Open access repositories as channel of publication scientific grey literature

Tránsito Ferreras-Fernández
Francisco J. García-Peñalvo
José A. Merlo-Vega

Contact
transiff@usal.es / @transi
fgarcia@usal.es / @frangp
merlo@usal.es / @merlovega

GRIAL Research Group, University of Salamanca
“Theses and dissertations are the most useful kinds of invisible scholarship and the most invisible kinds of useful scholarship. Because of their high quality and low visibility, the access problem is worth solving”

(Peter Suber, 2012)
Our Goals

• Reporting the benefits of scientific communication model through open access repositories

• Reporting the benefits for scientific grey literature: deposited and disseminated theses through GREDOS repository

• Presenting the fundamentals, the state of the art, trends and benefits of open access, understood as a radical change in the system of scientific communication
Scientific communication

- Technological development facilitates the communication of scientific knowledge, allowing expand distribution channels and significantly reducing transmission costs of the investigation.

- New paradigms in the process of scientific communication.

- Internet as a powerful tool to disseminating knowledge.

- Open academic and research content Scientific production world is available to Society.
Open Access movement

- The **Open Access movement** was launched to improve access to publicly funded research and help libraries dealing with rampant inflation in journal subscription prices.

- According to **Peter Suber** the Open Access movement is dependent on Internet technologies and the consent of the author or copyright holder.
Institutional Repositories (IRS)

- Institutional repositories are a product of OA movement.

- Repositories capture the scientific research of the faculty, staff, and students of universities or academic disciplines, and assure perpetual and free access to it.

- The most Institutional repositories (IRS) disseminate electronic unpublished theses and dissertations (ETD).
IRS and Grey literature

• So far this lack of dissemination of this unpublished research works made be considered as grey literature, understanding as the literature what is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers.

• The repositories are a tool for the dissemination of research in universities through the dissemination theses and dissertations.
IRS and Theses

• The practice of making digital versions of theses and dissertations available online is growing internationally.

• Repositories of Electronic Theses and Dissertations (ETDs) are now becoming common in universities across the world.

• Data from these repositories suggest a dramatic increase in the use and citation of doctoral theses in current research activity, and this is related to the enhanced resource discovery and access that the digital surrogate confers.
IRS and Theses

• For researchers and institutions is essential to achieving visibility, citation and impact of their work.

• Therefore it is very important that repositories ensure interoperability and digital preservation in order to ensure visibility, accessibility and long-term use of their digital content.

• The interoperability of repositories and mandate policies positively affect visibility, citations and impact of research works deposited in institutional repositories.
IRS and Theses

University of Salamanca deposit of theses in the open access repository:
- Mandate 2008
IRS and Theses

• **Open Access is more easy for Theses** (ETDs) than for any other kind of research literature:
  – Authors have not yet transferred rights to a publisher.
  – Theses and dissertations are now born digital

• **Visibility:** By giving authors a foreseeable, real audience beyond the dissertation committee an OA policy strengthens existing incentives to do rigorous, original work.

• **Sharing theses** and dissertations that meet the school’s high standard reflects well on the institution and benefits other researchers in the field.
Open Access and citation

• Some researchers made a study based on a huge dataset of over one million papers published in the period of 1992-2003 in ten academic fields

• They concluded that Open Access papers have consistently received more citations than non-Open Access papers

Technological Ecosystems for Enhancing Multiculturality
Open Access Theses: visibility and citation

• And What’s the matter with theses? Do Open Access Theses have a citation advantage?

• Larivière, Zuccala and Archambault (2007) studied the impact of theses and the results of their study were paradoxical because while theses are becoming increasingly accessible to scholars in electronic format, their overall scientific impact seems to be declining.
Open Access Theses: visibility and citation

- Our aim is to prove through a simple analysis how e-theses have a high degree of use and visibility, and also a small but significant citation rate.
- The results prove that the open access repositories are a valid way to disseminate scientific grey literature of institutions.
- The OA repositories are a valid channel of publication scientific grey literature.
Interoperability

• OA repositories have been increasing and have become an important component of the infrastructure of e-research (COAR, 2012)
• The degree of visibility the repositories are closely related to interoperability of them
• All repositories need to support interoperability in order for their repository contents to be part of the global aggregate of OA knowledge.
• COAR (2015). *Roadmap Future Directions for Repository Interoperability*
Areas of GREDOS interoperability
Dissemination

- To increase discoverability and availability of electronic doctoral theses (ETDs) repositories should network and interconnect. This interconnection can be on different levels, like regional or state-wide, national or international (Schöpfel, 2013)
INTEROPERABILITY, DISSEMINATION AND VISIBILITY

