

Dear Guest Editors of the SAVE-SD Special Issue of the Data Science journal,

Please find enclosed the camera ready version of our manuscript entitled "Enabling text search on SPARQL endpoints through OSCAR".

We would like to thank all the reviewers for their insightful comments and suggestions. The paper has been carefully revised accordingly to their concerns. Please find all the answers to the specific issues raised by the reviewers as follows.

Best regards,

Ivan Heibi
Silvio Peroni
David Shotton

Reviewer 1

Reviewer's comment: Doesn't really A/B test against the other text search or query-building engines described in Related Work.

Authors' response: The usability of OSCAR has already been treated in our previous publication through a dedicated user testing session. Nevertheless, after reviewing the other notes received from the reviewers we decided to write a dedicated subsection in "Related works" to compare OSCAR with the searching/query-building tools presented, based on a set of relevant features/aspects. A comparison test with different tools is out of the scope of the present publication.

Reviewer's comment: A set of textual errors.

Authors' response: We have fixed all the notes.

Reviewer 2

Reviewer's comment: The title only focuses on one aspect of the tool (text search) and could be modified to indicate that OSCAR has more general capabilities.

Authors' response: OSCAR inputs are only textual in both types of searches: free-text search, and the advanced search. By writing "Enabling text search ..." we mean both searching options handled.

Reviewer's comment: Full citation of the previous version is not appropriate for an abstract and should be removed.

Authors' response: The hyperlinks in the abstract have been removed, although they are cited later in the paper.

Reviewer's comment: The five-point list of enhancements which is included in the Introduction should be moved to section 3 or the summary and conclusions.

Authors' response: This work is an extension of a previous paper, presented at the SAVE-SD 2018 workshop. Thus, we think it is appropriate to clarify since the beginning which additions we have made in this extended version of our paper.

Reviewer's comment: ELDA (the Epimorphics Linked Data API) is an additional general purpose tool for masking a SPARQL endpoint, with a simple configuration method, which should perhaps be added to the review of related work.

Authors' response: We have added and discussed ELDA in the 'Related works' section. In addition, ELDA has also been included in the comparison table.

Reviewer's comment: There is no real comparison of the capabilities of OSCAR with related tools - just a relatively cursory overview. A tabulation of the features of each compared with OSCAR would be helpful.

Authors' response: We have created a dedicated subsection in "Related works" in which we compare OSCAR with the other searching and query-building tools presented, based on a set of relevant features/aspects.

Reviewer's comment: Section 3 contains some rather detailed technical description of the tool. Some of sections 3.2 and 3.3 is possibly too much for a journal or conference paper, and perhaps should be left as references to technical documentation.

Authors' response: We think this part is essential to let users understand some technical aspects of OSCAR, and understand how to correctly customise and re-use it for their own purposes.

Reviewer 3

Reviewer's comment: in relation with the novelty of contribution, what does OSCAR provide that the others do not? A deeper comparison between OSCAR and the tools mentioned in the related work would substantiate the novelty of contribution from the scientific perspective.

Authors' response: We have created a dedicated subsection in "Related works" in which we compare OSCAR with the other searching and query-building tools presented, based on a set of relevant features/aspects.

Reviewer's comment: In relation to OSCAR usefulness, I have no problem believing it is useful. However, showing the Usage statistics retrieved from the OpenCitations website logs is a very weak proof of its usefulness. It does not distinguish between the usefulness of the content served by OpenCitations and the actual usefulness of OSCAR. Of course, it is better than nothing, but actual user satisfaction using OSCAR should be more systematically investigated.

Authors' response: We have reworded the text in order to explicitly mention that the usage statistics reported demonstrate the OSCAR 'usage' rather than its 'Usefulness', as wrongly reported in the previous version of this paper.

Reviewer's comment: Selected Keywords (i.e., OpenOffice, ODT to RASH) do not relate to the content of the paper.

