Preprint submitted to Scientometrics
Final version published at 10.1007/s11192-023-04787-2

Bibliometric denialism

Daniel Torres-Salinas, Wenceslao Arroyo-Machado and Nicolas Robinson-Garcia
torressalinas@ugr.es; wences@ugr.es; elrobin@ugr.es
0000-0001-8790-3314; 0000-0001-9437-8757; 0000-0002-0585-7359
EC3 research group, Department of Information and Communication Sciences, University of Granada, Spain

Dear *Scientometrics* Editors,

For some time now we are noting a new type of denialism in the bibliometric and research evaluation communities which, although it has always existed, is becoming more and more pronounced. We are referring to bibliometric denialism, a new trend that can be added to climate change denialism, COVID denialism, vaccine denialism, or flat-earth movements. "Denialism" is a concept which refers to the intentional denial of a corpus of empirical evidence (Hansson, 2017). The characteristics of denialism are well established, including a dogmatic rejection of scientific results, the selection of data for convenience and, above all, the formation of impossible expectations about what science can provide (Lucena, 2022). In some cases denialists adopt unfounded positions based on a distorted version of the original message, while in other cases, it is a deliberate rhetoric strategy to persuade others. Currently, there are various opinions that present some of these characteristics, as they categorically deny the usefulness of bibliometric indicators, strongly discouraging their use in scientific evaluation processes.

The evidence for the existence of metric denialism, particularly in the European context, is evident in discussions regarding the obliteration of research metrics in evaluation exercises, favoring the use of narrative CVs or directly banning their use. For instance, in Germany the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) has positioned itself strongly against the use of these indicators, pointing out that "[a] narrow focus in the system of attributing academic reputation – for example based on bibliometric indicators – not only has a detrimental effect on publication behaviour, it also fails to do justice to scholarship in all its diversity"[1]. Among the measures adopted by the DFG to carry out a content-related qualitative evaluation, include the use of a binding CV template and the consideration of new research outputs such as preprints and datasets[2]. Another example is of the Spanish Research Evaluation Agency (AEI), which instructs reviewers of project applications to "consider during the evaluation process that the AEI has signed the DORA Declaration by which it commits to conduct qualitative assessments, without considering bibliometric indexes" (own translation). Furthermore, the CV template used for Spanish calls has now been modified removing the option of including any type of bibliometric indicator, justifying such change by stating that 'it was modified in order to advance on the implementation of the DORA principles"[3]. These

[1] https://www.dfg.de/en/service/press/press_releases/2022/press_release_no_15/index.html
[2] https://www.dfg.de/en/research_funding/announcements_proposals/2022/info_wissenschaft_22_61/index.html
[3] https://www.aei.gob.es/convocatorias/buscador-convocatorias/proyectos-generacion-conocimiento-2021/convocatoria

radical positions are confronted with evidence that shows that both, purely qualitative as and purely quantitative national performance-based research funding systems have their own limitations and can be potentially harmful to the scientific system (Zacharewicz et al., 2019).

Bibliometric denialism has not appeared in a vacuum but responds to a simplistic reinterpretation of DORA[4]. Our understanding is that DORA challenges the abuse and misuse of the Journal Impact Factor in individual evaluation exercises. Its criticism is well-founded and can be extended to any other journal-based indicator applied in individual assessment exercises. But it does not extend to the use of any other type of metric, nor denies the importance of taking informed decisions in evaluation exercises. This misinterpretation along with the positive response DORA had, has led many institutions to subscribe to the declaration. However, we do observe two types of signatures: those of institutions who sign just for image-related reasons and then continue business as usual, and those considering the use of any bibliometric indicator as potentially harmful.

