The LSE, The Blogs and The Metadata

Helen Williams
Metadata Manager
LSE Library
Higher Education Innovation Funding

Knowledge exchange between Universities & the Wider World which result in Economic & Social benefit to the UK
Knowledge Exchange and Impact

“Supporting knowledge exchange throughout the research lifecycle”

LSE Impact Blog World-leading discussion about the impact of social sciences
A systematic identification and analysis of scientists on Twitter

Fig 3

Top scientific domains by disciplines.

We extract URLs from tweets and retweets and then count the appearances of the scientific domains in each of the 10 disciplines.

https://doi.org/10.1371/journal.pone.0175368
LSE Library Strategy

2.iii Secure the collection and preservation of the complete intellectual output of the School in partnership with academic and professional services colleagues.
“Metadata enrichment for research output and visibility measurement”
Lesson 1

• Short term niche recruitment may be a challenge

• What are the options for using internal skills and backfilling existing work instead?
Lesson 2

- Planning planning planning

ALWAYS PLAN AHEAD.
IT WASN'T RAINING WHEN NOAH BUILT THE ARK.
-RICHARD C. CUSHING
<table>
<thead>
<tr>
<th>Staff time involved: Current</th>
<th>Staff time involved: Project estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDFs have already been created, and are supplied by email. Each PDF is saved, by Metadata Services staff in the appropriate \libfile\repository\Content LSE Blog XXX folder, filed by year and month (Each PDF will automatically save with the right file name, so staff do not need to ‘name’ the PDF files. Cover sheets are not required for LSE blog posts.)</td>
<td>Working from live blog, check blog post is suitable for addition Scroll down and select print friendly, followed by create pdf, download pdf, select open, drag to libfile and save in appropriate folder</td>
</tr>
<tr>
<td>1.5 mins</td>
<td>2.5 mins</td>
</tr>
<tr>
<td>Metadata record created by locating template record, selecting ‘Record Administration’ and then ‘Use as Template’ on the Actions tab.</td>
<td>Metadata record created by locating template record, selecting ‘Record Administration’ and then ‘Use as Template’ on the Actions tab.</td>
</tr>
<tr>
<td>1 min</td>
<td>1 min</td>
</tr>
<tr>
<td>A new duplicate record is created and the following fields are altered to create a new record, by selecting ‘Edit’ on the Details section: Title, Abstract, Creators, Date, Published Item URL, LSE Experts page, LC subject classification, JEL code (if economics related). Additional information – copyright.</td>
<td>A new duplicate record is created and the following fields are altered to create a new record, by selecting ‘Edit’ on the Details section: Title, Abstract, Creators, Date, Published Item URL, LSE Experts page, LC subject classification, JEL code (if economics related). Additional information – copyright.</td>
</tr>
<tr>
<td>5 mins</td>
<td>5 mins</td>
</tr>
<tr>
<td>The relevant PDF is uploaded by clicking on the Upload tab, selecting ‘Browse’ and locating the PDF previously saved in libfile, selecting ‘Open’ and when the link appears in the New Document box, clicking on ‘Upload’.</td>
<td>The relevant PDF is uploaded by clicking on the Upload tab, selecting ‘Browse’ and locating the PDF previously saved in libfile, selecting ‘Open’ and when the link appears in the New Document box, clicking on ‘Upload’.</td>
</tr>
<tr>
<td>0.5 mins</td>
<td>0.5 mins</td>
</tr>
<tr>
<td>The item is deposited by clicking on the Deposit tab, selecting ‘Deposit Item Now’, before moving to the Review page and clicking on the green tick ‘Move to Repository’ icon.</td>
<td>The item is deposited by clicking on the Deposit tab, selecting ‘Deposit Item Now’, before moving to the Review page and clicking on the green tick ‘Move to Repository’ icon.</td>
</tr>
<tr>
<td>0.5 mins</td>
<td>0.5 mins</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
</tr>
<tr>
<td>9 mins</td>
<td>10 mins</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Top 10</th>
<th>Current Research</th>
<th>Archived</th>
<th>Community</th>
<th>Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blog posts to be added</td>
<td>5829 / 11,626</td>
<td>5627</td>
<td>699</td>
<td>2505</td>
<td>2413</td>
</tr>
<tr>
<td>Min totals</td>
<td>5829</td>
<td>11,456</td>
<td>12,155</td>
<td>14,560</td>
<td>17,073</td>
</tr>
<tr>
<td>Max totals</td>
<td>11,626</td>
<td>17,253</td>
<td>17,952</td>
<td>20,457</td>
<td>22,870</td>
</tr>
</tbody>
</table>
Lessons 3 – 5

