DSpace and the standardization of the information: names of Spanish authors

Fernanda Peset, Antonia Ferrer, Tomas Baiget, José-Manuel Rodríguez-Gairín
Universidad Politécnica de Valencia: mpesetm@upv.es, Universidad Politécnica de Valencia: anfersa@upv.es,
Institut d’Estadística de Catalunya: baiget@idescat.es
Universitat de Barcelona: rodriguez.gairin@ub.edu

DSpace User Group Meeting 2007, FAO, Rome, Italy, October 17th–19th

Introduction:
Spaniards use officially two family names and sometimes also a compound given name. IraLIS is campaigning to make Spanish authors aware of the negative consequences of the traditional signature format.

Definition:
International/standardized form of the name: Only one first name and one family name. Middle initial optional.
Examples for Josep Manel Rodríguez Gairín:
Jose Rodriguez or
Jose-Manuel Rodriguez-Gairin or
JoseManuel RodriguezGairin or
JosepManel RodriguezGairin or
Pepe RodriguezGairin...

Objective of this study:
To identify all the author’s names variations of the main Spanish DSpace repositories.

Materials and method:
– Source:
Wiki Spanish DSpace User Group GUDE

In Spain there are 7 DSpace installations. We analysed 4 of them grouped under RECERCAT, gathering more than 4,400 authors.

<table>
<thead>
<tr>
<th>Institution</th>
<th>Project name</th>
<th>N° of authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consorcio de Biblioteques Universitaries de Catalunya</td>
<td>RecerCat – Diposit de la Recerca a Catalunya</td>
<td>4424</td>
</tr>
<tr>
<td>Universidad Carlos III de Madrid (UC3M)</td>
<td>e-archivo</td>
<td>809</td>
</tr>
<tr>
<td>Universidad de Alcalá (UAH)</td>
<td>e-buah</td>
<td>208</td>
</tr>
<tr>
<td>Universidad de Alcante (UA)</td>
<td>RUA Research</td>
<td>106</td>
</tr>
<tr>
<td>Universidad de Barcelona (UB)</td>
<td>Diposit digital de la UB</td>
<td>194</td>
</tr>
<tr>
<td>Universitat de Girona (UdG)</td>
<td>Diposit de la UdGora</td>
<td>106</td>
</tr>
<tr>
<td>Universitat Politecnica de Catalunya (UPC)</td>
<td>DSpace. E-prints UPC</td>
<td>617</td>
</tr>
</tbody>
</table>

Intended results:
% usage of the “international form” of the names, the consistency between signature forms, the most frequent typing errors...

Main results:
Forms of the signatures of the Spanish authors:

<table>
<thead>
<tr>
<th>Non standardized</th>
<th>Standardized</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1,131</td>
<td>521</td>
</tr>
<tr>
<td></td>
<td>1,652</td>
<td></td>
</tr>
</tbody>
</table>

Regarding authority control:
• The repositories analysed include personal (91%) and corporate authorities (9%) names
• Less than 2% of errors of wrong typing

Regarding signature form:
• 25% of the authors in our Spanish repositories are foreign researchers
• Only 25% of the Spanish authors use an international/standardized form
• Standardization Index: UPC is the university more standardized, and the UPF the less

Conclusions:
It’s needed an awareness campaign stressing the importance of signing in a standard form to improve the retrieval precision

The Spanish science needs also some technical developments that help to correct this situation.

The presentation of this paper in the Conference has been funded by the Programa de Apoyo a la Investigación y Desarrollo (PAID-04-07) of Universidad Politécnica de Valencia.

International Registry for Authors: Links to Identify Scientists
See you in the IraLIS web!
http://www.iralis.org