In order to turn DORA into real action, the "Agreement on Reforming Research Assessment" promoted by CoARA[5] was launched. Both, CoARA and DORA, are based on undeniable principles for any bibliometrician, such as, for example, not considering only journal publications, assessing social impact or promoting a more inclusive system. However, they also imply a rejection towards the informative power of metrics and a denial of their usefulness, as the new system is decidedly and unilaterally committed to peer review. In this sense, the European Commission's Directorate-General for Research and Innovation (DG RTD) is categorical, institutions ascribed to the agreement must reform their evaluation system introducing qualitative approaches. Nowhere in the CoARA document are the words bibliometrics, informetrics or scientometrics explicitly mentioned. Perhaps they have deleted them because they understand that “quantitative indicators” is a less offensive term and they seem to consider their use in extreme cases, in a responsible manner and taking multiple precautions. It is preferable that evaluators work blindly, without data, rather than tarnish their judgements with vile indicators.

Two things strike us about this. First, it ignores a complete body of literature already starting in the 1980s in which many of the recommendations, warnings, and precautions on the use of bibliometric indicators was already present in what was coined as Evaluative Bibliometrics (Moed et al., 1985). Here, indicators were already considered support tools rather than an alternatives to peer review processes. Indeed, we agree with Henk F. Moed when indicating that “informetric indicators as a support tool in peer review processes rather than as a replacement of such processes still has a great potential” (Moed, 2017, p. x). CoARA reformulates many of these ideas but misses to mention that they are quite old, as if this body of work was a victim of a new culture of metric cancellation. While it is true that ‘informed peer review’ still needs to be further developed (Wouters, 2014, p.19), it makes little sense to simply ignore its utility instead of working towards its improvement.

The second issue has to do with the romanticism surrounding the notion of peer review and qualitative judgment. One wonders whether CoARA is also missing on the great body

[4] https://sfdora.org/read/
[5] https://coara.eu/agreement/the-agreement-full-text/

of literature dealing with the limitations of peer review and whether they have read the amount of scientific evidence showing the problems with peer review. Furthermore, the proposed alternative, narrative CVs, still has a long way to go until we can understand how they really improve the current assessment system (Bordignon, Chaignon & Egret, 2023). Just as the limits of Bibliometrics are pointed out, the limits of peer review should also be pointed out. Leaving aside the possible problems of subjectivity, discrepancy in the evaluations, impartiality, etc., what concerns us is whether the more practical implications of the systematic application of peer review have been considered. For example, the cost of the system, the growth of bureaucracy, the problem of resolving arbitrariness, the obvious lack of experts in some subjects and the absence, for the moment, of tangible and universal evaluative solutions in the short term. In the era of Big Data and data-informed decisions, it seems that the opposite is being promoted in our sector, precisely when we have more bibliometric methods and indicators at our disposal. When we are most aware of attending to the singularities and social impact of research, the more we seem to crouch in the unilateral trench of the experts. Instead of denying a whole corpus of literature, it might have been more productive to work together towards its improvement.

References

Bordignon, F., Chaignon, L., & Egret, D. (2023). Promoting narrative CVs to improve research evaluation? A review of opinion pieces and experiments. *Research Evaluation*, 1-8. https://doi.org/10.1093/reseval/rvad013

Hansson, S.O. (2017). Science denial as a form of pseudoscience. *Studies in History and Philosophy of Science Part A*, *63*, 39-47. https://doi.org/10.1016/j.shpsa.2017.05.002

Lucena, A.D. (2022, enero 20). *Negacionismo, anticiencia y pseudociencias: ¿en qué se diferencian?* The Conversation. http://theconversation.com/negacionismo-anticiencia-y-pseudociencias-en-que-se-diferencian-174831

Moed, H.F., Burger, W.J.M., Frankfort, J.G., & van Raan, A.F.J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy*, *14*(3), 131-149. https://doi.org/10.1016/0048-7333(85)90012-5

Moed, H.F. (2017). *Applied Evaluative Informetrics*. Springer International Publishing. https://doi.org/10.1007/978-3-319-60522-7

Zacharewicz, T.., Lepory, B., Reale, E., Jonkers, K. (2019). Performance-based research funding in EU Member States – a comparative assessment. *Science and Public Policy, 46*(1), 105-115 https://doi.org/10.1093/scipol/scy041