The CWTS Leiden Ranking and other advanced bibliometric tools

Paul Wouters (with Nees Jan van Eck, Ed Noyons and Mark Neijssen)

Road to Academic Excellence – Russian universities in rankings by subject
Mocow, 9-10 April 2015
Centre for Science and Technology Studies (CWTS)

• Research center at Leiden University focusing on quantitative studies of science (bibliometrics and scientometrics)

• Bibliometric contract research
  – Monitoring & evaluation
  – Advanced analytics
  – Training & education
International research front
## CWTS Leiden Ranking 2014

**Select field and region/country**

- **Field:** All sciences
- **Region:** All regions
- **Country:** All countries

**Select indicators**

- **Type of indicators:** Impact
- **Indicator used for ranking:** PP(top 10%)
- **Minimum number of publications:**

**Advanced parameters**

- Calculate size-independent indicators
- Calculate impact indicators using fractional counting
- Show stability intervals

### Rankings

<table>
<thead>
<tr>
<th>Rank</th>
<th>University</th>
<th>Country</th>
<th>P</th>
<th>PP(top 10%)</th>
<th>Stability interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rockefeller Univ</td>
<td></td>
<td>1033</td>
<td>29.1%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>MIT</td>
<td></td>
<td>9149</td>
<td>25.2%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Harvard Univ</td>
<td></td>
<td>29693</td>
<td>23.0%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Univ Calif - Berkeley</td>
<td></td>
<td>11384</td>
<td>22.5%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Stanford Univ</td>
<td></td>
<td>13399</td>
<td>22.3%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Caltech</td>
<td></td>
<td>5072</td>
<td>22.2%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Princeton Univ</td>
<td></td>
<td>5017</td>
<td>21.9%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Univ Calif - Santa Barbara</td>
<td></td>
<td>4246</td>
<td>21.2%</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Univ Calif - San Francisco</td>
<td></td>
<td>9990</td>
<td>20.2%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Yale Univ</td>
<td></td>
<td>9775</td>
<td>20.0%</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Rice Univ</td>
<td></td>
<td>2324</td>
<td>19.2%</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Univ Calif - Santa Cruz</td>
<td></td>
<td>1945</td>
<td>18.9%</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Northwestern Univ</td>
<td></td>
<td>9306</td>
<td>18.8%</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Univ Calif - San Diego</td>
<td></td>
<td>11300</td>
<td>18.7%</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Univ Colorado - Boulder</td>
<td></td>
<td>4833</td>
<td>18.6%</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Univ Texas - Southwestern Med Ctr</td>
<td></td>
<td>4059</td>
<td>18.4%</td>
<td></td>
</tr>
</tbody>
</table>
CWTS Leiden Ranking 2014 team

The following individuals have contributed to the CWTS Leiden Ranking 2014.

**Data collection**
Sophie Boisvert-Hearn
Clara Calero-Medina (coordinator)
Martijn Visser

**Impact indicators**
Nees Jan van Eck
Ludo Waltman (coordinator)

**University-industry collaboration indicators**
Robert Tijsen (coordinator)
Alfredo Yegros

**Geographical collaboration distance indicators**
Robert Tijsen
Nees Jan van Eck
Ludo Waltman

**Website**
Henri de Winter
Nees Jan van Eck (coordinator)
Ludo Waltman
CWTS Leiden Ranking

- Provides bibliometric indicators of:
  - Scientific impact
  - Scientific collaboration
- Calculated based on Web of Science data
- Includes the 750 largest universities worldwide in terms of Web of Science publication output
Differences with other university rankings

• No aggregation of different dimensions of university performance (research, teaching, etc.) into a single overall performance indicator

• Exclusive focus on measuring universities’ scientific performance

• No dependence on survey data or data provided by universities

• Advanced bibliometric methodology
Advanced bibliometric methodology

• Percentile-based indicators to properly deal with highly skewed citation distributions