• There may be funding restrictions
• Continuous employment may impact on salary
• Regulations can change!
Automation

• The project used manual processes

• As part of the project the creation of PDFs was automated...

And the creation of metadata was semi-automated
5 most commonly used DC elements for archiving web content

• Description
• Creator
• Subject
• Title
• Date
** LSE Brexit
* -----------------------------------------------------------------------
*/
@Article{
    author = "Blog Author1 and Blog Author2",
    abstract = "Enter blog abstract here...",
    title = "Enter LSE Brexit blog title here...",
    publisher = "London School of Economics and Political Science",
    url = "http://blogs.lse.ac.uk/brexit/",
    journal = "LSE Brexit",
    year = "2017",
    note = "© YEAR The Author(s)",
}

/**
* Impact of Social Sciences
* -----------------------------------------------------------------------
*/
@Article{
    author = "Blog Author1 and Blog Author2",
    abstract = "Enter blog abstract here...",
    title = "Enter Impact of Social Sciences blog title here...",
    publisher = "London School of Economics and Political Science",
    url = "http://blogs.lse.ac.uk/impactofsocialsciences/",
    journal = "Impact of Social Sciences Blog",
    year = "2017",
    note = "© YEAR The Author(s) CC BY 3.0",
}
/**
Guess the results...
The results are in...

• 11,665 blog posts added to LSERO

• Total unique content from

  • 42 current research blogs
  • 8 archived research blogs
  • 6 partner blogs
Sustainability number crunching

Average blog posts x Time to create record + Time to check duplicate content = 12 hours a week 😞

Average top 10 blog posts x Time to create record + Time to check duplicate content = 6 hours a wk – 2.5 hours spent on old procedures = 4 hours a wk ÷ 4 people = 1 hour a wk each 😊
Project evaluation

• Knowledge Exchange and Impact Strategy
  archive all official blog output

• Library Strategy
  secure the collection and preservation of the complete intellectual output of the School
LSERO item type downloads as percentage of total content

- LSERO monograph downloads:
  - 2015: 40%
  - 2016: 36%
  - 2017: 30%
  - 2018: 29%

- LSERO article downloads:
  - 2015: 34%
  - 2016: 36%
  - 2017: 33%
  - 2018: 32%

- LSERO blog downloads:
  - 2015: 11%
  - 2016: 14%
  - 2017: 25%
  - 2018: 26%

- LSERO book chapter downloads:
  - 2015: 10%
  - 2016: 9%
  - 2017: 7%
  - 2018: 8%
Feedback

Delighted

Fantastic opportunity

Brilliant

Very timely... responsible closure

Many thanks

Honoured

Excellent initiative

Great
Phase 2 ideas

- Review metadata standards for live blog posts to improve discoverability

- Add DOIs to give increased stability to blog content, and to consider impact measurement via Altmetric

- Consider auto-classification for this content to scale up metadata work in this area
LSE blogs: http://blogs.lse.ac.uk/

LSERO: http://eprints.lse.ac.uk/

H.K.Williams@lse.ac.uk

@HelsKRW
Photo credits

Slides 2, 4-7, 9-12, 14-26: CC0, 3dman-eu: Pixabay
Slide 3: Beaver, LSE Library
Slide 11: Always Plan Ahead, CC BY 2.0, Raphael Love: Flickr
Slide 14: British Pound Sign, CC BY 4.0, Rareclass: Flickr