Wright State University

CORE Scholar

Computer Science and Engineering Faculty Publications

Computer Science & Engineering

Open and Transparent: The Review Process of the Semantic Web Journal

Krzysztof Janowicz

Pascal Hitzler
pascal.hitzler@wright.edu

Follow this and additional works at: https://corescholar.libraries.wright.edu/cse

Part of the Bioinformatics Commons, Communication Technology and New Media Commons, Databases and Information Systems Commons, OS and Networks Commons, and the Science and Technology Studies Commons

CASE STUDY

Open and transparent: the review process of the Semantic Web journal

Krzysztof Janowicz, University of California, Santa Barbara, USA

Pascal Hitzler, Kno.e.sis Center, Wright State University, USA

ABSTRACT: While open access is established in the world of academic publishing, open reviews are rare. The Semantic Web journal goes further than just open review by implementing an open and transparent review process in which reviews are publicly available, the assigned editors and reviewers are known by name, and are published together with accepted manuscripts. In this article we introduce the steps to realize such a process from the conceptual design, over the implementation, a overview of the results so far, and up to lessons learned.

Introduction

The Semantic Web [1,4] is a highly multidisciplinary research area bringing together theoretical and applied researchers and companies ranging from computer and information science over cognitive science up to geography, cultural heritage research, and the life sciences. While Semantic Web research spans several topics, they all share the vision and need for the on-the-fly integration of large and heterogeneous data sets. For instance, the long-term vision of Geographic Information Science is a Digital Earth where data from multiple sources, topics, and perspectives can be queried and combined to answer complex scientific questions. Bioinformatics has a similar vision – the Digital Cell. Both rely on automatic approaches to integrate massive amounts of information. Initiated in the 90s by Sir Tim Berners-Lee, inventor of the World Wide Web, the Semantic Web offers methods and technologies to overcome the integration problem by expressing the meaning of information using so-called ontologies and formal logical deduction to discover, match, and translate between data. In recent years, research results from the Semantic Web have been integrated into numerous commercial applications and also IBM's Watson system that won in the game of Jeopardy against the best human players in early 2011 [3]. Most recently, researchers, industry, and government agencies have started to provide so-called Linked Data on the Web, which has grown into a vast and rapidly growing collection of Semantic-Web-enabled data on a wide variety of topics from governmental and scientific data to trivia such as movies and popular music [2]. Ten years of focused research funding from major agencies across the globe has resulted in a large, dynamic, active, and interdisciplinary field which draws major investments from IT companies and application developers.

The Semantic Web journal (SWJ),[1] published by IOS Press[2] and AKA-Verlag,[3] started in 2010 as a new journal entirely dedicated to Semantic Web research. Its subtitle Interoperability, Usability, Applicability refers to the three pillars of the journal. Interoperability refers to research on ontologies, reasoning, data integration, and discovery, usability refers to work on closing the gap between the Semantic Web and end-users, while applicability refers to software tools and systems as well as to application reports published by the journal to demonstrate the impact of Semantic Web technologies in large-scale applications. The editorial board consists of emerging, young scholars flavored with some key Semantic Web seniors.

---

[1]   http://www.semantic-web-journal.net/
[2]   http://www.iospress.nl/
[3]   http://www.aka-verlag.com/

In contrast to the vast majority of computer science journals, the Semantic Web journal combines a pre-press open access with an open and transparent review process. All submitted manuscripts are published online and are accessible to the public without registration. The responsible editors and solicited reviewers are given by name and are published in the header of accepted papers. Additionally, registered and pre-approved users can add open, not-solicited reviews via the journal's Web page. Authors can submit their responding letters to the reviewers in public to document the changes made when revising their manuscript if they like to. The journal follows an opt-out approach which allows reviewers to stay anonymous on request. To implement a *digital forgetting* [5] the reviews and manuscripts of rejected submissions are available online only for a limited time. While the Semantic Web journal makes pre-press manuscripts available online, it is not a classical open-access journal. The final, typeset versions are only available on subscription via IOS Press as printed issues or PDF files. The journal appears twice a year and will start to appear quarterly in 2012. Within the first 18 months, more than 150 manuscripts have been submitted for open and transparent review at the Semantic Web journal. The journal kicked-off with a double issue of 23 fully reviewed vision statements of the editorial board members.

Rationales for open and transparent reviewing

The Semantic Web journal has developed an open and transparent review process to improve accountability. First, peer reviews are a crucial component of quality assurance in science. By investing a substantial amount of time in carefully reviewing a submitted manuscript, reviewers directly contribute to the quality of the final article. In many cases, their contributions go beyond quality control and reviewers share their perspectives, visions, and insights with the authors. However, due to the dramatically increasing time pressures on potential reviewers, reviewing can be seen as a time-consuming exercise with rather small incentive. This development has a negative impact on the elaborateness and overall quality of the average review, a trend which can already be observed for workshop and conference reviews, which are an important part of computer and engineering science. Reviews for the Semantic Web journal are non-anonymous and publicly available by default. Additionally, reviewers of accepted manuscripts are explicitly acknowledged in the final, printed versions of the articles. Consequently, reviewers receive visibility for their contribution and the quality of their work can be assessed by others, thus providing an incentive for high-quality reviewing.

Second, the review process has to minimize a positive or negative bias of reviewers and editors and especially it has to uncover conflicts of interest. Just by choosing reviewers, editors can influence the likelihood that a paper gets accepted or rejected. However, keeping track of such things as coauthorship, supervision, and joint research or service between authors and potential reviewers over many years is not reasonable. Even if this were possible, it cannot address the problem that the same reviewers may be assigned to the papers of a specific author over and over again. This would require that journals and conferences share metadata about their review processes. In theory, anonymous reviewers could promote or reject manuscripts not based on quality but on other, conflicting interests, e.g., whether the reviewer's work is cited in the submitted manuscript. The transparent review process of the Semantic Web journal resolves this problem by publishing the names of the reviewers as well as of the responsible editor. Consequently, the selection of reviewers can be tracked by authors, readers, and also editors of other journals.

Another important side-effect is that editorial decisions become more transparent as well. The objectivity of editors in accepting or rejecting manuscripts can be judged by anybody because of the public availability of the reviews.

Third, the classical blind-review process offers only partial anonymity for reviewers. Editors, steering committee members, program chairs, and sometimes other reviewers have access to the reviews, names of the reviewers, and the resulting decisions. As a consequence, most editors and organizers of conferences and workshops are hesitant or not allowed to submit their own work to these events or journals. Unfortunately, this makes editing work less attractive to scholars as they cannot contribute their own work. While not a key issue for established research fields, this is troublesome for emerging or highly specialized research areas such as those usually addressed at workshops. Therefore, submissions by the organizers are often tolerated in workshops and even journal special issues. This puts the burden on the reviewers as they have to rate a manuscript from one of the organizers or editors. No matter whether such papers get accepted or not, the underlying decision process is not transparent for other authors contributing to the same special issue or workshop. The Semantic Web journal review process makes such decisions transparent and, therefore, does not provoke allegations.

Summarizing, an open and transparent review process eases the assignment of editors and reviewers to manuscripts, acknowledges the contribution of reviewers in improving scientific articles, and moreover shields editors and reviewers from conflicts of interests and allegations.

Conceptual design

While open or voluntary reviews have been adopted by journals, as far as we know the precise model presented above has not been proposed nor implemented before. Consequently, the workflow is not covered by any journal management system. The basic principles of an open and transparent review process have to be accompanied by detailed instructions for editors, reviewers, and authors. Conceptually, the workflow for an open and transparent review process should be designed as follows.

Paper submission, selection of editors and reviewers, solicited reviews, as well as the communication between editors, reviewers, authors, and the publisher should follow established paths and use an existing journal management system. These systems have been established over many years and provide (more or less) intuitive, Web-based user interfaces. The new review process should put no additional burden on any of the involved parties.

