definition of $\mathbb{L}rocrate$.

Considering the recommendation in updated [section 2.2.3](#): “a referenced contextual entity should always be described within the RO-Crate Metadata File with at least a type and name, even where their PID might resolve to further Linked Data”, then such a requirement should be added also to the first order logic rules.

Investigating this, you also helped uncover a variable name typo in the [section 9.2](#) Class declaration:

$\forall e$ Entity(e) \Rightarrow Class(e)

In the updated manuscript, the above has been replaced with:

$\forall e$ Entity(e) \Rightarrow type(e, c) \wedge Class(c)

$\forall e$ ContextualEntity(e) \Rightarrow name(e, n)

The updated mapping to $\mathbb{L}rd\mathbb{f}$ adds this relatively harmless fallback, to ensure a triple is produced for untyped URI references:

ContextualEntity(e) \Rightarrow type(e, <http://schema.org/Thing>)

The mapping for any arbitrary ContextualEntity(E) will then proceed as in the license example.

An alternative rule, which we dismissed as it is adding RDFS inferencing, would be to represent the property <http://schema.org/name> which is defined with domain <http://schema.org/Thing>. This rule would then be triggered by the previous ContextualEntity existential requirement of n in name(e, n):

$\forall e \forall n \text{ name}(e, n) \Rightarrow \text{type}(e, \langle \text{http://schema.org/Thing} \rangle)$

> A couple of final questions, that can be addressed in authors' future work, is about the **profiles**. Could they limit the re-usability of RO-Crate and the data entities in it in different domain? Would the proliferation of profiles hinder the simplicity and set barriers to the adoption of the community tools that have been developed so far?

We appreciate these questions, and refer to the new [subsection 6.1](#) that includes:

Specialization comes hand-in-hand with the principle of *graceful degradation*; RO-Crate applications and users are free to choose the semantic detail level they participate at, as long as they follow the common syntactic requirements.

Even though there is common ground across profiles thanks to graceful degradation, the reviewer is right to point out that a proliferation would move the interoperability goal posts and risk different RO-Crate extensions diverging with different vocabularies and conventions. We have thus expanded new [section 7](#) to include:

By discovering commonalities in emerging usage (e.g. additional schema.org types), the community helps to reduce divergence that could otherwise occur with proliferation of further RO-Crate profiles.

Response to Reviewer 3

By anonymous

The reviewer provided helpful perspectives on Linked Open Data potential of RO-Crate and reminded us that we should not underplay this. We agree on the need for a user study and have added this to the planned future work, as well as highlighting the development methodology.

The suggestions on expanding the related work in a couple of directions are understandable. However, after careful consideration, we have concluded that to give justice in such a review we would need to significantly increase the scope and length of this article. We therefore rather propose to address this in a future work with a wider review, and have cited relevant existing review articles in [Section 5](#).

For the benefit of future readers we have expanded on our reasoning below.

Overview of changes

- Reference added to [Section 5](#)
- Clarified user engagement strategy in new [2.1 Development methodology](#).
- Added [Section 7 Future Work](#), including plans for a user study
- Highlighted potential for consuming RO-Crate as RDF/knowledge graph in [section 2.2.1](#), [section 4.3](#), [section 7](#)
- Expanded [section 5.3](#) on practical limitation of workflow provenance
- Clarifications in to [section 2.2.6](#), [section 6](#)

Detailed response per comment

> *Background: Incomplete or inappropriate*

In [Section 5](#) we have added a reference to newly published <https://doi.org/10.1016/j.patter.2021.100322> which we have contributed to, and which provides an extensive list of metadata standards with importance for reproducible computational research. A similar review on data packaging literature would be welcome, but we feel that could get too extensive for the scope of this manuscript as a Resource Article.

> *Although the idea behind the RO-Crate seems interesting, in its current shape that is designed for a limited community with limited purposes.*

> *The impact of RO-Crate for large uptake and long term could be formulated better.*

We realised from this comment and other reviewers that we should be much more explicit in our user engagement strategy, which was to start with a small group and build firm foundations and making sure that the approach worked before we set out for wider adoption. We have added this in a new section [2.1 Development methodology](#).

