Reviewer’s report

Title: Peer2ref: a peer-reviewer finding web tool that uses author disambiguation

Version: 5 Date: 9 August 2012

Reviewer: Sean O'Donoghue

Reviewer’s report:

The manuscript has been greatly strengthen by the included benchmark, which now gives a much clearer picture of the value and accuracy of this method, both in comparison to other similar tools, as well as in comparison to the obvious alternative of simply using the cited references of any paper as a source of potential referees. However, the results of this benchmark should be more clearly described.

Major Compulsory Revisions

None.

Minor Essential Revisions

(1) The benchmark results are critical and therefore need to be given more attention in the main body of the manuscript. A short table summarizing the benchmark results should be included in the main text. It should list for each tool the number of ‘top authors’ that are also cited.

(2) In the Results, where the benchmark results are described, the authors currently state: "The top suggestion was identified as a cited author also in 15 cases both for eTblast and Jane." - the wording here is very confusing - I believe they mean to say '... in exactly 15 cases for each tool (peer2ref, eTblast and Jane).'.

(3) "The results showed that the three tools perform similarly in this task." - this sentence seems to be misleading. The results in fact show that eTblast and Jane (both 15/40) both perform better than peer2ref (12/40) when it is run with no options. To obtain the same performance (15/40), the authors needed to manually enter specific topics. This obviously goes against the whole premise of peer2ref, which is about a fully automated method of finding potential referees. Also, it is to some extent an unfair comparison, as both eTblast and Jane may have also manual options that could be used to improve their performance.

(4) In the Discussion: "We think that this indicate that the tools are somewhat complementary and that peer2ref distinguishes itself more than eTBlas and Jane" - I do not see any evidence for the statement that 'peer2ref distinguishes itself more than eTBlas and Jane'. From the benchmark, it appears that peer2ref performs similarly, but not as well, as the other two tools when all are run completely automatically (no options). The authors should either provide
evidence for this statement, or omit it.

(5) In Table 1, the most important columns are the ones labelled 'Ref' - they are described in the legend as listing the 'number of references coauthored by suggested authors'. This is confusing. I suggest they authors list only a 1 or 0, indicating whether the suggested author did or did not occur in the bibliography of a paper. The description needs to modified to make it clearer, perhaps 'indicates whether the suggested author did or did not occur in the bibliography of each paper'.

(6) "the recall for all tools is lower that one might expect" - lower than expected based on what? This needs to be explained further.

(7) "In principle, we do not have reasons to believe that the proposed reviewers are worse in all these cases," - this is not such a large set - it would be feasible to have a small group (2-3 people) make a subjective evaluation of whether the suggested 'top authors' would indeed make good referees.

(8) "Indeed, this could be indicative that using the references as a source of reviewers may not be always an optimal strategy" - maybe, however I believe the most straight-forward interpretation of these results is that all three methods are not very accurate, as only 15/40 = 37.5% of suggested 'top authors' have been 'proven' to be good referees. The authors should at least acknowledge this possibility, or alternatively disprove it, for example as suggested in the point above.

Discretionary Revisions

* "in 15 cases the fist suggestion was a cited author" - change to 'the first suggestion'

* First sentence in the abstract: "due to the progressive growth of the body of knowledge that carries specialization onto narrower areas of research" - the meaning is clear, but the phrasing is rather odd.

* Introduction "For example, BioMed Central editorial uses this approach by proposing to the associate editor that would handle the submission the cosine values between the abstract of the submitted manuscript and abstracts referenced in MEDLINE." - this sentence seems to have both language and content problems, with the result that is quite unclear what is being asserted here.

* " but the system is not fool-proof (personal observation)" - to say that any given system is not fool-proof is almost a tautology - I would suggest dropping this phrase.

* "journal subject class (i.e. journals in the area of cardiology)" - should be e.g., not i.e.

* "The complete list of journals is available in the tool's website." - change to '... on the Peer2ref website'
Accordingly we selected 40 of the latest (at that date) published papers in PubMed Central, which makes full text available."

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests