Beyond the impact factor: possibilities of scientometrics to understand science and society

Rodrigo Costas

Centre for Science and Technology Studies (CWTS-Leiden University),
the Netherlands

Centre for Research on Evaluation, Science and Technology (CREST-Stellenbosch University), South Africa

25 August 2020
Outline

• Scientometrics as data science

• Scientometrics to understand science...

• ... and society
  - Altmetrics and social media metrics
Scientometrics as a branch of Data Science

– “Data science is a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data” (Wikipedia)

– Scientometric data sources and applications conform a powerful approach to analyzing: scientific dynamics, trends, activities, interactions, impacts, etc. of multiple scholarly (and non-scholarly) actors through ‘traces’ recorded in scientific publications (and indexed in global bibliographic scientific databases)
Why Scientometrics?

A bibliographic record from any scientometric database is a rich source of **multiple ‘traces’ of information**, all of them with diverse analytical possibilities.

Let’s see an example from the Web of Science...
What do university rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications

Abstract
University rankings by fields are usually based on the research output of universities. However, research managers and rankings consumers expect to see such fields reflect the structure of the institutions. In this study, we address such misinterpretation by developing the research profile of the organizational units of two Spanish universities: University of Granada and Pompeu Fabra University. We use two classification systems: the subject categories offered by Thomson Reuters, which are commonly used in bibliometric studies, and the 33 disciplines displayed by the Spanish FGR Rankings and the UGR Rankings, which are constructed from an aggregation of the former. We also describe the internal processes of universities when working with address data of top research groups and we show differences between universities' structures found from the interdisciplinary organizational forms of new managements. We conclude by highlighting that rankings by fields should clearly state methodology for the construction of such fields. We indicate that the construction of research profiles may be a good solution for universities for finding out levels of discrepancy between organizational units and subject fields.

Keywords
University rankings; Fields; Address data; Institutional structure; Subject classification

Categories / Classification
- Research Areas: Computer Science; Information Science & Library Science
- Web of Science Categories: Computer Science, Interdisciplinary Applications, Information Science & Library Science
Actors (authors, institutions, countries, journals, Funders)
Other factors: time, types of outputs, informal relationships, languages, Open Access
Impact and interactions
Beyond the Journal Impact Factor (or h-index)

Scientometrics

• Science of science
  – Maps of science

• Research management
  – Workforce indicators
  – Mobility of scholars

Altmetrics

• Science communication
  – Social media landscapes

• Societal interactions
  – Heterogeneous couplings
Maps of science

- Hematology
- Cardiovascular system
- Cell biology
- Anesthesiology
- Emergency medicine
- Clinical neurology
- Neurosciences
- Biochemistry & molecular biology
- Toxicology
- Plant sciences
- Biology
- Multidisciplinary sciences
- Chemistry, organic
- Chemistry, multidisciplinary
- Crystallography
- Chemistry, physical
- Oceanography
- Environmental sciences
- Geosciences, multidisciplinary
- Physics, applied
- Physics, interdisciplinary
- Physics, particles & fields
- Engineering, civil
- Engineering, electrical & electronic
- Mathematics, applied
- Mathematics
- Logic
- Transportation
- Ergonomics
- Statistics & probability
- Archaeology
- Education, scientific disciplines
- Health policy & services
- Public, environmental & occupational health
- Psychology, multidisciplinary
- Psychology, educational
- Education & educational research
- Business, finance
- Literature
- History
- Law
- Religion
- Dance
- VOSviewer
Figure 1. Landscape of science (data: Web of Science 2000–2017). Circles represent clusters of publication (areas), size represents relative volume (numbers of publications), color represents main fields, disciplines.
Term maps
HIV research [Dimensions] - VOSviewer
Research management: Workforce analysis

• 2008 onwards: author-affiliation linkage in Web of Science publications
Mobility analysis

2008 onwards: author-affiliation linkage in WoS publications
Conceptualizing mobility