Authors' response: We have now changed the set of keywords to the following ones: OSCAR, OpenCitations, OCC, COCI, SPARQL, Free-text search, Scholarly data, Advanced search

Reviewer's comment: please state explicitly that OSCAR can be applied beyond scholarly data portals, as it is configurable on any SPARQL endpoint and RDF schema.

Authors' response: We have added a sentence in the final part of the Introduction section to emphasize this aspect.

Reviewer's comment: In the related works, I would consider citing YASGUI as an example of an interface for semantic web literate. Laurens Rietveld, Rinke Hoekstra, The YASGUI family of SPARQL clients. Semantic Web 8(3): 373-383 (2017)

Authors' response: We have added YASGUI in the discussion about the related works, and added it to the new comparison table.

Reviewer's comment: Configuration examples might be not easy to grasp, I think a reference to the configuration instructions and the inclusion of some comments in the configuration file might help.

Authors' response: We currently have a brief documentation on the OSCAR repository (see <https://github.com/opencitations/oscar/tree/master/doc>). We plan to extend the current one so as to provide more details about all the available configuration options and parameters. In addition, we have generated a new file (<https://github.com/opencitations/oscar/blob/master/static/js/conf-guidelines.js>) to summarise such configuration parameters.

Reviewer's comment: I suggest adding a discussion of the limitations and applicability of OSCAR. For example, in a separate section. In such a section, the authors might want to answer the following questions: When configuring OSCAR is less handy than writing a custom user interface? Under what kind of licence OSCAR

is made available? is there any assistance in case one has any difficulties when configuring/using OSCAR? etc.

Authors' response: We extended the text so as to address all the points raised. In particular, we stated that the current license of OSCAR is ISC, and we have added some final sentences addressing users that wish to use and integrate OSCAR with their services.

Reviewer's comment: One thing I've noticed playing with the results from <http://opencitations.net/search?text=machine+learning> : if one sorts the results by the number of citations, and limits the number of results visualized, he gets the first ten results in the result set, not the first ten most-cited papers. This is extremely counterintuitive! I suggest to fix it in the next release of OSCAR.

Authors' response: Thanks for the hit, it is a bug. We have added an issue in the GitHub repository so as to address this point in future releases.

Reviewer's comment: The Data science journal requires that all used and produced data are openly available in established data repositories, as mandated by FAIR and the data availability guidelines (<https://journals.plos.org/plosone/s/data-availability>). As far as I understand the guidelines have not followed for the statistics regarding the accesses to OSCAR used in section 5. Please fix it or make clear how you have met the availability guidelines.

Authors' response: The data of all the charts have been uploaded on Figshare, so as to address the data availability guidelines. We have mentioned this and appropriately cited this new source in the paper.

Meta-Reviewer

Reviewer's comment: Please address the concerns by all reviewers about including a more in-depth comparison with other tools and describe what are OSCAR's distinctive advantages. For example, consider including a table comparing OSCAR's functionality against the functionality provided by other tools. Do include in the comparison of the new tools suggested by reviewers. Please, also address the distinction between the tool and its data content and provide ways of evaluating both. In addition, please make available all the material required to evaluate the tool and content (e.g. the usage statistics).

Authors' response: We have added a dedicated subsection in "Related works" in which we compare OSCAR with the other searching and query-building tools presented, based on a set of relevant features/aspects.

Reviewer's comment: As regards the availability of associated material, while OSCAR development is open and you provide the GitHub repository

(<https://github.com/opencitations/oscar>), there are currently no releases in that repository. I recommend you create a release and use the GitHub/Zenodo association to obtain a DOI and make the code citable (see <https://guides.github.com/activities/citable-code/>). Thus, please include a citation to your software using the Zenodo DOI.

Authors' response: We have published the first release of OSCAR (version 2) using Zenodo/Github, getting the DOI <https://doi.org/10.5281/zenodo.2587541>, which is now cited in the paper.