inorms

Research Evaluation Working Group

Fair and responsible university assessment:
Application to the global university rankings and beyond
Version 5. July 2022

Introduction

The International Network of Research Management Societies (INORMS) established a two-year Research Evaluation Working Group (REWG) in 2018. It consisted of representatives from a range of global member research management societies all seeking to work towards better, fairer and more meaningful research evaluation. One of the group's two areas of focus is the burgeoning influence of the global university rankings on the behaviours of universities despite often poor methodological approaches and practices. The purpose of this work-package was to consider what we, as an international group of research managers and other HE community members, think the characteristics of a fair and responsible university assessment system should look like. The idea was to then 'turn the tables' on the rankings and rate them against our agreed criteria.

An initial set of criteria went out for consultation in May 2019 and received helpful feedback from the community. This is the revised ranking criteria, themed under four key headings: Good governance, transparency, measure what matters, and rigour.

Good governance

- **Engage with the ranked.** Has a clear mechanism for engaging with both the academic faculty at ranked institutions and their senior managers, for example, through an independent international academic advisory board, or other external audit mechanisms. (BP, Bilder et al)
- **Self-improving.** Regularly applies measures of quality assurance to their ranking processes (BP).
- **Declare any conflict of interests.** Provides a declaration of potential conflicts of interest as well as how they actively manage those conflicts.
- **Open to correction.** Data and indicators should be made available in a way that errors and faults can be easily corrected. Any adjustments that are made to the original data and indicators should be clearly indicated. (BP)
- **Deal with gaming.** Has a published statement about what constitutes inappropriate manipulation of data submitted for ranking and what measures will be taken to combat this. (DORA)

Transparency

- **Transparent aims.** States clearly the purpose of the ranking, what it seeks to measure, and their target groups. (BP)
- **Transparent methods**. Publishes full details of their ranking methodology, so that given the data a third party could replicate the results. (CWTS, DORA, BP)
- **Transparent data availability.** Provides detailed descriptions of the data sources being used, inclusion and exclusion parameters, date data snapshots were taken, and so on.
- **Open data.** Makes all data on which the ranking is based available in an open standard non-proprietary format and, where possible, use open standard definitions and classifications (e.g. for subjects, publication types, etc.) to aid interoperability and comparability, and so that those being evaluated can verify the data and analysis. (LM, DORA, BP)
- **Financially transparent.** Publishes details of all sources of income from consultancy services, training, events, advertising, and so on including financial outgoings, e.g. sponsorships.

Measure what matters

- **Drive good behaviour.** Seeks to enhance the role of universities in society by measuring what matters, driving positive systemic effects and proactively seeking to limit any negative impacts such as over-reliance on rankings for decision-making.
- **Measure against mission.** Accepts that different universities have different characteristics – mission, age, size, wealth, subject mix, geographies, etc, and makes visible these differences, so that universities can be clustered and compared fairly. (LM, BP, Blank, Shen)
- **One thing at a time.** Does not combine indicators to create a composite metric thus masking what is actually being measured. (YG1) (CWTS) **LL**
- **No monotonic indicators** for which a good value will depend on the mission of a university.
- **Tailored to different audiences.** The ranking provides different windows onto the data that may be relevant to different audiences. For example, by providing an opportunity to focus in on teaching elements for students.
- **No unfair advantage.** Makes every effort to ensure the approach taken does not discriminate against organisations by size, disciplinary mix, language, wealth, age and geography. (NUF). For example, the use of bibliographic databases that do not have global representation. (LM, Maheu & Lacroix).

Rigour

- **Rigorous methods.** Data collection and analysis methods should pass tests of scientific rigour, including sample size, representation, normalisation, handling of outliers, etc. (BP)
- **No sloppy surveys.** Limit use of unverifiable survey information and ensures that where they are used that the methods are sound and unbiased, e.g. samples are large, representative and randomly selected; questions are reliability-tested and measure what they seek to measure.
- **Defines "University".** When using multiple data sources to take measurements, uses a consistent definition of university across the different data sources. (E.g., universities with multiple campuses - including off-shore, or those with teaching hospitals) (CWTS)
- **Validity.** Indicators have a clear relationship with the characteristic they claim to measure. For example, teaching quality should not solely be indicated by staff-student ratios. (BP, YG2)

- **Sensitivity.** Indicators are sensitive to the nature of the characteristic they claim to measure. (YG1)
- **Honest about uncertainty.** The types of uncertainty inherent in the methodologies used, and of the data being presented should be described, and where possible, clearly indicated using error bars, confidence intervals or other techniques, without giving a false sense of precision. (LM)

References

LM=Leiden Manifesto http://www.leidenmanifesto.org/ (2016)

YG1=Yves Gingras. (2014). *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge, Mass.: MIT Press.

BP=Berlin Principles on Ranking of HE Institutions
https://www.che.de/downloads/Berlin_Principles_IREG_534.pdf (2006)

CWTS=CWTS 10 Principles for Responsible Use of University Rankings
http://www.leidenranking.com/information/responsibleuse (2017)

DORA= Declaration on Research Assessment https://sfdora.org/read/
(2013)

Bilder, Lin & Neylon, Principles for Open Scholarly Infrastructures
https://figshare.com/articles/Principles_for_Open_Scholarly_Infrastructures_v1/1314859 (2015)

Adam, Edmund, New ranking to look at universities' contributions to sustainable development
https://www.universityaffairs.ca/opinion/in-my-opinion/new-ranking-to-look-at-universities-contributions-to-sustainable-development/ (February 2019)

Blank, Kim, University rankings: The Emperor has at least some clothes,
https://www.universityaffairs.ca/opinion/in-my-opinion/university-rankings-emperor-least-clothes/ (November 2016)

YG2 = Gingras, Yves, Academic rankings: The university's new clothes? https://www.universityaffairs.ca/opinion/in-my-opinion/academic-rankings-universitys-new-clothes/ (November 2016)

Shen, Anqi, Consultations are underway for a 'made-in-Canada' Athena SWAN program,
https://www.universityaffairs.ca/news/news-article/consultations-are-underway-for-a-made-in-canada-athena-swan-program/ (October 2018)

Maheu, Louis & Robert Lacroix, The university rankings roller coaster, https://www.universityaffairs.ca/opinion/in-my-opinion/university-rankings-roller-coaster/ (February 2015)

JW = Wilsdon, James. (2019) Deliver us from rankers. https://wonkhe.com/blogs/deliver-us-from-rankers/

INORMS Research Evaluation Working Group

For more information on the INORMS Research Evaluation Working Group please go to:
https://inorms.net/activities/research-evaluation-working-group/