Shifting data sources in the rankings of universities
Chun-Kai (Karl) Huang, Cameron Neylon, Lucy Montgomery, Katie Wilson, Richard Hosking, Alkim Ozaygen, Chloe Brookes-Kenworthy
Curtin University

Key features:
- Large scale university-level comparison on data sources: Web of Science (WoS), Scopus, Microsoft Academic (MSA)
- Compares rankings on open access (OA) and average citation count (CC), calculated from Unpaywall and OpenCitations
- Presence of drastic changes due to shifts in data source
- Mid to lower ranked, non-English language, non-European universities are the most affected.
- This is likely driven by sources’ differences in DOI coverage (e.g., disciplines, output types, language) and metadata records (e.g., publication year).

Methodology:

Result 1:
- 15 universities
- Top 3 remain unchanged by shifting sources
- Ranks in average citation and open access (OA) do not necessarily correlate
- Drastic differences in ranks observed for some universities (mostly non-English speaking universities)

Result 2:
- 155 universities
- Each boxplot depicts changes in ranks when shifting from one source to another
- High central peak
- Long tails (extremes)
- Extreme cases are often mid to lower ranked, non-English speaking universities

Main message:
There is a need to combine and supplement data sources for more robust and fair evaluation metrics and frameworks!

Contact: karl.huang@curtin.edu.au

1 The sample of 15 universities includes: Cairo University, Curtin University, Dalian University of Technology (DUT), Indian Institute of Science Bangalore (IISc), Institut Teknologi Bandung (ITB), Loughborough University (LU), Massachusetts Institute of Technology (MIT), Moscow State University (MSU), National Autonomous University of Mexico (UNAM), University College London (UCL), University of Cape Town (UCT), University of Giessen, University of Sao Paulo (USP), University of Tokyo, Wayne State University (WSU).

This work is part of the Curtin Open Knowledge Initiative. Click here for more details.