1. e-Social Sciences

Extraction of data from the web and their analysis through the use of Social Network Analysis and Semantic Analysis techniques

- Virtual communities
- Open Source Software communities
- Open innovation communities
- Website structure analysis
- Knowledge networks
1. e-Social Sciences

2. Web content

1. Information extraction
2. Keywords selection
3. Structuring and processing information (LSA)
4. Categorization
2. Web content

Oss projects mailing list

Distribution of topics per year

<table>
<thead>
<tr>
<th>Year</th>
<th>Topic 1</th>
<th>Topic 2</th>
<th>Topic 3</th>
<th>Topic 4</th>
<th>Topic 5</th>
<th>Topic 6</th>
<th>Topic 7</th>
<th>Topic 8</th>
<th>Topic 9</th>
<th>Topic 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>data</td>
<td>security</td>
<td>software</td>
<td>web</td>
<td>design</td>
<td>tutorial</td>
<td>training</td>
<td>research</td>
<td>development</td>
<td>maintenance</td>
</tr>
<tr>
<td>2011</td>
<td>cloud</td>
<td>big data</td>
<td>internet</td>
<td>mobile</td>
<td>UI</td>
<td>analytics</td>
<td>e-commerce</td>
<td>social</td>
<td>content</td>
<td>privacy</td>
</tr>
<tr>
<td>2012</td>
<td>AI</td>
<td>IoT</td>
<td>privacy</td>
<td>content</td>
<td>UX</td>
<td>AI</td>
<td>data science</td>
<td>security</td>
<td>e-commerce</td>
<td>privacy</td>
</tr>
<tr>
<td>2013</td>
<td>machine learning</td>
<td>blockchain</td>
<td>data science</td>
<td>AI</td>
<td>IoT</td>
<td>data</td>
<td>cloud</td>
<td>big data</td>
<td>internet</td>
<td>mobile</td>
</tr>
</tbody>
</table>

GCC 4.7 is now the default for x86 architectures
2. Web content

Research trends in ITS
2. Web content
2. Web content

Members:
- Alias – e-mail
- Analysis of the headings

Arcs:
- Valued graphs

Community of ARM Debian Linux (2007)

Vote
Discuss
2. Web content

Data extracted from IdeaStorm

- Idea name
- Author
- Date
- Comments
  - Number of comments
  - Authors who posted these comments
- Promotions
  - Number of received promotions
  - Authors who suggested promotions of the idea
- Demotions
  - Number of received demotions
  - Authors who suggested demotions of the idea

1482 ideas
2. Web content

Comment network
N = 1361; n = 1153 + 208
In-degree = 0 ⇒ 808
In-degree > 1 ⇒ 24
Out-degree = 0 ⇒ 953
Out-degree > 1 ⇒ 406
Out-degree > 10 ⇒ 25

Promotion network
N = 2151; n = 1153 + 998
In-degree = 0 ⇒ 1153
In-degree > 1 ⇒ 341
Out-degree = 0 ⇒ 6
Out-degree > 1 ⇒ 1281
Out-degree > 10 ⇒ 25

Demotion network
N = 1459; n = 1153 + 306
In-degree > 10 ⇒ 22
Out-degree = 795 ⇒ 1

• Aim: patterns of collaboration and internationalization of Universities
• Methodology: Co-authorship analysis + Social Network analysis
• Scope: England Universities

Knowledge Networks

Longitudinal study
2006
2010
3. Web structure

University website structure

<table>
<thead>
<tr>
<th>Total</th>
<th>Average</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subdomains</td>
<td>2438</td>
<td>30.47</td>
</tr>
<tr>
<td>External domains</td>
<td>30500</td>
<td>381.25</td>
</tr>
<tr>
<td>Web pages</td>
<td>718272</td>
<td>8978.40</td>
</tr>
<tr>
<td>Out-links</td>
<td>4429231</td>
<td>55365.38</td>
</tr>
</tbody>
</table>

80 corporate Spanish Universities

www.ub.edu
www.uab.es
www.uvic.es
www.umh.es
www.upc.edu
www.uclm.es

Representation of the University of Seville website

Domain network

Page network
3. Web structure

University website structure

- Factor 1: Distributed structure
  - Information finding requires browsing through several web pages

- Factor 2: Centralized structure
  - Accessibility of information
3. Web structure

University website structure

- Factor 3: Egocentric structure
  - Sum of more or less independent subnetworks

- Factor 4: High sized websites
  - Long navigation process to achieve the information
3. Web structure

Factor 5: Small sized websites
- Information provided through external references

Factor 6: Subnetwork structure
- One subnetwork which contains the most relevant information
DATA EXTRACTION, MODELING AND APPLICATIONS IN E-SOCIAL SCIENCES

M.R. Martínez Torres
University of Seville
Seville, Spain
rmtorres@us.es