GLOBAL UNIVERSITY RANKINGS AND THEIR IMPACT

Andrejs Rauhvargers

http://www.eua.be
Purpose and principles of review

- Addresses the most popular global university rankings
- Providing universities with analysis of the methodologies
- Only publicly accessible information was used
- Efforts were made to discover
  - what is actually measured,
  - how the scores for indicators are calculated
  - how the final scores are calculated, and
  - what the results actually mean.
Global rankings cover not more than 3-5% of world’s universities

Number of universities

Other 16,500 universities

Top 20

Top 500

Next 500

...3...
Decrease of scores within the Top 400 universities

How big can be the scores of remaining for 16’600 universities?
Indicators covering elite research universities only

• “Quality of faculty” = staff winning Nobel prizes (Shanghai-ARWU)
• “Highly Cited” = belonging to worlds Top 200 in 21 areas, i.e. 4200 altogether (ARWU)
• “Peer review” = nominating 30 best universities from pre-selected list (QS-based rankings)
• Teaching reputation survey(s) = nominating 30 best (QS-based, THE-TR)
• Universities considered: pre-selection from elite group of universities: ARWU, THE, Leiden
Indicator scores are usually not the indicator values themselves.

Each indicator has a dimension or denominator, e.g.: articles count, staff numbers, citations per academic unit.

To make indicator scores dimensionless, either:

- Values are expressed as percentage of the result of the “best” university:

\[
Score = \frac{R_x}{R_{\text{best}}} \times 100
\]

- \textit{Z-score} is calculated as being the difference between the measure \(x\) and the mean value \(X\) divided by standard deviation \(\sigma\):

\[
Z = \frac{x - X}{\sigma}
\]
Simple arithmetics have huge influence on scores

- Where a composite score is calculated from several indicators, ranking providers assign weights to each indicator in the overall score.

- This means that the ranking provider’s subjective judgement determines which indicators are more important (e.g. citations – 10%, reputation – 40%)

- If a ranking predominantly uses absolute values (ARWU, Webometrics), its scores are size-dependent, i.e. the ranking favours large universities.
Can rankings assess quality of the research mission of universities? Indicators:

- Publication count SCI & SSCI, Scopus: - production
- Publication count in *Nature* & *Science* - excellence
- Publications per staff - staff research productivity
- Citations (count) – overall force of HEI
- Citations - per paper or per staff - impact
- Citations to articles in the top impact journals – excellence
- Research income
- Research reputation surveys

*But there are also biases and flaws ...*
Rankings and the teaching. Indicators:

- Alumni who have been awarded a Nobel Prize
- Staff/Student ratio
- Reputation surveys (academics, students, employers)
- Teaching income
- Dropout rate
- Time to degree
- PhD/ undergraduate ratio

*All of the above are distant proxies, some questionable*

- Learning outcomes – *are we there yet?*
BIASES AND FLAWS

Natural sciences and medicine vs. social sciences and humanities bias

- Bibliometric indicators primarily cover journal publications and conference proceedings
  - Natural and life scientists primarily publish in journals,
  - Engineering scientists - in conference proceedings,
  - Social scientists and humanists – in books

Several indicators count by 21 broad area
### 21 broad subject areas as defined by ISI

| 1. Agricultural Sciences                        | 12. Mathematics                       |
| 2. Biology & Biochemistry                      | 13. Microbiology                      |
| 3. Chemistry                                   | 14. Molecular Biology & Genetics      |
| 5. Computer Science                            | 16. Pharmacology                      |
| 8. Engineering                                 | 19. Psychology/Psychiatry             |
| 10. Immunology                                 | 21. Space Sciences                    |
| 11. Materials Science                          | ...11...                              |
Different publication and citation cultures in different fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Papers per faculty</th>
<th>Citations per faculty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological Sciences</td>
<td>7.62</td>
<td>59.62</td>
</tr>
<tr>
<td>Physical Sciences and Mathematics</td>
<td>6.39</td>
<td>31.94</td>
</tr>
<tr>
<td>Engineering</td>
<td>6.04</td>
<td>17.83</td>
</tr>
<tr>
<td>Social and Behavioral Sciences</td>
<td>2.14</td>
<td>5.47</td>
</tr>
<tr>
<td>Arts and Humanities</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

*Source: presentation of Cheng at IREG 2010 conference in Berlin*
Field normalisation - solutions and issues

Field-normalised citations per publication indicator (Leiden ‘Crown indicator’)

\[
\frac{CPP}{FCSm} = \frac{C_1 + C_2 + C_3 \ldots}{e_1 + e_2 + e_3 \ldots}
\]

- \(C_i\) is the number of citations of the publication \(i\)
- \(e_i\) is the expected number of citations of publication \(i\) given the field and the year

Criticisms – prefers older publications,
- blurs the picture
Mean-normalisation - solutions and issues

- New attempt (2010) - mean-normalised citation score (MNCS)

\[ MNCS = \frac{1}{P} \times \left( \frac{C_1}{e_1} + \frac{C_2}{e_2} + \frac{C_3}{e_3} + \ldots \right) \]

- Good idea, **but**: now the results are unstable for the very newest publications

- To avoid the new flaw MNCS indicator is used which **leaves out** publications of the last year

- And then it appears that a single publication may substantially change university’s score

- But after all, it just improves mathematics, not the issue that WoS and Scopus insufficiently cover books
‘Peer review’ biases and flaws

- Why calling reputation surveys “Peer reviews”?
- ‘Peers’ are influenced by previous reputation of the institution (including positions in other rankings)
- Limiting the number of universities nominated (THE, QS based rankings) makes approach elitist – and strengthens previous reputation dependence
- Using pre-selected lists rather than allowing ‘peer’s’ free choice results in leaving out huge numbers of institutions
- Is 5-10 % response rate a sufficient result?
- How does opinion survey work when used internationally?
The risks of overdoing

- Rankings encourage universities to improve their scores
- Universities are tempted to improve performance specifically in areas measured in rankings
- There are risks that universities will concentrate funds and efforts to the above aspects and pay less attention to issues that are not rewarded in ranking scores such as:
  - quality of teaching, regional involvement, widening access, lifelong learning, social issues of students and staff etc.
Can rankings be improved?

- There will be no improvement from extending 5 distant proxies to 25 – they will still remain proxies...
- Improve coverage of teaching – most probably through measuring learning outcomes,
- Lift biases, eradicate flaws of bibliometric indicators: field, language, regional,
  but first of all – address non-journal publications properly!
- Change rankings so that they in reality help students to make their choices.
- Addressing elite only, ranking results impact life all universities – it is time to produce rankings that cover more universities!

...17...
The new developments: classifications, multi-indicator tools and comparing learning outcomes
U-map classification profiles (EU)

Neutral process indicators – *not* for value judgments or ranking

- Teaching and learning – levels and orientation of degrees, subject range
- Student profile - mature, distance, part-time,
- Research activity,
- Knowledge exchange,
- International orientation,
- Regional engagement.
The new developments: U-map

- U-Map has two visualisation tools allowing to classify HEIs and to make detailed comparison of selected HEIs.
U-map classification

333 universities have U-map profiles:

- EU – 291
- non-EU European - 16
- Other parts of the world – 26

Goal for 2013: 1000 universities in U-map

Threats:

- Using self-reported data on international level
- Funding
Multi-indicator tool  U-Multirank (EU)

- **Performance indictors**
- Ranking based on one indicator, scores in other indicators displayed
- No overall score calculated
Performance profiles (institutional level)

Source: Multirank final rep, 2011

CPD = continuous profes. development
Multirank: default set of 15 indicators

<table>
<thead>
<tr>
<th>Code of Institution</th>
<th>Teaching &amp; Learning</th>
<th>Research</th>
<th>Knowledge Transfer</th>
<th>International Orientation</th>
<th>Regional Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>student staff ratio</td>
<td>graduation rate</td>
<td>qualification of academic staff</td>
<td>research output</td>
<td>external research</td>
</tr>
<tr>
<td>148</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>293</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>196</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>111</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>222</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>98</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>152</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
«personalizing» the ranking

<table>
<thead>
<tr>
<th>Teaching &amp; learning</th>
<th>Teaching &amp; learning; student satisfaction</th>
<th>Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>student staff ratio</td>
<td>overall judgement of program</td>
<td>external research income</td>
</tr>
<tr>
<td>graduation rate</td>
<td>evaluation of teaching</td>
<td>research publication output</td>
</tr>
<tr>
<td>investments in laboratories</td>
<td>facilities (libraries)</td>
<td>doctorate productivity</td>
</tr>
<tr>
<td>qualification of academic staff</td>
<td>facilities (IT)</td>
<td>field normalised citation rate</td>
</tr>
<tr>
<td>rel. graduate unemployment rate</td>
<td>organisation of program</td>
<td>highly cited research publications</td>
</tr>
<tr>
<td>interdisciplinarity of programs</td>
<td>research orientation of ed. program</td>
<td></td>
</tr>
<tr>
<td>inclusion of employability issues</td>
<td>inclusion of work experience</td>
<td></td>
</tr>
<tr>
<td>inclusion of work experience in program</td>
<td>quality of courses</td>
<td></td>
</tr>
<tr>
<td>computer facilities: internet access</td>
<td>social climate</td>
<td></td>
</tr>
<tr>
<td>student gender balance</td>
<td>support by teachers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>opportunities to stay abroad</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Knowledge transfer</th>
<th>International orientation</th>
<th>Regional engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>ac. staff with non-HE experience</td>
<td>incoming and outgoing exch students</td>
<td>graduates working in the region</td>
</tr>
<tr>
<td>joint research contracts priv sector</td>
<td>international orientation of programs</td>
<td>degree theses with req. enterprise</td>
</tr>
<tr>
<td>university-industry joint publications</td>
<td>international academic staff</td>
<td>regional participation in continuing ed.</td>
</tr>
<tr>
<td></td>
<td>international research grants</td>
<td>summer schools sec.ed.students</td>
</tr>
<tr>
<td></td>
<td>international joint research publ.</td>
<td>student internships in region</td>
</tr>
<tr>
<td></td>
<td>% international students</td>
<td></td>
</tr>
<tr>
<td></td>
<td>internat doctorate graduation rate</td>
<td></td>
</tr>
</tbody>
</table>

Source: Multirank presentation, 2011
Full use of combined U-Map + Multirank

1) using U-map (neutral indicators) select a group of HEIs

2) compare the selected group of HEIs by Multirank using personalized set of performance indicators
This is the result of the personalized comparison of pre-selected group of HEIs

<table>
<thead>
<tr>
<th>code of institution</th>
<th>% income third party funding</th>
<th>Incentives for KT</th>
<th>CPD courses offered</th>
<th>Res contract with regional firms</th>
<th>% Income from regional sources</th>
<th>Student internships in region</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>152</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>293</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>98</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>222</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>196</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>148</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>111</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Source: Multirank presentation, 2011
Name of institution: 293

Address

URL

Mission statement

U-Map profile

U-Multirank profile

Source: Multirank presentation, 2011
Concerns/threats

- There are still no sufficient performance indicators for teaching and learning
- How well self-reported will work in international context
- How well student satisfaction data will work in international context,
- Whether other parties will turn Multirank into a league table and what will be the consequences
The new developments: AHELO

- OECD’s AHELO project is an attempt to compare HEIs internationally on the basis of actual learning outcomes.
- Three testing instruments will be developed within AHELO: one for measuring generic skills and two for testing discipline-specific skills, in economics and engineering.
- Questions yet to be answered are: whether it is possible to develop instruments to capture learning outcomes that are perceived as valid in diverse national and institutional contexts.
New visualisations of global rankings incl. classifications and multi-indicator tools

- ARWU «Rankings Lab»: possibility to chose indicators and asign various weights
- ARWU GRUP Global Research University Profiles self-submitted data collection, 231 universities
- ARWU «Ranking by indicator(22)»: resembles Multirank
- ARWU- field
- Times Higher Education - subject rankings
- Times Higher Education - «Lab» tool
New visualisations of global rankings incl. classifications and multi-indicator tools

- Thomson-Reuters The Profiles Project (reputation, funding, faculty characteristics)
- QS subject rankings – 33 of 52 subjects ranked already
- QS Classifications (size, subject range, research intensity, age of university)
- QS Ranking by indicator («multirank»)
- QS «stars»: (8 criteria)
- QS Country Guides
Main conclusions

1. Since arrival of global rankings then universities cannot avoid national and international comparisons, and this has caused changes in the way universities function.

2. Rankings so far cover only some of university missions. Lack of suitable indicators is most apparent when measuring teaching performance. The situation is better when evaluating research, but even the bibliometric indicators have their biases and flaws.

3. At present, it would be difficult to argue that the benefits that rankings provide are greater than the negative effects of the so-called ‘unwanted consequences’ of rankings.

4. Higher education policy decisions should not be based solely on rankings data.
The elephant is here to remind that information we get from ranking indicator may be correct but it is still partial.

Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong!

by John Godfrey Saxe (1816–1887)
Thanks for your attention