It is right to point out that RO-Crate has a relatively small user base today, however it is also a fairly broad base in terms of scientific domains involved, which is part of our argument for RO-Crate being applicable for general purpose description of research artefacts.

We have added [Section 7 Future Work](#) on how we are working on increasing uptake, current outreach activities and new connections.

> *In the related work section, a better explanation and connection of RO-Crate to the relevant and known platforms is expected.*

We have focused this article on RO-Crate as a platform-neutral mechanism to capture data, metadata and their context. The sphere of Open Science-supporting platforms/services is very exciting and important for progress on the practice of FAIR and reproducible data, but we are concerned that comparing a metadata format with these varied platforms could get confusing to the reader.

> *Though a resource paper, there is no **evaluation or statistics** for comparison of RO-Crate to any other similar service. Even a **user study** from the LOD community would be helpful.*

To clarify this we have added to new [section 2.1 on our user methodology](#) that we have stabilized the RO-Crate specification, and are now at the stage of building wider adoption.

We agree on the suggestion of a user study, particularly after development of tutorials. We are planning a user study with the Linked Open Data community or as part of EOSC Bring-your-own-Data training events, and have included this in the new [section 7 Future Work](#).

> *- Following the given example, how do the authors deal with the massively missing parts of the output files, sequence analysis code, resulting data and reports of the existing research results of the past or even future? not every code and related artefacts of research outcomes are always available. Providing a statistic from repositories will be helpful.*

We recognize this limitation, and have expanded [section 5.3](#) to add:

It is important to note that having all these computational details in order to represent them in an RO-Crate is an ideal scenario -- in practice there will always be gaps of knowledge, and exposing all provenance details automatically would require improvements to the data sources, workflow, workflow engine and its dependencies. RO-Crate can be seen as a flexible annotation mechanism for augmenting automatic workflow provenance. Additional metadata can be added manually, e.g. for sensitive clinical data that cannot be publicly exposed¹, or to identify prototype software without a web presence. This inline *FAIRifying* allows researchers to achieve “just enough FAIR” to explain their computational experiments.

¹: FAIR principle A2: *Metadata are accessible, even when the data are no longer available.*

[\[10.1038/sdata.2016.18\]](https://doi.org/10.1038/sdata.2016.18)

We recognise the value of gaining understanding of how RO-Crate users deal with missing data and descriptions. This can be discussed as part of the user study mentioned above, which will be a good opportunity to get direct feedback from users that might not be fully familiar with RO-Crate.

Regarding statistics from existing repositories, we have left out of the scope of this article the analysis of the current practices in Open Science Registries and repositories. This is, however, an intriguing line of work that may motivate even further the need for RO-Crates.

> - Although the primary purpose of the RO-Crate specification is to assist developers in leveraging LOD, how can individual researchers benefit from RO-Crate or is it just a community-wide service for LOD developers?

Individual researchers benefit from RO-Crates by having access to bundles of information with connected resources, in context (and through the applications aimed to help their consumption). While this is available in a structured human-readable HTML format, RO-Crate JSON is intended to be machine-readable, meaning software is needed to consume it, primarily with existing RO-Crate tools ([section 3](#)), general Linked Data tools ([section 2.2.1](#)) or alternatively with new code at JSON level.

We have thus so far, in the development of specifications for writing RO-Crate as JSON, primarily targeted developers who may not be familiar with Linked Open Data principles, but which are familiar with Web technologies. We describe this further in new [section 2.1](#).

As part of this we have also prototyped user-level applications for individual researchers (e.g. Describo) as well as integration into data platforms. The next step, as shown in new [section 7](#), is to further develop both of these aspects along with training material to onboard wider communities of users which are not interested in aspects of JSON serialization.

> - It is mentioned that ‘RO-Crate is used across multiple areas’, but how can other less-technical domains benefit from RO-Crate?

Firstly, we would like to object to the term “*less technical domains*”, as for instance humanities, palaeontology, art, history, psychology and language studies are all examples of non-STEM domains

which are adapting digital data management practices in their every-day research, as well as using high tech instruments for a variety of purposes.