• Exclusion of publications not aimed at international research front

• Normalization for differences between fields in citation and collaboration practices:
  – Field definitions based on an algorithmically constructed publication-level classification system of science
  – Fractional counting of co-authored publications
Average-based vs. percentile-based indicators
Average-based vs. percentile-based indicators
Exclusion of publications

The following publications are excluded:

• Non-English publications
• Publications in national scientific journals, trade journals, and popular magazines
• Publications in fields with a low citation density
• Retracted publications
Exclusion of publications and the effect on PP(top 10%)
Publication-level classification system of science

- Fields are defined at the level of individual publications rather than journals
- Publications are clustered into ~800 fields based on citation relations
- Smart local moving community detection algorithm
Full vs. fractional counting

Full counting is biased in favor of fields with a lot of collaboration and a strong citation advantage for collaborative publications.
### U- Multirank

#### University Comparison

**Your selection: 854 universities**

<table>
<thead>
<tr>
<th>Teaching &amp; Learning</th>
<th>Research</th>
<th>Knowledge Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bachelor graduation rate</td>
<td>Masters graduation rate</td>
<td>Graduating on time (bachelors)</td>
</tr>
<tr>
<td>U California Santa Cruz</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>U Siegen</td>
<td>DE</td>
<td>--</td>
</tr>
<tr>
<td>MIT</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>Princeton U</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>UAS Wiener Neustadt</td>
<td>AT</td>
<td>--</td>
</tr>
<tr>
<td>U California Berkeley</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>Stanford U</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>U Chicago</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>Caltech</td>
<td>US</td>
<td>--</td>
</tr>
<tr>
<td>Harvard U</td>
<td>US</td>
<td>--</td>
</tr>
</tbody>
</table>

**For students**  |  **Compare**  |  **At a glance**  |  **Readymade**
Some conclusions

• There is no such thing as ‘overall university performance’; do not mix up different dimensions of university performance

• Use an appropriate bibliometric methodology:
  – Use percentile-based indicators, not average-based indicators
  – Do not blindly follow the selection of journals made by database producers; exclude non-scientific journals and journals with a national focus
  – Use fractional counting, not full counting
National research missions
Research leaders face key questions

• How should we monitor our research?
• How can we profile ourselves to attract the right students and staff?
• How should we divide funds?
• What is our scientific and societal impact?
• What is actually our area of expertise?
• How is our research trans-disciplinary connected?
Research leaders need more, not less, strategic intelligence

• Increasing demand for information about research:
  – hyper competition for funding
  – globalization
  – industry – academic partnerships
  – interdisciplinary research challenges
  – institutional demands on research & university management

• Increased supply of data about research:
  – web based research
  – deluge of data producing machines and sensors
  – increased social scale of research: international teams
  – large scale databases of publications, data, and applications
New trends in assessment

• Increased bibliometric services at university level available through databases
• Increased self-assessment via “gratis bibliometrics” on the web (h-index; publish or perish; etc.)
• Emergence of altmetrics
• Increased demand for bibliometrics at the level of the individual researcher
• Societal impact measurements required
• Career advice – where to publish?
Two examples of comparisons (using new normalization)

- Leiden Ranking of universities
  - University ranking
  - Only dimension is RESEARCH
  - Limited set of 750 (most productive) universities world-wide
  - Yearly update

- Brazilian Research Ranking
  - Including all organizations (not only universities)
  - Distribution over 5 regions (N, NE, CO, SE and S)
  - Selection of 110 organization meeting current criteria (P>100)
Exploring the results

• All organizations in Brazil with more than 100 publications in ten years registered in the WoS

