Why is wrong to evaluate researchers in CS through journals and what can we do about it?
Microsoft Academic Search might help

Ricardo Jimenez-Peris,
Marta Patiño-Martinez
Universidad Politecnica de Madrid
Current Practices for Evaluating CS Research

• Many funding agencies and public organizations rely on Journal Citation Report from ISI-Thomson (JCR).

• Some organizations are using Scopus that also indexes conferences.

• Scopus is however expensive and many organizations cannot pay for it (e.g. our university).
A New Alternative: Microsoft Academic Search

- [http://academic.research.microsoft.com/](http://academic.research.microsoft.microsoft.com/)
- It indexes over 3 million CS papers and almost 7 million papers including other sciences.
- It provides ranking for conferences, journals, authors, publications and organizations.
- It also provides rankings per categories within CS.
- Time windows: global, last 10 years and last 5 years.
Microsoft Academic Search
What is especially good about Microsoft Academic Search?

• The main strength of Microsoft Academic Search is the conference and journal rankings based on the indexing of over 3 million CS papers (citeseer-X only indexes 1.6 million CS papers).

• They provide raw data, publications and citations what enables to compute the impact.
What needs to be improved Microsoft Academic Search?

• Author citation lists still have a heterogeneous coverage.
• For instance, for the authors of this presentation it covers slightly below 50% of the tracked citations by the authors.
• It also misses book citations that for some authors represent a large fraction of their citations.
• For tracking the citations of a single author Google Scholar is still the best source combined with the Publish-or-Perish tool.
• CIDS: Web-based citation analysis discerning self-citations from University of Lisbon (http://cids.di.fc.ul.pt) removes self citations from Google Scholar queries what enables to compare citations among researchers without manual intervention.
• Currently, it allows to upload non-indexed papers and correct paper data what will allow to improve significantly its quality that is currently quite good.
How to Use Microsoft Academic Search?

- The current data from the conference and journal rankings can be used to build impact rankings.
- Some filtering is needed:
  - ACM SIG Newsletters appear under journal incorporating citation data from the conferences they organize.
    - This can be fixed simply removing them from the journal ranking.
  - Some venues with a low number of publications are sometimes wrongly ranked.
    - A threshold for a minimal number of publications, say 200, might be used, filtering those not reaching it.
Looking at Microsoft Academic Search Data

• It is well known that using only journals is not the right way to evaluate CS research.
• However, it is assumed that JCR is a good source for ranking journals.
• Looking at the citation data in Microsoft Academic Search the conclusion is very different.
Comparing JCR and Microsoft Academic Search (MAS)

- We have looked at the ranking of traditional ACM Transactions in both JCR and MAS: TOCS, TODS, TOIS, TOPLAS, TOSEM, JACM, TOCHI, and TOG.
- JCR ranks them in 2009 in the following positions: 51st, 198th, 132nd, 281st, 92nd, 42nd, 216th, 12th, that is, only one is ranked among the top 25 journals in CS (TOPLAS is ranked 281st!!).
- Whilst MAS ranks them as follows: 1st, 6th, 7th, 8th, 13th, 14th, 18th, and 23rd. That is, all of them are ranked among the top 25 as most researchers consider them.
Comparing JCR and Microsoft Academic Search (MAS)

• Why is JCR providing a so different ranking?
• It simply dismisses all conference citations and most journal citations due to they refer to papers published two years before.
• Taking into account that many journals have a publication time of 2 years or more, this results in discarding most citations...
• The new 5 year impact factor in JCR might improve the situation.
• Publications like proceedings from WSEAS appear now indexed in JCR...
Quality in Citations Analysis vs. Quantity Citation Analysis

- Citation analysis enable to measure the impact of the research.
- However, the bulk number of citations is not good enough since it can be biased.
- This can be fixed by looking at the venues at which a researcher is cited and by whom.
- A good quality factor is to look at the highest impact venues a researcher work is cited in MAS.
- Another one is to identify which of 1000 top-cited researchers at MAS cite the researcher work.
How can we improve the situation?

• In the short and medium term we can lobby the funding agencies to take into account conferences and Microsoft Academic Search as ranking for conferences and journals.
• Informatics Europe might provide conference and journal rankings based on MAS that can be used as a reference for European funding agencies.
• Australia has performed a similar effort with CORE ranking but unfortunately they are done by human perception as opposed to an objective method.
• However, in the longer term the way to avoid a continuous struggle is to lobby ACM, IEEE CS, USENIX and other major CS organizations to index the conference proceedings in JCR.
Conclusions

• Microsoft Academic Search provides a good quality updated ranking of conferences and journals.
• In combination with other tools such as Google Scholar, Publish-or-Perish and CIDS provides a free alternative to evaluate CS Research.
• A ranking provided by Informatics Europe based on MAS can help to establish this as main CS ranking in Europe.
• Quality citation analysis should be promoted.
• Lobbying ACM, IEEE CS, USENIX and other major CS research organizations to index their proceedings in JCR will help to improve the situation in the longer term.