The Editors-in-Chief should be responsible for making incoming submissions available online together with metadata such as the assigned editorial board member. Once all solicited and voluntary reviews are available, they should be published together with the editorial decision. The Editors-in-Chief are also responsible for ensuring that the names of anonymous reviewers are not disclosed and that voluntary reviews follow certain rules, e.g., use appropriate language.
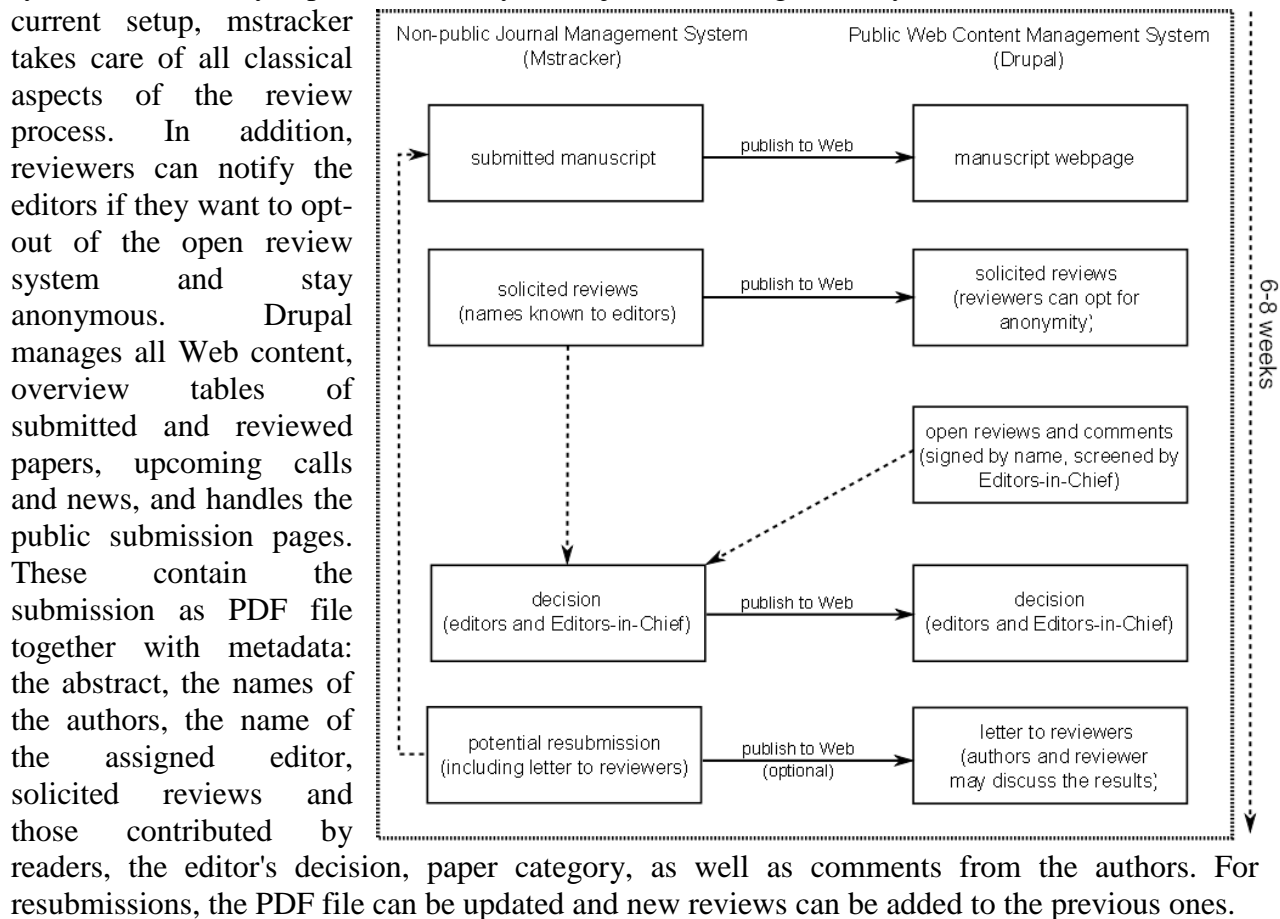
Based on discussions with authors, reviewers, and editorial board members, and independent of whether articles were accepted or not, reviews describing shortcomings of earlier versions of manuscripts should not be publicly available forever. Therefore, rejected papers and their reviews have to be de-published after a certain time. However, all data should remain visible to the editorial board members.

In contrast to the classical review process, where only final versions of accepted papers are announced on the journal's web page, information on the Semantic Web journal's site is updated

frequently, e.g., new submissions and revisions are uploaded, editors are chosen, solicited and voluntary reviews are published, or authors react to the reviewers' comments. To ensure that these changes are communicated to the research community, a push-based news system as well as overview tables for various information have to be established.