Dissemination

• The collection of doctoral theses of the University of Salamanca, disseminated through GREDOS, is consulted from many places on the Web, thus increasing the number of visits and downloads of its contents

• Metadata of community doctoral theses of GREDOS is collected by thematic harvesters: TDR, DART-Europe and OATD

• Metadata are also disseminated through other service providers
Visibility

- The visibility results of the GREDOS Repository:
  - use of tools interoperability
  - and the description schemas that are compatible with
    - The Repository Networks
    - The harvesters
    - Google Scholar and the reliability of web indicators (search engines)
INTEROPERABILITY, DISSEMINATION AND VISIBILITY

Visibility

• The GREDOS Repository has a high level of visibility at Internet according the Ranking Web of Repositories

In the 17th edition of the Ranking Web (July 2015), GREDOS is ranked 71th position into the institutional repositories of the world (2.193 repositories)
Methods of analysis

• Tracking of the Open Access theses
  – Analysis of use statistics of open access theses in GREDOS to measure the visits and downloads module University of Tasmania. Through data visits and data downloads of theses we can analyzing their visibility on the Internet
  – Citation analysis of Open Access theses using the data provided by Google Scholar Citation. It provides a simple way to track of citations of works
# E-Theses in Gredos: Analysis of Visibility and Citation

## Resumen estadísticas Año/Mes

<table>
<thead>
<tr>
<th>Fecha</th>
<th>Visitas</th>
<th>Descargas</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-07</td>
<td>118599</td>
<td>222689</td>
</tr>
<tr>
<td>2015-06</td>
<td>194920</td>
<td>396284</td>
</tr>
<tr>
<td>2015-05</td>
<td>200415</td>
<td>414031</td>
</tr>
<tr>
<td>2015-04</td>
<td>196995</td>
<td>324389</td>
</tr>
<tr>
<td>2015-03</td>
<td>157215</td>
<td>319093</td>
</tr>
<tr>
<td>2015-02</td>
<td>121778</td>
<td>237990</td>
</tr>
<tr>
<td>2015-01</td>
<td>168106</td>
<td>199819</td>
</tr>
<tr>
<td>2014-12</td>
<td>226892</td>
<td>333595</td>
</tr>
<tr>
<td>2014-11</td>
<td>150104</td>
<td>247027</td>
</tr>
<tr>
<td>2014-10</td>
<td>161795</td>
<td>288863</td>
</tr>
<tr>
<td>2014-09</td>
<td>166783</td>
<td>209466</td>
</tr>
<tr>
<td>2014-08</td>
<td>119861</td>
<td>192124</td>
</tr>
<tr>
<td>2014-07</td>
<td>156860</td>
<td>252773</td>
</tr>
<tr>
<td>2014-06</td>
<td>156684</td>
<td>325422</td>
</tr>
<tr>
<td>2014-05</td>
<td>157055</td>
<td>567665</td>
</tr>
<tr>
<td>2014-04</td>
<td>130961</td>
<td>329413</td>
</tr>
<tr>
<td>2014-03</td>
<td>145598</td>
<td>235676</td>
</tr>
<tr>
<td>2014-02</td>
<td>142830</td>
<td>171425</td>
</tr>
<tr>
<td>2014-01</td>
<td>133186</td>
<td>158100</td>
</tr>
</tbody>
</table>

Statistics provided by the Statistics module University of Tasmania

Technological Ecosystems for Enhancing Multiculturality TEEM’15
E-THESSES IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

Data collection and sample selection of theses

• Theses defended at the University of Salamanca in the period 2006-2007 / 2010-2011

• Data on the theses defended at the University of Salamanca: TESEO database

• Data to know if the theses are Open Access: GREDOS
E-THeses IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

<table>
<thead>
<tr>
<th>Theses USAL 2006-2007 / 2010-2011</th>
<th>No OA Theses</th>
<th>OA Theses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1064</td>
<td>461</td>
<td>603</td>
</tr>
</tbody>
</table>

Test sample

<table>
<thead>
<tr>
<th>CATEGORIES</th>
<th>OA</th>
<th>No OA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Sciences</td>
<td>176</td>
<td>126</td>
<td>302</td>
</tr>
<tr>
<td>Experimental Sciences</td>
<td>85</td>
<td>63</td>
<td>148</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>210</td>
<td>178</td>
<td>388</td>
</tr>
<tr>
<td>Technological Sciences</td>
<td>23</td>
<td>16</td>
<td>39</td>
</tr>
<tr>
<td>Humanities</td>
<td>109</td>
<td>178</td>
<td>287</td>
</tr>
</tbody>
</table>

Fields of knowledge classification

Technological Ecosystems for Enhancing Multiculturality  TEEM’15
E-THESES IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

• For this research we analyzed 125 theses from the 1064 theses read in the indicated time period.

