

A Critical Look at the ResearchGate Score as a Measure of Scientific Reputation

Peter Kraker
Know-Center
Inffeldgasse 13/VI
8010 Graz, Austria
pkraker@know-center.at

Elisabeth Lex
Graz University of Technology
Inffeldgasse 13/V
8010 Graz, Austria
elisabeth.lex@tugraz.at

ABSTRACT

In this paper, we present an assessment of the ResearchGate score as a measure of a researcher's scientific reputation. This assessment is based on well-established bibliometric guidelines for research metrics. In our evaluation, we find that the ResearchGate Score has three serious shortcomings: (1) the score is intransparent and irreproducible, (2) the score incorporates the journal impact factor to evaluate individual researchers, and (3) changes in the score cannot be reconstructed. Therefore, we conclude that the ResearchGate Score should not be considered in the evaluation of academics in its current form.

Categories and Subject Descriptors

D.2.8 [Software Engineering]: Metrics; H.2.8 [Database Applications]: Scientific databases

General Terms

Measurement

Keywords

Science 2.0, bibliometrics, ResearchGate, evaluation, composite indicators, Journal Impact Factor, reproducibility

1. INTRODUCTION

ResearchGate is an academic social network that revolves around research papers, a question and answering system, and a job board. Researchers are able to create a profile that showcases their publication record and their academic expertise. Other users are then able to follow these profiles and are notified of any updates. Launched in 2008, ResearchGate was one of the earlier Science 2.0 platforms on the Web. In recent years, ResearchGate has become more aggressive in marketing its platform via e-mail. In default settings, ResearchGate sends between 4 and 10 e-mails per week, depending on the activity in your network. This high number of messages proves to be very successful: according to a recent study by Nature [8], ResearchGate is the most

well known social network among researchers; 35% of surveyed researchers say that they signed up for ResearchGate “because they received an e-mail”.

One of the focal points in their e-mails is a researcher's latest ResearchGate (RG) Score. Updated weekly, the ResearchGate Score is a single number that is attached to a researcher's profile. The RG Score claims to be “a metric to measure your scientific reputation”; it was designed to “help you measure and leverage your standing within the scientific community”. According to ResearchGate, the RG Score includes the research outcomes that you share on the platform, your interactions with other members, and the reputation of your peers (i.e., it takes into consideration publications, questions, answers, followers). The ResearchGate Score is updated every Thursday and communicated along with usage stats in a weekly e-mail. In addition, the score is displayed very prominently on every profile, alongside the photo, name and basic information of a researcher (see Figure 1).

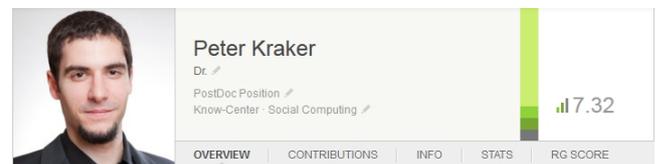


Figure 1: The ResearchGate Score (right) is displayed prominently in a researcher's profile.

In this paper, we take a critical look at the ResearchGate score as a measure of scientific reputation of a researcher, based on well-established bibliometric guidelines for research metrics.

2. EVALUATION

In our evaluation, we found that the ResearchGate Score has serious shortcomings. Following, we give three main reasons for this assessment.

2.1 Reason 1: The score is intransparent and irreproducible

The ResearchGate Score is a composite indicator. According to the ResearchGate, the ResearchGate score incorporates the research outcomes that you share on the platform, your interactions with other members, and the reputation of your



peers. The exact measures being used as well as the algorithm for calculating the score are, however, unknown¹. The only indication that ResearchGate gives its users is a breakdown of the individual parts of the score (see Figure 2), i.e., publications, questions, answers, followers (also shown as a pie-chart), and to what extent these parts contribute to your score.

Unfortunately, there is not enough information given in the breakdown to reproduce one's own score. There is an emerging consensus in the bibliometrics community that transparency and openness are important features of any metric [3]. One of the principles of the recently published Leiden Manifesto for research metrics [4] states for example: "Keep data collection and analytical processes open, transparent and simple", and it continues: "Recent commercial entrants should be held to the same standards; no one should accept a black-box evaluation machine."