Implementation

The conceptual design is realized using two major components, mstracker[4] as journal management system and a highly customized version of the free and open source Drupal[5] 6 Content Management System. A flowchart representing these systems is given below. Both systems are entirely separated and any other journal management system could be used. In our current setup, mstracker takes care of all classical aspects of the review process. In addition, reviewers can notify the editors if they want to opt-out of the open review system and stay anonymous. Drupal manages all Web content, overview tables of submitted and reviewed papers, upcoming calls and news, and handles the public submission pages. These contain the submission as PDF file together with metadata: the abstract, the names of the authors, the name of the assigned editor, solicited reviews and those contributed by

readers, the editor's decision, paper category, as well as comments from the authors. For resubmissions, the PDF file can be updated and new reviews can be added to the previous ones.

All submission relevant data is entered by the Editors-In-Chief using a manually defined content type form – a part of Drupal's build-in customization functionality. Each field of the form is linked to additional actions using Drupal's Rules, Twitter and Content Construction Kit (CCK) modules. For instance, the editorial board members can be selected from a pull-down list based on their user role defined during registration. Adding a new submission, changing the status of a paper, setting the editor's decision, submitting a public comment, or uploading a review triggers different tweets. These tweets are used as a push-based notification system and can be followed on Twitter;[6] the last 5 tweets are also directly visible on the journal's Web page. For instance,

---

[4]   http://mstracker.com/
[5]   http://drupal.org/
[6]   http://twitter.com/SW_Journal

adding a review leads to the following automatically generated tweet: A review has been added (or updated) to a submission to the Semantic Web Journal (#swj): http://tinyurl.com/3y6mycg.

Submission pages are publicly available during and after the review and readers can request an account at the Semantic Web journal to add their non-solicited reviews using Drupal's comment function. Authors can reply to these comments and upon revision of a manuscript they can also upload their letters to the reviewers as public comments. However, this is not mandatory and these letters can also be submitted using mstracker. Drupal offers fine-grained user right control and, therefore, all user reviews are manually approved by the Editors-In-Chief, while editorial board members can directly publish their comments. An important implementation decision was whether comments should be made available during the review process or only after all solicited reviews have arrived. To motivate the research community and avoid that users have to wait weeks before their non-solicited reviews are published, the comments are made available immediately after approval by the editors. Comments cannot be published anonymously and users have to register with the full name and an email address pointing to their affiliation. The Editors-In-Chief contact new users if their data cannot be verified.

Accessing, browsing through, and de-publishing submissions is another important implementation detail and made possible using Drupal's Views module. The Semantic Web journal Web site offers several pre-defined views, in the form of tables, to provide overviews of all submitted manuscripts, all reviewed manuscripts, all submissions that are currently under review, all submissions pre-sorted by number of views, as well as a table of all manuscripts that are awaiting a resubmission. In each view, users can sort the papers by title, last action date, editorial decision, or total views. Additionally, the Under Review view shows which papers have comments (and how many) as well as whether solicited reviews have been published. We decided not to display the number of page views per paper in the Under Review view, however, they can be seen by directly accessing a submission. This avoids sorting manuscripts that are under review by views to estimate the popularity or interest in a paper. There is no mechanism that would prevent authors from visiting their submission page over and over again to raise the number of views. Furthermore, the Semantic Web is a very heterogeneous community and some topics attract a larger group of readers than others, so the number of views is by no means an indicator for the relevance or quality of a manuscript. Nevertheless, the total view count is available in all tables listing reviewed submissions. It is motivating for the authors to see that readers are constantly visiting their submission page and downloading their papers while it cannot influence the review process.