• In this article we selected 25 OA theses for each of the knowledge area. The sample was randomly drawn from the total number of OA theses.
E-THESSES IN GREDOSS: ANALYSIS OF VISIBILITY AND CITATION

OA/No OA Theses by Knowledge areas
E-THESES IN GREDOSS: ANALYSIS OF VISIBILITY AND CITATION

Percentage OA Theses per Knowledge areas

Technological Ecosystems for Enhancing Multiculturality TEEM’15
E-THESES IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

Analysis

- Life Sciences theses
  - Total views: 5938
  - Total downloads: 42057
  - Media views: 194
  - Media downloads: 689
  - Presence Google Scholar: 100%
  - Citations: 16%

Life Sciences theses scatter plot
E-THESIS IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

Analysis

- Experimental Sciences Theses
  - Total views: 5378
  - Total downloads: 63399
  - Media views: 194
  - Media downloads: 595
  - Presence Google Scholar: 100%
  - Citations: 32%

Experimental Sciences theses scatter plot
Analysis

• Humanities theses
  – Total views: 5851
  – Total downloads: 39663
  – Media views: 218
  – Media downloads: 502
  – Presence Google Scholar: 100%
  – Citations: 20%

Humanities theses scatter plot
E-THESES IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

Analysis

• Social Sciences theses
  – Total views: 7411
  – Total downloads: 86146
  – Media views: 253
  – Media downloads: 1267
  – Presence Google Scholar: 100%
  – Citations: 20%

Social Sciences theses scatter plot
E-THESIS IN GREDOS: ANALYSIS OF VISIBILITY AND CITATION

Analysis

- Technological Sciences Theses
  - Total views: 7155
  - Total downloads: 52540
  - Media views: 207
  - Media downloads: 747
  - Presence Google Scholar: 100%
  - Citations: 4%

Technological Sciences theses scatter plot
## Results

<table>
<thead>
<tr>
<th></th>
<th>Views</th>
<th>Downloads</th>
<th>Media views</th>
<th>Media downloads</th>
<th>GS</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Life Sciences</strong></td>
<td>5938</td>
<td>42057</td>
<td>194</td>
<td>689</td>
<td>100%</td>
<td>16%</td>
</tr>
<tr>
<td><strong>Experimental Sciences</strong></td>
<td>5378</td>
<td>63399</td>
<td>194</td>
<td>595</td>
<td>100%</td>
<td>32%</td>
</tr>
<tr>
<td><strong>Humanities</strong></td>
<td>5851</td>
<td>39663</td>
<td>218</td>
<td>502</td>
<td>100%</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Social Sciences</strong></td>
<td>7411</td>
<td>86146</td>
<td>253</td>
<td>1267</td>
<td>100%</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Technological Sciences</strong></td>
<td>7155</td>
<td>52540</td>
<td>207</td>
<td>747</td>
<td>100%</td>
<td>4%</td>
</tr>
</tbody>
</table>

### Analysis results

Technological Ecosystems for Enhancing Multiculturality  TEEM’15
Conclusions

• Among the main reasons that the last decade saw an explosion of repositories include precisely the benefits derived in visibility, impact, use and dissemination of scientific production
• The study conclusions
  1. Open Access repositories can obtain information on the use (visibility), and citation (impact) of the doctoral thesis, information cannot be obtained in the case of theses that are not on open access
  2. Open Access theses longer grey literature in view of usability measured by visits and downloads them
  3. Open Access institutional repositories become a new form of scientific communication for dissertations, grey literature considered
  4. The recovery is enhanced through harvesters and search engines (Google) directly through the repository, as evidenced by data downloads
  5. The theses available in Open Access increase your visibility and use but the citation index is low
  6. Based on the sample analyzed it seem that there is not a correlation between the number of views / downloads and the citation rate
References


References

• Ferreras-Fernández, T., 2013. La interoperabilidad: el pegamento técnico para conectar repositorios. Notice in Blok de Bik. Available at: http://www.ub.edu/blokdebid/es/content/la-interoperabilidad-el-pegamento-t%C3%A9cnico-para-conectar-repositorios


References


• This paper may be cited

Open access repositories as channel of publication scientific grey literature

Tránsito Ferreras-Fernández
Francisco J. García-Peñalvo
José A. Merlo-Vega

Contact
transiff@usal.es / @transi
fgarcia@usal.es / @frangp
merlo@usal.es / @merlovega

GRIAL Research Group, University of Salamanca