Research Data Management Literacy at Three Spanish Universities

Alicia Arias-Coello
Clara Simón-Blas
Pablo Arráñz-Val
José Simón-Martín

Universidad Complutense de Madrid (Spain)
Universidad Rey Juan Carlos de Madrid (Spain)
Universidad de Burgos (Spain)
Universidad Complutense de Madrid (Spain)
UCM: Founded in XVI century.
+ 70,000 students approx.
6,000 Univ. professors approx.

URJC: Founded in 1996
+ 30,000 students approx.
1,400 university professors approx.

UBU: Founded in 1994
+ 6,500 students approx.
800 university professors approx.
1 Situation in Spain

2 Methodology

3 Results

4 Discussion and Conclusions
Situation in Spain

Legislation
Instrument

- An on-line survey coordinated by
  Serap Kurbanoğlu
  Joumana Boustany

Our Sample

- Research professors and student research fellows at:
  - UCM
  - URJC
  - UBU

Total Questionnaires

- 3,799
  - UCM 2,740
  - URJC 634
  - UBU 425

Number Responses:

- Received: 828
  - Fully completed: 591
DEMOGRAPHIC RESULTS

Questionnaires
Research academic staff
501 (84.6%)

Questionnaires
Student research fellows
90 (15.2%)
### Demographic Results

1

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>269</td>
<td>45.5</td>
</tr>
<tr>
<td>Male</td>
<td>316</td>
<td>53.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-25</td>
<td>17</td>
<td>2.9</td>
</tr>
<tr>
<td>26-35</td>
<td>85</td>
<td>14.4</td>
</tr>
<tr>
<td>36-45</td>
<td>159</td>
<td>26.9</td>
</tr>
<tr>
<td>46-55</td>
<td>198</td>
<td>33.5</td>
</tr>
<tr>
<td>56-65</td>
<td>113</td>
<td>19.1</td>
</tr>
<tr>
<td>65+</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>No reply</td>
<td>1</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Table 1. Demographic data of respondents**

Profesor’s group

82.5% older than 36 years age and men

Student group

100% less than 30 years old and men
Years involved in research

- > 20 years: 38.2% (Number: 226)
- 16-20 years: 18.6% (Number: 110)
- 11-15 years: 14.4% (Number: 85)
- 5-10 years: 13.7% (Number: 81)
- <5 years: 14% (Number: 83)

- Never involved: 1% (Number: 6)
Table 3. Work discipline

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture Science</td>
<td>13</td>
<td>2.2</td>
</tr>
<tr>
<td>Engineering</td>
<td>42</td>
<td>7.1</td>
</tr>
<tr>
<td>Humanities</td>
<td>71</td>
<td>13.1</td>
</tr>
<tr>
<td>Medical and Health</td>
<td>66</td>
<td>11.1</td>
</tr>
<tr>
<td>Natural Sciences</td>
<td>145</td>
<td>24.5</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>209</td>
<td>35.3</td>
</tr>
<tr>
<td>Other</td>
<td>45</td>
<td>7.6</td>
</tr>
</tbody>
</table>
DATA FILES AND STORING DEVICES

RESULTS
TYPE OF DATA FILES AND STORING DEVICES

- Standard office documents (text,...)
- Internet and web-based data (webpages,...)
- Images (JPEG, GIF, TIFF, PNG, etc.)
- Archived data (ZIP, RAR, ZAR, etc.)
- Structured scientific and statistical data...
- Non digital data (paper, films, slides,...)
- Databases (e.g. in Access, Oracle, MySQL,...)
- Software applications (modelling tools,...)
- Configuration data (parameter settings,...)
- Raw (machine-generated) data
- Audio files
- Source code (scripting, Java, C, C++, etc.)
- Encoded text (XML, SGML, etc.)
- Structured graphics (CAD, CAM, VRML,...)

Fig. 1. Type of data files used and produced in research

Results
DF-SD
1
DATA MANAGEMENT PLAN (DMP) RESULTS
Fig. 2. Opinions about DMP
KNOWLEDGE AND USE METADATA RESULTS
Fig. 3. Opinions about knowledge of metadata
Uses file naming convention or standard
Uses datasets that are tagged with standard metadata
Uses your own/in-house (your research team) tags and metadata
Uses metadata standard for tagging your data

Fig. 4. Use of metadata. AW (Almost Always), OF (Often), ST (Sometimes), RA (Rarely), NE (Never)
DATA LITERACY TRAINING
RESULTS
DATA LITERACY TRAINING

Fig. 5. Formal training received in various areas of data management

Results training
1
DISCUSSION AND CONCLUSIONS
Our Study

Limitations

¿DMP???

Metadata...
Final conclusion

A growing need for data management = The necessary training for better quality and more competitive research.
aarias@ucm.es & jsimon@ucm.es

clara.simon@urjc.es

parranz@ubu.es