Scientific culture: Public perception of Science and Technology in Spanish Wikipedia

Ángel F. Zazo Rodríguez, Carlos G. Figuerola,
José L. Alonso Berrocal

Institute of Science and Technology Studies
University of Salamanca
{ angelzazo, figue, berrocal }@usal.es

2014-11-25
Scientific culture: Public perception of Science and Technology in Spanish Wikipedia

- Aspects to consider:
  - Selecting articles on Sci&Tech
  - Thematic fields and relations
  - Editing contents in a collaborative environment
Data (v. 2013-11-01)

Only articles after cleaning

- 1,025,168 articles
- 200K categories
- 30M links
- 49M editions
- 5,7M editors (mostly anonymous)
- 30M article views per day
Selecting articles on Sci & Tech

- 1,025,168 articles
- We can use interlinks between articles to apply SNA techniques
- Community detection: *sets of highly inter-connected nodes*
  - InfoMap algorithm, very fast
  - It produces 1,314 communities, 342 have more than 10 articles
  - They can be evaluated by hand, 23 were assigned to S&T
- Performs additional checks over links to tune articles assignment
  - 119,686 articles on S&T

E-I index: \[
\frac{EL - IL}{EL + IL} = -0.69
\]
Sci & Tech articles

Community detection again

1. Astronomy (7,108)
2. Health Sci (13,259)
3. Inf. Tech. (9,793)
4. Maths (3,008)
5. Paleonthology, Earth Sci. (6,079)
6. Zoology (35,142)
7. Chemistry, Physics, Engineering (12,628)
8. Bothanics (27,553)
Sci & Tech articles

Community detection again

1. Astronomy (7,108)
2. Health Sci (13,259)
3. Inf. Tech. (9,793)
4. Maths (3,008)
5. Paleontholopy, Earth Sci. (6,079)
6. Zoology (35,142)
7. Chemistry, Physics, Engineering (12,628)
8. Bothanics (27,553)
### Sci & Tech articles: links between articles

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Astronomy</strong> (204.971)</td>
<td></td>
<td>91%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td><strong>Health Sci</strong> (352.349)</td>
<td></td>
<td>79%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5%</td>
<td>9%</td>
</tr>
<tr>
<td><strong>Inf. Tech.</strong> (190.555)</td>
<td></td>
<td></td>
<td></td>
<td>92%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td><strong>Maths</strong> (53.116)</td>
<td></td>
<td>5%</td>
<td></td>
<td>84%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td><strong>Paleoarth., Earth Sci.</strong> (230.631)</td>
<td></td>
<td>4%</td>
<td></td>
<td></td>
<td>65%</td>
<td>26%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zoology</strong> (800.308)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>91%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chem. Phys. Eng.</strong> (309.645)</td>
<td></td>
<td>3%</td>
<td>7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82%</td>
<td></td>
</tr>
<tr>
<td><strong>Bothanics</strong> (766.200)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24%</td>
<td>71%</td>
<td></td>
</tr>
<tr>
<td><strong>Military Tech.</strong> (97.654)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7%</td>
<td>87%</td>
</tr>
</tbody>
</table>

\[
\frac{EL - IL}{EL + IL} = -0.63
\]
## Non-S&T → S&T links

<table>
<thead>
<tr>
<th>Non-S&amp;T articles</th>
<th>% links from Non-S&amp;T (3.593.348)</th>
<th>Average indegree from Non-S&amp;T articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astr.</td>
<td>3,2%</td>
<td>16,4</td>
</tr>
<tr>
<td>Health S.</td>
<td>6,9%</td>
<td>18,7</td>
</tr>
<tr>
<td>Inf.Tech</td>
<td>9,0%</td>
<td>33,0</td>
</tr>
<tr>
<td>Maths</td>
<td>1,1%</td>
<td>12,8</td>
</tr>
<tr>
<td>Pal.&amp;Earth</td>
<td>1,3%</td>
<td>7,5</td>
</tr>
<tr>
<td>Zool.</td>
<td>35,9%</td>
<td>36,7</td>
</tr>
<tr>
<td>Chem&amp;</td>
<td>25,7%</td>
<td>73,0</td>
</tr>
<tr>
<td>Bothan.</td>
<td>10,7%</td>
<td>14,0</td>
</tr>
<tr>
<td>Mil.</td>
<td>6,3%</td>
<td>44,0</td>
</tr>
</tbody>
</table>
Editing contents in a collaborative environment

Differences between Non-S&T and S&T articles

- Articles
- Categories
- In-links
- Out-links
- Editors
- Reverts
- Origin of editors
- Page views

<table>
<thead>
<tr>
<th>Editors</th>
<th>#</th>
<th>New Articles</th>
<th>Editions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered (no admin)</td>
<td>467,199</td>
<td>388,743</td>
<td>13,904,570</td>
</tr>
<tr>
<td>Admin (sysop, bureaucrat, rollbacker, checkuser...)</td>
<td>784</td>
<td>492,964</td>
<td>9,505,670</td>
</tr>
<tr>
<td>Bots</td>
<td>366</td>
<td>3,832</td>
<td>10,559,405</td>
</tr>
<tr>
<td>Unregistered</td>
<td>5,217,187</td>
<td>141,629</td>
<td>15,191,681</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>5,685,536</td>
<td>1,027,168</td>
<td>49,161,326</td>
</tr>
</tbody>
</table>
Articles, editors and editions

Number of Non-S&T articles

Number of S&T articles

Articles, editors and editions

Number of Non-S&T articles

Editions of Non-S&T articles

Number of S&T articles

Editions of S&T articles
Categories, size and in\&out links

Frequency distribution of categories

Frequency distribution of article length

Number of categories

Number of articles (%)

Number of articles (%)
Categories, size and in\&out links

Frequency distribution of categories

Frequency distribution of article length

Frequency distribution of out-links

Frequency distribution of in-links

User activity: editors, reverts, origin

Frequency distribution of editors

Non-S&T
S&T

Number of articles (%)

Human editors

0% 5% 10% 15% 20% 25%

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
User activity: editors, reverts, origin

Frequency distribution of editors

Frequency distribution of reverts

User activity: editors, reverts, origin

Frequency distribution of editors

Frequency distribution of reverts

Origin of editions (detected for ~50%)
Article views

Average article views per month (2012-2013)

- Non-S&T
- S&T
Article views

**Average article views per month (2012-2013)**

**Total article views per community per month (2012-2013)**
II Seminar on Indicators of Scientific and Technological Culture

Scientific culture: Public perception of Science and Technology in Spanish Wikipedia

Ángel F. Zazo Rodríguez, Carlos G. Figuerola, José L. Alonso Berrocal

Institute of Science and Technology Studies
University of Salamanca
{angelzazo, figue, berrocal}@usal.es