<table>
<thead>
<tr>
<th>NOT MOBILE</th>
<th>MIGRANT</th>
<th>TRAVELER</th>
<th>NON DIRECTIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Year 4</td>
</tr>
<tr>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
<tr>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Year 4</td>
</tr>
<tr>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image21.png" alt="Image" /></td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Year 4</td>
</tr>
<tr>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
<td><img src="image28.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
</tr>
<tr>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Year 4</td>
</tr>
<tr>
<td><img src="image33.png" alt="Image" /></td>
<td><img src="image34.png" alt="Image" /></td>
<td><img src="image35.png" alt="Image" /></td>
<td><img src="image36.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image37.png" alt="Image" /></td>
<td><img src="image38.png" alt="Image" /></td>
<td><img src="image39.png" alt="Image" /></td>
<td><img src="image40.png" alt="Image" /></td>
</tr>
<tr>
<td>Year 1</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Year 4</td>
</tr>
<tr>
<td><img src="image41.png" alt="Image" /></td>
<td><img src="image42.png" alt="Image" /></td>
<td><img src="image43.png" alt="Image" /></td>
<td><img src="image44.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image45.png" alt="Image" /></td>
<td><img src="image46.png" alt="Image" /></td>
<td><img src="image47.png" alt="Image" /></td>
<td><img src="image48.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Mobility analytics – flows, profiles, networks

F) SOUTH AFRICA

<table>
<thead>
<tr>
<th>Emigrating to</th>
<th>Immigrating from</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA 182</td>
<td>GBR 228</td>
</tr>
<tr>
<td>GBR 122</td>
<td>USA 203</td>
</tr>
<tr>
<td>AUS 90</td>
<td>IND 118</td>
</tr>
<tr>
<td>DEU -51</td>
<td>FRA 85</td>
</tr>
<tr>
<td>CAN -45</td>
<td>DEU 79</td>
</tr>
<tr>
<td>NLD -35</td>
<td>NGA 69</td>
</tr>
<tr>
<td>FRA -31</td>
<td>CAN 54</td>
</tr>
<tr>
<td>CHE -27</td>
<td>AUS 53</td>
</tr>
<tr>
<td>NGA -27</td>
<td>NLD 41</td>
</tr>
<tr>
<td>KEN -25</td>
<td>ZWE 40</td>
</tr>
<tr>
<td>NZL -23</td>
<td>ITA 39</td>
</tr>
<tr>
<td>IND -20</td>
<td>CHE 28</td>
</tr>
<tr>
<td>SAU -18</td>
<td>CHN 27</td>
</tr>
<tr>
<td>ZWE -17</td>
<td>ESP 23</td>
</tr>
<tr>
<td>BWA -17</td>
<td>BEL 21</td>
</tr>
</tbody>
</table>
Altmetrics & Social media metrics
What are altmetrics & social media metrics?


- **Difficult to define:**
  - Working definition: *events on social and mainstream media platforms related to scholarly content or scholars [...] and are not the same as [...] citations* (Haustein, Bowman, Costas, 2015)
  - Heterogeneity!

- Main challenge: *what do they mean?*
How does it look like in real life?
Altmetric landscapes:
HIV research [Dimensions] – VOSviewer (tweets)
Towards the conceptualization of ‘heterogeneous couplings’

Bibliographic coupling (co-linking)*

Co-citation (co-linked)*

* Webometric parallels (Bjorneborn & Ingwersen, 2004)

(Boyack & Klavans, 2010)
Heterogeneous couplings on Twitter – tweets

Tweet coupling

Co-tweet linked

Links to

URL1 (paper 1)

Links to

URL2 (paper 2)

Tweet 1

Tweet 2

Links to

URL1 (paper1)

Links to

URL2 (paper2)
Heterogeneous couplings on Twitter – tweeter

Tweeter coupling

@user1

URL1 (paper 1)

tweets

URL2 (paper 2)

tweets

Co-tweeter linked

@user1

tweets

URL1 (paper1)

@user2

tweets

URL2 (paper2)
Tweeter coupling based on ZA-research (Community of attention)

Applications:
- Communities of attention
- Diversity of audiences
- Identification of stakeholders
- Cognitive bridges
Co-tweeter linkage of journals based on ZA-research

Applications:
- New maps of science
- *Societal* clusters of publications/topics
- Cognitive gaps
- Misinformation
The future

Social-media studies of science

Wouters, Zahedi, Costas (2019)
Thank you very much