A large part of RO-Crate's work by the Australian partners is focused on this *eScience* cross-over with humanities and social sciences, both by building user-level tooling for describing individual data files (Describo, ro-crate-html-js, ro-crate-excel) as well as integration into existing data platforms like PARADISEC ([section 4.3](#)).

One challenge for a new domain is to use RO-Crate beyond the bibliographic level, to form what is effectively a new RO-Crate profile with additional contextual types and properties that are meaningful for their desired metadata.

We have found that, to a large degree, existing terms from schema.org can be used, however their match is not always obvious to the domain scientist and may still need help from a LOD expert. Tools like Describo help by providing a user interface for a given type, and we have experimented with supporting the full schema.org vocabulary without otherwise dealing with low level Linked Data details.

However we did find that, without the filtering from a profile, it can be overwhelming for users to navigate or understand how a type could be used. For instance, <https://schema.org/HowTo> is seemingly explained as a generalization of food recipes which sounds unrelated to science, but on closer inspection this type can work equally well to structure the steps of an industrial biochemical process or how to prepare an insect specimen in a museum collection.

When we identify such common solutions within the RO-Crate Community we "lift" a practice from specific profiles to the general RO-Crate specification, adding the explanation "glue" on how it would apply to different domains, e.g. for [equipment used to create files](#).

While we have added some of these details to [section 2.2.6](#) and [section 7](#), we plan to expand further on this in a separate future work detailing RO-Crate profile development and formalization.

> - *How RO-Crate relates and differs from ORKG (<https://www.orkg.org/orkg/>) and other similar platforms?*

ORKG, as stated in <https://www.orkg.org/orkg/about> and <https://www.orkg.org/orkg/about/2/Features>, focuses on scholarly literature, the description of granular elements reported in scientific articles and the comparison across articles.

RO-Crate aims at packing all sorts of research outcomes as a coherent unit that can be FAIRly shared, and does not require a corresponding scholarly article.

In terms of granularity, an RO-Crate is generally a structured "map" to other digital resources, which themselves convey knowledge in other formats, rather than aiming to be a container of all the captured scientific knowledge.

A knowledge graph can be generated/derived from multiple RO-Crates and queried in detail, which can be particularly powerful if the crates follow the same profile. We have perhaps not sufficiently emphasized this opportunity in the [current specification](#), although the manuscript briefly suggests use of SPARQL queries in [section 2.2.1](#) and [section 4.3](#).

We will explore the consumption using knowledge graphs further in developing Jupyter Notebook tutorials, and have highlighted this in new [section 7 Future work](#).

With relation to ORKG and RO-Crate, we do however not see how we can do a fair comparison, as we think they are pursuing different aims at different granularities. Other related work in this field, delving into formalizing arguments from literature and representing individual scientific facts, include [nanopublications](#) and [FAIRSCAPE](#). Expanding the related work would naturally then also need to cover these, which we deem out of scope for this work as a resource article.

> - RO-Crate is presented as 'a structured archive' and not called a knowledge graph (KG) as it seems to be more than a KG, but looking into the details of the service, it looks like a multimodal scholarly KG - what are the authors' viewpoints here?

We like the categorization of *multimodal scholarly knowledge graph*, and have taken the liberty of adding this point to [section 6](#).

We think of RO-Crate mainly as a way to structure data and its metadata, but clearly the power of knowledge graph is one reason why we have adapted Linked Data technologies. We have emphasized this point stronger in updated [section 2.2.1](#).

In new [section 7 Future Work](#) we have also highlighted the plans for demonstrating more of consumption using knowledge graph technologies:

We will also collect and build examples of *RO-Crate consumption*, e.g. Jupyter Notebooks that query multiple crates using knowledge graphs.

Meta-review by editor

> I have to confess it has been a pleasure reading this paper since it presents, with a lot of details, an important resource for the scholarly community at large, in particular in the context of current practices (such as those concerning FAIR and EOSC).

> All the reviewers agreed on this aspect, even if they provided some insights and questions that should be appropriately addressed in the camera-ready and that I will carefully check before publication.

We thank the editors and the reviewers for the insightful comments and suggestions, which we feel have significantly improved the article. We hope you will find our edits to address the issues raised to your satisfaction.