The Potential of Collaborative Document Evaluation for Science

Joeran Beel[1] and Bela Gipp[1]

[1] Otto-von-Guericke University, Department of Computer Science,
Universitätsplatz 2, 39106 Magdeburg, Germany
j@beel.org, bela@gipp.com

Abstract. Peer review and citation analysis are the two most common approaches for quality evaluations of scientific publications, although they are subject to criticism for various reasons. This paper outlines the problems of citation analysis and peer review and introduces Collaborative Document Evaluation as a supplement or possibly even a substitute. Collaborative Document Evaluation aims to enable the readers of publications to act as peer reviewers and share their evaluations in the form of ratings, annotations, links and classifications via the internet. In addition, Collaborative Document Evaluation might well enhance the search for publications. In this paper the implications of Collaborative Document Evaluation for the scientific community are discussed and questions are asked as to how to create incentives for scientists to participate.

Keywords: open peer review, citation analysis, alternative, research policy

1 Introduction

Searching and evaluating scientific publications is a time-consuming activity. Synonyms, a growing number of publications and ambiguous nomenclature impede the search for relevant documents. Sometimes nomenclature itself changes over time. Therefore, researchers using keyword-based search engines miss out documents, if they do not know all relevant keywords or authors did not use the commonly known terms in their publications.

Once a publication is found, the reader needs to assess its quality and credibility. Usually scientists assume a publication's quality from the reputation of the issuing journal. Journals in turn select publications based on their peer reviewers' recommendations. However, peer review is often criticised for leading to non-objective decisions caused by incompetent reviewers and reviewers following own interests, whether due to competition, alliances or economical reasons [1,2]. The increasing amount of interdisciplinary articles encumbers the peer review process, too. Imagine an empirical study about the influence of music on online shop visitors' behaviour. A thorough evaluation would require experts in the fields of music, psychology, neuroscience, business, computer science and statistics. Hardly any journal has access to reviewers that could competently review such an interdisciplinary paper. This is especially the case for conferences.

Due to the limitations of the peer review process, scientists attempt to evaluate the quality of a publication by its citation counts. The assumption is that the more often

publications and authors are cited, the better they are. However, citation analysis is subject to criticism as well. Citation databases are incomplete and sometimes erroneous; citation counts are spoiled by ceremonial citations, self and negative citations, cronyism, citation oblivion and the fact that authors tend to cite secondary sources rather than the original authors; authors are biased and do not cite all influences, while they sometimes cite publications they have never read [3,4]. Most importantly, citation counts can only measure 'impact', but impact does not necessarily correlate with quality [5].

Another drawback of citation analysis and peer review is their lack of capability to accomplish post-publishing quality evaluations. Once a paper is published, it is associated with the journal's reputation even if at a later point in time new insights lead to a different assessment. For instance, John Darsee published dozens of articles in reputable journals. Later, most of his articles were proven to be fraudulent or at least questionable. Nevertheless, his flawed work was cited 298 times during the following ten years. An astonishing 86% of the citations approved of his work [6]. Apparently, the citing authors were unaware of the flaws and relied on the reputation of the publishing journals.

These shortcomings demonstrate the need for improving the existing quality evaluation approaches for scientific publications. The Otto-von-Guericke University is researching 'Collaborative Document Evaluation' as part of the *Scienstein.org* project. Collaborative Document Evaluation aims to let the scientific community evaluate publications and share the gathered information for everyone's benefit.

2   Related Work

Some attempts have been made, by letting the scientific community rate and/or annotate papers – including others such as Arxiv.org, Bibsonomy.org, Naboj and Nature [7]. However, none of these attempts has been totally successful. Incentives for participation are not sufficient and the competence and trustworthiness of those participating are unclear. As a consequence, very few annotations or ratings exist and their reliability remains unclear.

Nevertheless, in other domains comparable projects have been successful. For instance, Wikipedia succeeded to let the 'crowd' create and evaluate content in a decent manner [8] and the United States Patent and Trademark Office introduced successfully the public reviewing of patent applications [9].
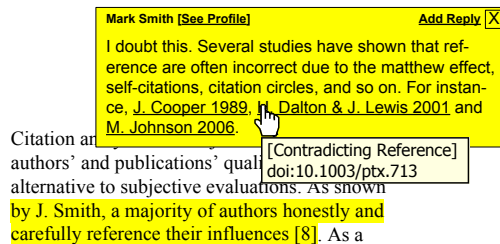
We believe that Collaborative Document Evaluation could be equally successful for evaluating publications.

3   Scienstein and Collaborative Document Evaluation

Collaborative Document Evaluation is about creating and sharing metadata of scientific papers by the scientific community via the internet [10]. The metadata gathered in the Scienstein project includes ratings, annotations, links, classifications and highlighted passages within documents. *Collaborative ratings* are quantitative ratings given in different categories such as *originality, significance, readability,*

correctness of methods and analysis and *overall quality*. *Collaborative annotations* are comments for entire documents or parts of it. They can be classified, for instance as *critique*, *addition* or *misc* and may include *collaborative links*. These links can point generally to other documents or to specific passages, just as hyperlinks do. In contrast to hyperlinks, collaborative links can also be classified, similar to collaborative annotations. *Collaborative classifications* are similar to tags, but more structured [10].