Openness and transparency is the only way scores can be put into context and the only way biases - which are inherent of all socially created metrics - can be uncovered. Furthermore, intransparency makes it very hard for outsiders to detect gaming of the system. For example, in ResearchGate, contributions of others (i.e., questions and answers) can be anonymously downvoted. Anonymous downvoting has been criticised in the past as it often happens without explanation [2]. Therefore, online networks such as Reddit have started to moderate downvotes [1].

2.2 Reason 2: The score incorporates the JIF to evaluate individual researchers

When a researcher adds a paper to his or her profile that was published in a journal that is included in the Journal Citation Reports of Thomson Reuters, and thus receives a Journal Impact Factor (JIF), the ResearchGate score increases by the impact factor of this particular journal². This reasoning is flawed, however. The JIF was introduced as a measure to guide libraries' purchasing decisions. Over the years, it has also been used for evaluating individual researchers. There are, however, many good reasons why this is a bad practice. For one, the distribution of citations within a journal is highly skewed; one study found that articles in the most cited half of articles in a journal were cited 10 times more often than articles in the least cited half [7]. As the JIF is based on the mean number of citations, a single paper with a high number of citations can therefore considerably skew the metric [6].

In addition, the correlation between JIF and individual citations to articles has been steadily decreasing since the 1990s [5]. Furthermore, the JIF is only available for journals; therefore it cannot be used to evaluate fields that favor other forms of communication, such as computer sci-

¹A fact that has already been criticised by researcher in the social Web, see e.g. <https://francescientist.wordpress.com/2014/04/20/researchgate-impact-on-the-ego/> and https://www.researchgate.net/post/ResearchGate_RG_SCORE_inclusion_in_the_personal_Curriculum_Vitae_CV

²This seems to be only true for the first article of a journal that is being added. For reasons outlined in 2.1, however, we cannot verify this.

ence (conference papers) or the humanities (books). But even in disciplines that communicate in journals, there is a high variation in the average number of citations which is not accounted for in the JIF. As a result, the JIF is rather problematic even when evaluating journals; when it comes to single contributions it is even more questionable.

There is a wide consensus among researchers on this issue: the San Francisco Declaration of Research Assessment (DORA) that discourages the use of the JIF for the assessment of individual researchers has garnered more than 12,300 signees at the time of writing.

The problem with the way ResearchGate incorporates the JIF runs even deeper. To understand this problem, one needs to know that the Journal Impact Factor for journal x in year y is calculated as the average number of citations an article in journal x from the years $y-1$ and $y-2$ received in year y . Therefore, in order to have any connection between the JIF and an individual paper at all, one needs to look at the JIF of the accompanying journal in the two years after publication. In case of a paper published in 2012, one needs to consider the JIF either from the years of 2013 or 2014.

In order to see whether this was true for ResearchGate, we checked all paper instances that were reported by the RG search interface from both the "Journal of Informetrics" and "Journal of the Association for Information Science and Technology (JASIST)"³, two of the higher ranked journals in the field of library and information science. From the results of this search it seems that each journal has a fixed impact factor that is assigned to the paper regardless of when it was published. In both instances, this fixed number corresponded to the highest impact factor for the journal since 2010⁴.

Therefore, ResearchGate does not only follow the very problematic practice of assigning the Journal Impact Factor to individual papers - and also to individual scientists; unless the paper was published in the two years before, the assigned JIF has nothing to do with the paper being assessed, as the impact factor being assigned does not incorporate the citations to this particular paper.

2.3 Reason 3: Changes in the ResearchGate Score cannot be reconstructed

The way the ResearchGate score is calculated is changing over time. That is not necessarily a bad thing. The Leiden Manifesto [4] states that metrics should be regularly scrutinized and updated, if needed. Also, ResearchGate does not hide the fact that it modifies its algorithm and the data sources being considered along the way. The problem with the way that ResearchGate handles this process is that it is not transparent and that there is no way to reconstruct it. This makes it impossible to compare the RG scores over time, further limiting its usefulness.

As an example, we have plotted the ResearchGate score of

³Formerly known as "Journal of the American Society for Information Science and Technology"

⁴In case of "Journal of Informetrics", it was the highest impact factor since its inception, which was reached in 2011

7.32
RG Score

A new way to measure scientific reputation.

