Bibliometric visualization using VOSviewer

Ludo Waltman
Centre for Science and Technology Studies (CWTS), Leiden University

Research Output & Impact – New Tools and Concepts
Technical University of Denmark
September 14, 2017
Centre for Science and Technology Studies (CWTS)

- Research center at Leiden University focusing on science and technology studies
- Strong emphasis on bibliometric and scientometric research
- Provider of commercial scientometric services
- History of more than 25 years
- Currently about 30 staff members
Scientometric databases at CWTS

- Web of Science
- Scopus
- PATSTAT
- PubMed
- CrossRef
- ORCID
- Mendeley
- Altmetric.com
- DataCite
- Full-text databases (Elsevier, PubMed Central, Springer, Wiley)
CWTS Leiden Ranking 2017 – Moving beyond just ranking

The CWTS Leiden Ranking 2017 offers key insights into the scientific performance of over 900 major universities worldwide. A sophisticated set of bibliometric indicators provides statistics on the scientific impact of universities and on universities' involvement in scientific collaboration.

Moving beyond just ranking

The CWTS Leiden Ranking aims to move beyond just ranking. The Leiden Ranking therefore can be explored from three different perspectives. In addition to the traditional list view, where universities are presented in a list ordered based on a selected indicator, the Leiden Ranking offers two additional perspectives, the chart view and the map view. The chart view presents universities in a scatter plot, enabling two selected indicators to be compared. The map view presents universities in an interactive geographical map.

Subscribe to the Leiden Ranking newsletter

Responsible use of university rankings

CWTS has developed principles for the responsible use of university rankings.
Outline

• VOSviewer
  – Introduction
  – Visualizing the large-scale structure of science
  – New data sources

• CitNetExplorer

• Principles and challenges for scientometric visualization

Most work presented today was done jointly with CWTS colleague Nees Jan van Eck
VOSviewer

• Any type of (scientometric) network
• Time dimension is of limited importance
• Restricted to small and medium-sized networks

CitNetExplorer

• Only direct citation networks of publications
• Time dimension is of key importance
• Support for large networks
VOSviewer: Introduction
VOSviewer

Authors: moed, hf
Title: measuring contextual citation impact of scientific journals
Source: journal of informetrics, 4(8), 265-277
Year: 2010
Co-authorship map
Term co-occurrence map
Growing use of VOSviewer

![Number of VOSviewer publications per year](chart)
VOSviewer: Visualizing the large-scale structure of science
Identifying the structure of science

• Publications in Web of Science are clustered based on direct citation links (Waltman & Van Eck, 2012):
  – About 20 million publications in the period 2000–2016
  – Over 300 million citation links

• Clustering is performed using the smart local moving algorithm (Waltman & Van Eck, 2013), an improvement of the well-known Louvain algorithm

• The Leiden algorithm, to be released soon, offers further improvements, allowing all publications in Web of Science to be clustered within one hour
Louvain vs. Leiden

Louvain algorithm

Leiden algorithm
4000 micro-level fields of science

Social sciences and humanities
Biomedical and health sciences
Life and earth sciences
Mathematics and computer science
Physical sciences and engineering
Cold vs. hot topics

Climate change

Complex networks

Obesity
Activity of Yonsei University
Relative strengths of Yonsei University
Relative strengths of Korea University
Relative strengths of Yonsei University: Zooming in on social sciences
Focus of Yonsei within scientometrics
Focus of Leiden within scientometrics
CWTS Leiden Ranking

Field-normalized indicators based on 4000 micro-level fields

### Time period, field, and region/country

- **Time period:** 2012-2015
- **Field:** All sciences
- **Region/country:** Denmark
- **Min. publication output:** 100

### Indicators

- **Type of indicators:** Impact
- **Indicators:** P, P(top 10%), PP(top 10%)
- **Order by:** P
- **Calculate impact indicators using fractional counting:**

<table>
<thead>
<tr>
<th>University</th>
<th>P</th>
<th>P(top 10%)</th>
<th>PP(top 10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univ Copenhagen</td>
<td>11090</td>
<td>1432</td>
<td>12.9%</td>
</tr>
<tr>
<td>Aarhus Univ</td>
<td>6476</td>
<td>752</td>
<td>11.8%</td>
</tr>
<tr>
<td>Tech Univ Denmark</td>
<td>5206</td>
<td>701</td>
<td>13.5%</td>
</tr>
<tr>
<td>Univ Southern Denmark</td>
<td>3263</td>
<td>343</td>
<td>10.7%</td>
</tr>
<tr>
<td>Aalborg Univ</td>
<td>2381</td>
<td>251</td>
<td>10.9%</td>
</tr>
</tbody>
</table>
VOSviewer: New data sources
New data sources

- Full-text data
- Crossref (I4OC)
- Microsoft Academic
- oaDOI
- ...
Full-text data: Availability of Elsevier full-text publications
Term co-occurrence map based on meta data (i.e., titles and abstracts)

Term co-occurrence map based on full-text data
Initiative for Open Citations

How many citations are open today?

As of June 2017, the fraction of publications with open references has grown from 1% to more than 45% out of the nearly 35 million articles with references deposited with Crossref (to date).

We encourage all other scholarly publishers to follow the example of these trail-blazing publishers by making their reference metadata publicly available. Please contact Crossref Support (support@crossref.org) for more information, or to let them know that you are ready to open up your reference metadata now. See also our list of responses to frequently asked questions.
Journal co-citation map based on Crossref data for *Scientometrics*
Co-authorship map based on Crossref data for *Journal of Informetrics*
VOSviewer version 1.6.6 to be released within the coming weeks!
CitNetExplorer
CitNetExplorer

The image shows a screenshot of the CitNetExplorer software, which is a tool for visualizing citation networks. The left side of the interface has a selection area with options for current network and selection parameters. The right side displays a citation network diagram with nodes and edges representing publications and their citations. The diagram includes nodes labeled with author names such as van eck, bornmann, and wallman, indicating their connections through citation links.
Standing on the shoulders of giants...
Citation network of scientometrics publications
Literature reviewing using CitNetExplorer
Principles and challenges for scientometric visualization
Principles for scientometric visualization

1. Acknowledge different use cases of visualizations (providing insight vs. attracting attention)

2. Adjust visualizations to the mode of presentation (static vs. interactive)

3. Find an appropriate balance between technical sophistication and methodological transparency
Principles for scientometric visualization

4. Be aware that visualizations tend to give a simplified and incomplete representation of the underlying data

5. Combine visualizations with other pieces of evidence, including your own intuitive judgment

6. Test sensitivity of visualizations to methodological choices; handle these choices pragmatically
Challenges for scientometric visualization

- How to take advantage of new scientometric data sources?
- How to better link interactive visualizations to the underlying scientometric data?
- How to better handle large scientometric data sets?
- How to improve visualization literacy in scientometrics?
Thank you for your attention!