Collaborative Document Evaluation enhances document search by various techniques. Via annotations and classifications, new terms can be assigned to publications. This way, older publications can be updated with modified or currently-used terminology. Highlighted passages and annotations indicate which parts of a publication are considered particularly relevant by the scientific community. Accordingly, keyword-based search engines could attach greater significance to the words in the highlighted or annotated passages. Ratings given by researchers with similar interests

Fig 1. Annotation, Highlighted Passage and Links

can be used by a research paper recommender system [11]. Collaborative links not only show *that* publications are related to each other, but also *how* they are related, which enables a new type of search for related work.

Collaborative Document Evaluation may enhance quality evaluations in a number of ways. First, in contrast to citation counts, ratings measure the real (subjective) quality perception of the community members. Based on these ratings, the overall rating and the ratings of subgroups with similar interests can be displayed for each publication. Second, due to low entry barriers, new groups of people can act as reviewers. For instance, professionals could communicate their knowledge effortlessly via annotations. Since many readers from various research fields can participate, evaluating interdisciplinary work would be facilitated. Third, highlighted passages and annotations can provide more concise information about a publication in addition to the abstract. Finally, one community member would be sufficient to inform other scientists via annotations about new insights. Herewith, Collaborative Document Evaluation would enable the first continuous post-publishing quality evaluation.

Scienstein aims to motivate researchers to participate in Collaborative Document Evaluation with various incentives. It will be an open platform, available to everyone from every application. This way, metadata can be submitted and retrieved, for instance, from PDF-readers, reference managers or browsers. This is certainly more convenient than the current procedure. Moreover, researchers can directly benefit from participating. New software will help them to manage their electronic documents based on their evaluations. For instance, Scienstein could display all documents a researcher has recently read, classified as *peer review* and rated as *good*. Another positive aspect is that researchers could improve their visibility and the visibility of their publications within the scientific community by annotating, rating, and classifying publications of colleagues. Last but not least, ratings of publications can be used by research paper recommender systems to generate recommendations.

Accordingly, the more publications a researcher has rated, the better the recommendations. Also fundamental for the success of Collaborative Document Evaluation is the ability to determine the participants' competence and trustworthiness. These key success factors are covered in more detail in [10].

4 Summary

In this paper we presented Collaborative Document Evaluation as a supplement or even alternative to citation analysis and classic peer review. We outlined many advantages, such as an improved evaluation of (interdisciplinary) work, a continuous post-publishing quality evaluation of publications and improved search possibilities. As part of the Scienstein project, we are currently implementing the presented concept and are looking forward to the results. Particularly the researchers' motivation to participate and methods to determine their trustworthiness and competence will finally decide to what extent Collaborative Document Evaluation actually will be an effective and efficient method for evaluating scientific publications.

5 References

1. F. Godlee, C. Gale, and C. Martyn, "Effect on the Quality of Peer Review of Blinding Reviewers and Asking Them to Sign Their Reports", *JAMA*, 1998, pp. 237-240.
2. A.S. Relman, "Peer Review in Scientific Journals - What Good Is It?", *New England Journal of Medicine*, vol. 153, 1990, pp. 520-522.
3. D. Lee, K. Jaewoo, M. Prasenjit, L. Giles, and O. Byung-Won, "Are your citations clean?", *Communications of the ACM*, vol. 50, 2007, pp. 33-38.
4. M.H. MacRoberts and B. MacRoberts, "Problems of Citation Analysis," *Scientometrics*, vol. 36, 1996, pp. 435-444.
5. L. Yates, "Is Impact a Measure of Quality?", *European Educational Research Journal*, vol. 4, 2005, pp. 391-403.
6. C.A. Kochan and J.M. Budd, "The persistence of fraud in the literature: the Darsee case", *JASIS*, vol. 43, 1992, pp. 488-493.
7. "Nature's peer review trial", *Nature*, 2006, http://www.nature.com/nature/peerreview/ debate/nature05535.html.
8. P. Ball, "The more, the wikier", *Nature*, 2007; http://www.nature.com/news/2007/ 070226/full/news070226-6.html.
9. R. Nyblod and J. Byrne, "USPTO Extends and Expands Peer Review Pilot", July 2008; http://www.uspto.gov/web/offices/com/speeches/08-26.htm.
10. J. Beel and B. Gipp, „Collaborative Document Evaluation: An Alternative Approach to Classic Peer Review", in Proceedings of World Academy of Science, Engineering and Technology, vol. 31, pp. 410–413 (2008) ISSN 1307-6884.
11. B. Gipp, J. Beel and C. Hentschel, "Scienstein - A Research Paper Recommender System", in Proceedings of IEEE International Conference on Emerging Trends in Computing, to be published in January 2009.