Are Women Present, Absent or in Disguise? Analyzing Gender Bias in the Spanish WP

J. Minguillón, J. Meneses, S. Fàbregues, E. Aibar, N. Ferran-Ferrer
Universitat Oberta de Catalunya, Barcelona, Spain

Nov. 2018

**English (1st)** | **Spanish**
---|---
Articles | 5.7 M | 1.5 M (8th)
Edits | 864 M | 111 M (4th)
Editors | 35 M | 5.2 M (2nd)
Active | 135 K | 18 K (4th)

**Sample**

October the 1st, 2017 dump

Registered users (no IPs, no bots) that have edited during the last 5 years with at least 50 edits and user page

13,210 candidates to be coded

Stratified random sampling (4 x 4) by number of edits / user page size

<table>
<thead>
<tr>
<th>Not coded</th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>1,618</td>
<td>1,122</td>
<td>172</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td>746</td>
<td>907</td>
<td>0</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td>59</td>
<td>0</td>
<td>123</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2,423</td>
<td>2,029</td>
<td>295</td>
</tr>
</tbody>
</table>

**Discussion**

4,747 coded, few discrepancies manually solved

Krippendorff’s α = 0.9635 (2 coders, 104 r. s.)

Men are more likely to include their gender in profile (59.6% vs 51.4%)

Only 354 (7.5%) can be identified as women, 2,775 (58.4%) men, 34.1% remain unknown

Our projection: **11.3% are women**

EuroCSS 2018 – European Symposium Series on Societal Challenges: Bias and Discrimination
Köln, Germany, 5-7 December 2018