The RG Score takes all your research and turns it into a source of reputation.



PERCENTILE

Your score is higher than 32.5% of ResearchGate members'.

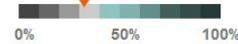


Figure 2: The ResearchGate Score breakdown (Source: ResearchGate profile of one of the co-authors.)

one of the co-authors below from August 2012 to April 2015 (see Figure 3). Between August 2012, when the score was introduced, and November 2012 his score fell from an initial 4.76 in August 2012 to 0.02. It then gradually increased to 1.03 in December 2012 where it stayed until September 2013. It should be noted that the co-author's behaviour on the platform has been relatively stable over this timeframe. He has not removed pieces of research from the platform or unfollowed other researchers. So what happened during that timeframe? The most plausible explanation is that ResearchGate adjusted the algorithm - but without any hints as to why and how that has happened, it leaves the researcher guessing. In the Leiden Manifesto [4], there is one firm principle against this practice: "Allow those evaluated to verify data and analysis".

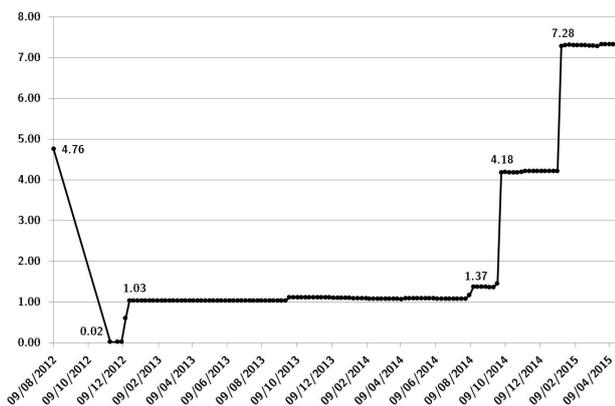


Figure 3: ResearchGate Score of one of the co-authors over time (9 Aug 2012 to 23 Apr 2015)

In comparison to the RG Score, the much criticized Journal Impact Factor fares a lot better. The algorithm is well known and documented, and it hasn't changed since its introduction in 1975. Nevertheless, the raw data is not openly available [5] and the JIF can therefore also not be fully reproduced.

3. CONCLUSIONS

In our evaluation, we found that the ResearchGate Score has serious limitations and should therefore not be considered in the evaluation of academics in its current form. Including research outputs other than papers (e.g. data, slides) is definitely a step into the right direction and the idea of considering interactions when thinking about academic reputation

has some merit. Approaches that take the quality or reputation of a node in a network into account when evaluating the directed edge (e.g., Google's PageRank) have proven to be useful and may also work for ResearchGate (e.g., answers of reputable researchers receive more credit than others). In our evaluation, however, we found several critical issues with the ResearchGate Score, which need to be addressed before it can be seen as a serious metric.

4. ACKNOWLEDGMENTS

The authors would like to thank Isabella Peters for her helpful comments on the manuscript. This work is supported by the EU funded project Learning Layers (Grant Agreement 318209). The Know-Center is funded within the Austrian COMET program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth, and the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

5. REFERENCES

- [1] S. Adali. *Modeling Trust Context in Networks*. Springer, 2013.
- [2] J. Cheng, C. Danescu-Niculescu-Mizil, and J. Leskovec. How community feedback shapes user behavior. *arXiv preprint*, 2014.
- [3] U. Herb. Offenheit und wissenschaftliche Werke: Open Access, Open Review, Open Metrics, Open Science & Open Knowledge. In U. Herb, editor, *Open Initiatives: Offenheit in der digitalen Welt und Wissenschaft*, pages 11–44. Saarland University Press, 2012.
- [4] D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, and I. Rafols. Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548):429–431, 2015.
- [5] G. A. Lozano, V. Larivière, and Y. Gingras. The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the American Society for Information Science and Technology*, 63(11):2140–2145, May 2012.
- [6] M. Rossner, H. Van Epps, and E. Hill. Show me the data. *The Journal of Cell Biology*, 179(6):1091–2, Dec. 2007.
- [7] P. Seglen. Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314(February):498–502, 1997.
- [8] R. Van Noorden. Online collaboration: Scientists and the social network. *Nature*, 512(7513):126–9, Aug. 2014.