• Only dimension is research: impact and collaboration

• Developed at CWTS in collaboration with Brazilian sandwich PhD’s

• http://BRR.CWTS.NL
## CWTS Brazilian Research Ranking

### Ranking

<table>
<thead>
<tr>
<th>Rank</th>
<th>Organization</th>
<th>Type</th>
<th>Region</th>
<th>P</th>
<th>PP(top 10%)</th>
<th>Stability interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Univ Estácio Sá</td>
<td>PrU</td>
<td>SE</td>
<td>113</td>
<td>14.7%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>IMPA - Inst Nacl Matemática Pura &amp; Aplicada</td>
<td>PuRO</td>
<td>SE</td>
<td>425</td>
<td>12.9%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>UESB</td>
<td>PuU</td>
<td>NE</td>
<td>135</td>
<td>10.7%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Inst. Agronômico (IAC)</td>
<td>PuRO</td>
<td>SE</td>
<td>118</td>
<td>7.1%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Embrapa Recursos Genéticos &amp; BioTecnol</td>
<td>PuRO</td>
<td>CO</td>
<td>176</td>
<td>7.1%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>UFABC</td>
<td>PuU</td>
<td>SE</td>
<td>427</td>
<td>6.8%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>PUCPR</td>
<td>PrU</td>
<td>S</td>
<td>441</td>
<td>6.6%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>UFPeL</td>
<td>PuU</td>
<td>S</td>
<td>654</td>
<td>6.5%</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>UNIFRAN</td>
<td>PrU</td>
<td>SE</td>
<td>103</td>
<td>6.3%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>UFSC</td>
<td>PuU</td>
<td>S</td>
<td>3251</td>
<td>6.1%</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Laboratório Nacional de Luz Síncrotron</td>
<td>PuRO</td>
<td>SE</td>
<td>271</td>
<td>6.1%</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>UCS</td>
<td>PrU</td>
<td>S</td>
<td>277</td>
<td>6.0%</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>UNAERP</td>
<td>PrU</td>
<td>SE</td>
<td>128</td>
<td>5.9%</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>EMBRAPA</td>
<td>PuRO</td>
<td>CO</td>
<td>1432</td>
<td>5.7%</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Museu Paraense Emílio Goeldi</td>
<td>PuRO</td>
<td>N</td>
<td>108</td>
<td>5.7%</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>UFAL</td>
<td>PuU</td>
<td>NE</td>
<td>436</td>
<td>5.5%</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>UFCG</td>
<td>PuU</td>
<td>NE</td>
<td>440</td>
<td>5.5%</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>HC/USP</td>
<td>H</td>
<td>SE</td>
<td>181</td>
<td>5.4%</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Petrobras</td>
<td>C</td>
<td>SE</td>
<td>281</td>
<td>5.3%</td>
<td></td>
</tr>
</tbody>
</table>
CWTS Monitor - Meaningful Metrics

• A new interactive way of bibliometric analyses

• Powerful web-based application:
  – User-friendly reporting interface
  – Robust cleaned WoS database run by CWTS
  – Fair and correct benchmarking by state-of-the-art indicators
  – Highly configurable to client’s specific needs

• Professional bibliometric reporting in your hands

• Scientists affiliated to the CTWS Institute of Leiden University provide expert support
CWTS Monitor: Select - Visualise - Conclude
The Leiden Manifesto (forthcoming)

- Quantitative evaluation should support qualitative, expert assessment.
- Measure performance in accordance with the research missions of the institution, group or individual researcher.
- Enable indicators to promote locally-relevant research.
- Keep data collection and analytical processes open, transparent and simple.
- Allow those evaluated to verify data and analysis.
- Account for variation by field in publication and citation practices.
- Base the assessment of individual researchers on qualitative judgment of their portfolio.
- Avoid misplaced concreteness and false precision.
- Recognize systemic effects of assessment and indicators.
- Scrutinize indicators regularly and update them.

In collaboration with: Diana Hicks (Georgia Tech), Ismael Rafols (SPRU/Ingenio), Sarah de Rijcke and Ludo Waltman (CWTS)
The best decisions are taken by combining robust statistics with sensitivity to the mission and nature of the research that is evaluated.

Decision making about science must be based on high quality processes informed by the highest quality data.