To implement a digital forgetting and support authors in resubmitting their work to other journals or conferences, rejected papers as well as those awaiting resubmission for a substantial amount of time (usually more than 8 weeks) are de-published by the Editors-In-Chief. De-published manuscripts are still visible with their title in all views but the submission pages with all metadata (e.g., the names of authors), reviews, and PDF files cannot be accessed. Editors can still see the documents and make them available again in case a new version gets resubmitted. It is important to note that the very concept of de-publishing is, in fact, not entirely possible. Web pages can be locally stored by any user, and companies such as Google maintain public caches of Web sites. Nevertheless, one would be wrong assuming that de-publishing submissions from the Semantic Web journal's page does not accommodate authors. The Google cache, for instance, is best used to access a previous version of a Web site found via Google Search. However, de-published submission pages cannot be found this way. They are also not directly available via the

Internet Archive (archive.org). This does not mean that cached versions of de-published submission pages cannot be accessed at all, but that this requires more than plain search on the Web. We believe that it is therefore highly unlikely that reviewers and editors of other journals or conferences (to which a rejected manuscript may have been resubmitted) would actively crawl Web caches for potential previous paper versions and reviews. In summary, while information on published submissions may remain available, it is not directly discoverable and, most importantly, there is no reason to crawl the Web for it.

Finally, we implemented additional pages for current and previous issues that link to submission pages of accepted manuscripts and to their final, typeset versions at IOS Press, as well as an overview page for currently open calls and their deadlines.

Leaving smaller patches and minor modifications from the editors aside, the following third party Drupal modules are used to realize the journal's functionality: CCK, FileField Paths, Flexifilter, Messaging, Notifications, Rules, Spam Control, Access Control, Views, Oauth, Pathauto, Skinr, Token, and Twitter.

The story so far

As of summer 2011, 18 months after the launch of the Semantic Web journal:

- ⚞ 167 manuscripts have been submitted, not counting resubmissions. 96 manuscripts have been resubmitted.

- ⚞ 23 manuscripts have been accepted for publication (not counting editorials and a first, fully reviewed issue on vision statements from the editorial board member) leading to an acceptance rate of 14%. Note that this figure does not account for the fact that some of the manuscripts currently still under review may eventually be accepted and published.

- ⚞ 3 issues have been published with a fourth issue in final preparation.

- ⚞ 212 reviewers have contributed to the journal so far, each with 1-11 reviews.

- ⚞ Accepted papers have 1381 page views on average. This number is constantly increasing as the submission pages and papers are publicly available. The most viewed paper has 2902 page views.

- ⚞ Out of 426 reviews in total (including those for resubmissions) 113 are anonymous, while 313 are not, i.e., 73% of all reviews contain the names of the reviewers.

- ⚞ The numbers above cannot be used to determine an average number of reviews per manuscript as several of the 167 manuscripts are under review, while others have been rejected during pre-screening. Accepted papers received 4.1 reviews on average.

- ⚞ Leaving notes from the editors aside, 52 comments have been published. This includes non-solicited reviews, open letters to the reviewers, and discussions between reviewers and authors.

- ⚞ The average turnaround time between the submission of a manuscript and the editorial decision is 63 days.

↳ While most papers have only been published for a few months, they have received a total of 135 citations so far (based on Google Scholar[7] and including 49 publications i.e. regular papers, editorials, as well as the initial vision statements issue).

Lessons learned

At present, there is hardly any software support for the editorial management of open and transparent review processes by existing journal management systems or content management systems. Rather, this currently requires a substantial amount of customization of these systems and indeed also the maintenance of two independent systems, namely the review management system and the Web site. While the additional workload for editors and authors is negligible, the Editors-In-Chief have to invest considerable effort for ensuring the (currently manual) information flow between the two systems, and for the administration of a customized Drupal system. Each of the additional Drupal modules follows different release cycles and the levels of support from maintainers and the community varies. For instance, a change in the Twitter Authorization API, required changes in Drupal modules and different fixes that were not immediately available. Due to the release cycles of different modules, changes in one module can render another one unstable until an update is available. Therefore, we test updates to Drupal before installing them on our production system. This should not be understood as a criticism of Drupal and the developer community. To the contrary, the complex workflow of the Semantic Web journal could very well be modeled and realized based on Drupal if further and substantial customizations of the system were put in place. This means, however, that technical skills and additional administration efforts are required to realize an open and transparent review process.

The Editors-In-Chief received strong support and positive feedback from the Semantic Web community in establishing the journal; however, a small number of scientists declined to review for the journal or noted that they would not contribute their own research. Two major arguments were raised, 1) open reviews may lead to conflict and pressure on reviewers not to reject work of prominent researchers; 2) authors may risk their reputation if flaws in their work are found and publicly discussed.

