Editorial

Open Science data and the Semantic Web journal

Pascal Hitzler [a], Krzysztof Janowicz [b], Cogan Shimizu [a], Lu Zhou [a] and Andrew Eells [a]

[a] *Department of Computer Science, Kansas State University, USA*
E-mails: hitzler@ksu.edu, cogan.shimizu@coganshimizu.com, luzhou@ksu.edu, andreweells@ksu.edu
[b] *Geography Department, University of California, Santa Barbara, USA*
E-mail: janowicz@ucsb.edu

Establishing Open Science as the default practice in the academic community is critical. Reporting a research result is more than just the manuscript. It is the actual process undertaken, the data (and metadata), the implemented software (and environment), and many other components. The manuscript provides insights, but is still just that – a report. In addition to these more traditional components, it is also important to ask questions pertaining to transparency and reproducibility, such as

– Why was a paper accepted?
– Who made these decisions and how were they made?
– To a reasonable extent, could the results be reproduced?
– What infrastructure (e.g., hardware and software dependencies) is required for reproducing the results?
– Are there indications of data tampering and/or bias?

And, most importantly, *Where can these resources be found*? By making the connections between these artifacts *explicit* and *accessible* (i.e. through Open Science practices, and in the spirit of the FAIR principles [3]), it helps to democratize research by lowering the barrier of entry to understanding (and replicating) cutting edge research, while simultaneously accelerating novel research by reducing the need to re-implement software or re-collect data.

Connecting these components and data in a useful and understandable way is a core mission of our field. We are concerned with data openness and sharing, including methods for knowledge graph and ontology development, deployment, and usage [1,2]. Open Science Data is thus a topic that is of concern for the Semantic Web research community. However, like many fields, the Semantic Web community has mostly addressed this concern through informal processes.

As a step towards addressing the broader Open Science Data challenges, the Semantic Web journal is implementing new requirements on the provision of resources – data and software – that accompany paper submissions. The corresponding changes will be rolled out shortly. While the details are not set in stone, and will likely be modified as we gain more experience with the process and guidelines, we anticipate the following set-up.

– Authors will be expected to provide data and software relevant for assessing a submission and for replicating experiments, whenever is feasible to do so.
– If relevant data or software was not included, the review will assess whether there are convincing reasons for this.
– Reviewers will include in their assessments quality, accessibility, and organization of the provided

data or software, as well as an indication whether the provided materials appear to be sufficient for the replication of experiments.

– Data and software, in particular after acceptance of a manuscript for publication, will be expected to be available long-term, without modifications, under stable URLs, while at the same time it will be backed up by the journal for long-term reference purposes.

We are aware that the modification of the established peer review procedures can initially cause additional and sometimes unexpected burden on our authors, reviewers, and editors. As such, we intend to keep them as minimally disruptive as possible, while acknowledging the need of advancing Open Science.

The specific requirements and other updated information will be kept up-to-date on the journal website, where we will also provide changes to our process in detail. In the meantime, please do not hesitate to contact the editors-in-chief if you have any questions or concerns.

References

[1] P. Hitzler, A review of the semantic web field, *Communications of the ACM* **64** (2021), 76–83. doi:10.1145/3397512.

[2] P. Hitzler and K. Janowicz, Linked data, big data, and the 4th paradigm, *Semantic Web* **4**(3) (2013), 233–235. doi:10.3233/SW-130117.

[3] M.D. Wilkinson et al., The FAIR guiding principles for scientific data management and stewardship, *Scientific Data* **3** (2016), 160018. doi:10.1038/sdata.2016.18.