From our experience so far, both arguments can be rebutted but not completely rejected. The journal has an acceptance rate below 15%, and not a single paper has been accepted in the first round of reviews; i.e., with an "accept as is" decision. Papers from prominent, senior researchers have been rejected. We could not observe an increased tendency to remain anonymous in case of critical reviews. Judging whether open and non-anonymous reviews are of a higher quality is difficult and a recent study points out that no significant difference could be found [6]. What we could observe, however, was a very low amount of inadequate reviews. We only received three reviews of two reviewers that did not met good scientific standards. Surprisingly, these reviews were non-anonymous. We could also not find a single review that was written in an aggressive and non-constructive manner. In general, our impression is that open reviews lead to a milder and more constructive choice of verbalization of criticism. It has to be noted that only the review text is available to the public and the authors. The 5-point scale for criteria such as significance to field, methodology, literature review, or writing style/clarity offered by mstracker are only visible to the editors. Similar to other journal management systems, mstracker offers a text field to add confidential remarks as well. While this field is not used to provide additional review information, reviewers can communicate valuable details, e.g., regarding the certainty of their

---

[7]   http://scholar.google.com/

assessment, expertise, or the undertone of the textual review. A tendency for rather positive reviews is easily counteracted with a more rigid evaluation from the responsible editors. In two cases, the Editors-In-Chief have decided to solicit an anonymous meta-reviewer which in one case eventually led to the rejection of the manuscript.

The reputation argument was one of the reasons for de-publishing submission pages of rejected papers after a certain time. In some cases, authors requested to hide their submission page earlier than the standard 8-week period. We believe that a research community is well aware of potential shortcomings in manuscripts and understands that they are an essential part of science. We do not expect that authors may be judged on shortcomings in their work in a different way than for closed reviews or published work (which can also contain errors).

In accordance with findings from other journals, voluntary (non-solicited) reviews play virtually no role. It is rather the theoretical possibility to react to a submission which seems to be important. We expect, however, that there will be a growing culture for such open reviews. Interestingly, authors started to use the comment function to publish their response letters to the reviewers, a development that we had not foreseen while setting up the journal.

Finally, in some cases, due to the open publication of manuscripts under review, editors have been notified of potentially overlapping publications and conflicts of interest which might otherwise have gone unnoticed.

Conclusions and Outlook

Based on our experience so far, we believe that an open and transparent review process is a valuable addition to the scientific publishing landscape and that transparency is of equal importance to openness. Our approach goes further than just making reviews publicly available, but brings the choice of editors, reviewers, editorial decisions, and author's responses also into the public space. We do not argue that the classical, (double-) blind, and closed review process should be completely replaced. To the contrary, if all journals would be based on a SWJ style review process, blind and closed reviews probably would have to be introduced.

In the future, we plan to replace mstracker and combine all necessary functionality within Drupal. Work on implementing such a system has already started and is supported by IOS Press, by a State of Ohio Research Incentive, and by the University of California, Santa Barbara. The integration of Semantic Web technologies and support for Linked Data to foster retrieval and inter-linkage are also part of our research and service agenda.

Finally, to motivate readers, authors, and reviewers, the journal's editorial board will select an outstanding reviewer and paper each year.

References

[1] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. Scientific American, 284(5):34–43, 2001.

[2] C. Bizer, T. Heath, and T. Berners-Lee. Linked Data – the story so far. International Journal on Semantic Web and Information Systems, 5(3):1–22, 2009.

[3] D. A. Ferrucci, E. W. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, J. M. Prager, N. Schlaefer, C. A. Welty. Building Watson: An Overview of the DeepQA Project. AI Magazine, 31(3):59-79, 2010.

[4] P. Hitzler, M. Krötzsch, and S. Rudolph. Foundations of Semantic Web Technologies. Chapman & Hall/CRC, 2009.

[5] V. Mayer-Schönberger. Delete: the virtue of forgetting in the digital age. Princeton University Press, 2009.

[6] S. van Rooyen, T. Delamothe, S. J. W. Evans, Effect on peer review of telling reviewers that their signed reviews might be posted on the Web: randomized controlled trial. British Medical Journal, 341:c5729, 2010. DOI:10.1136/bmj.c5729