Making scholarly statistics count in UK repositories

Ross MacIntyre, Mimas, The University of Manchester
Paul Needham, Cranfield University
November 2012
IRUS-UK

- Funded by JISC as part of UK RepositoryNet+
- Lead by Mimas
- Primary Project Team Members:
  - Mimas
  - Cranfield University
  - EvidenceBase, BCU
- IRUS-UK: Institutional Repository Usage Statistics - UK
IRUS-UK: history

- Emerged as an outcome of PIRUS2
  - Publisher and Institution Repository Usage Statistics project
  - [http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/](http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/)
  - Aimed to develop a global standard to enable the recording, reporting and consolidation of online usage statistics for individual journal articles hosted by Institutional Repositories, Publishers and others
  - Proved it was technically feasible, but thwarted by organisational and political issues
  - However, crisis/opportunity… we now knew it was possible to…
IRUS-UK: aim

- Enable UK IRs to share/expose usage statistics based on a global standard – COUNTER
  - Produced on the same basis as publishers
  - Filtered to remove robots and double clicks
  - Comparable
  - Reliable
  - Trustworthy
  - Authoritative
IRUS-UK: objectives

• Collect raw usage data from UK IRs for all item types within repositories
  • Downloads not record views
• Process those raw data into COUNTER-compliant statistics
• Return those statistics back to the originating repositories for their own use
IRUS-UK: objectives

- Give JISC (and others) a nation-wide picture of the overall use of UK repositories
  - demonstrate their value and place in the dissemination of scholarly outputs
- Offer opportunities for benchmarking
- Act as an intermediary between UK repositories and other agencies
  - e.g. global central clearinghouse, national shared services, OpenAIRE
IRUS-UK: gathering data

- Considered 2 scenarios for gathering data
  - ‘Tracker’ code
    - a server-side ‘Google Analytics’ for item downloads
    - Pushes metadata to a third-party server as OpenURL Key/Value strings
  - OAI-PMH harvesting
    - Used to by third parties to ‘pull’ metadata from repositories
    - Repurposed to expose usage events as OpenURL Context Objects
- Opted for the Tracker
  - Just easier
  - Plugins/patches available for DSpace and Eprints
IRUS-UK: gathering data

- IRUS logs receive OpenURL strings
  - 81.17.244.212 - - [06/Jul/2012:00:00:00 +0100] "HEAD /counter/?url_ver=Z39.88-2004&
  - url_tim=2012-07-05T22%3A59%3A59Z&
  - req_id=urn%3Aip%3A86.15.47.114&
  - req_dat=Mozilla%2F5.0+(iPhone%3B+U%3B+CPU+iPhone+OS+5_1_1+like+Mac+OS+X
    %3B+en-us)+ AppleWebKit%2F534.46.0+(KHTML%2C+like+Gecko)+CriOS%2F19.0.1084.60+Mobile
    e%2F9B208+Safari%2F7534.48.3&
  - rft.artnum=oai%3Aeprints.hud.ac.uk%3A8795&
  - svc_format=application%2Fpdf&
  - rfr_id=eprints.hud.ac.uk HTTP/1.1" 200 - "- "EPrints 3.2.2 (Chocolate Cake) [Born on
    2010-05-25]"
IRUS-UK: processing data

- Logs are processed daily

- Step 1: Perl script parses the logs
  - Processes entries from recognised IRs
  - Sorts and filters entries following COUNTER rules
  - Consolidates daily accesses for each item
  - Outputs to intermediate file

- Step 2: Perl script parses intermediate file
  - Looks up each item in the IRUS DB
    - If item is unknown to the system add item with (most) metadata “unknown”
  - Updates DB with new statistics (for both ‘known’ & ‘known unknowns’)

- Step 3: Obtain “unknown” metadata
  - For the ‘known unknowns’ uses an OAI GetRecord to retrieve
  - Updates the metadata to DB
**OVERALL SUMMARY STATISTICS**

**OVERALL SUMMARY AS AT 7TH NOV 2012**

*At a glance*
- **Repositories**: 6
- **Items**: 21,002
- **Downloads**: 567,358
- **Downloads up to Oct-2012**: 594,628
- **Downloads this month**: 42,730

**SUMMARY BY REPOSITORY**

<table>
<thead>
<tr>
<th>Repository</th>
<th>From Jul-2012 to Oct-2012</th>
<th>This month, so far</th>
<th>Total Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bournemouth University BURO</td>
<td>4,505</td>
<td>2,366</td>
<td>10,813</td>
</tr>
<tr>
<td>City Research Online</td>
<td>1,101</td>
<td>466</td>
<td>20,950</td>
</tr>
<tr>
<td>Cranfield University CERRS</td>
<td>4,895</td>
<td>2,896</td>
<td>201,914</td>
</tr>
<tr>
<td>University of Huddersfield Repository</td>
<td>3,815</td>
<td>1,928</td>
<td>76,680</td>
</tr>
<tr>
<td>University of Reading CIRUR</td>
<td>802</td>
<td>495</td>
<td>12,093</td>
</tr>
<tr>
<td>University of Salford USR</td>
<td>6,145</td>
<td>2,454</td>
<td>134,743</td>
</tr>
</tbody>
</table>
## Item Type Summary Statistics

### Summary by IRUS Item Type

<table>
<thead>
<tr>
<th>Item Type</th>
<th>From Jul 2012 to Oct 2012</th>
<th>This Month, so far...</th>
<th>Total Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IRUS Item Type</strong></td>
<td><strong>Items</strong></td>
<td><strong>Downloads</strong></td>
<td><strong>Items</strong></td>
</tr>
<tr>
<td>Article</td>
<td>6,890</td>
<td>187,278</td>
<td>3,812</td>
</tr>
<tr>
<td>Book</td>
<td>338</td>
<td>6,285</td>
<td>149</td>
</tr>
<tr>
<td>Book Chapter</td>
<td>720</td>
<td>13,377</td>
<td>368</td>
</tr>
<tr>
<td>Book Section</td>
<td>725</td>
<td>13,349</td>
<td>408</td>
</tr>
<tr>
<td>Conference or Workshop Item</td>
<td>3,634</td>
<td>73,286</td>
<td>2,005</td>
</tr>
<tr>
<td>Dataset</td>
<td>5</td>
<td>121</td>
<td>3</td>
</tr>
<tr>
<td>Exam Paper</td>
<td>2,356</td>
<td>6,600</td>
<td>358</td>
</tr>
<tr>
<td>Manual</td>
<td>3</td>
<td>34</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>1,897</td>
<td>45,157</td>
<td>886</td>
</tr>
<tr>
<td>Presentation</td>
<td>10</td>
<td>140</td>
<td>5</td>
</tr>
<tr>
<td>Report</td>
<td>389</td>
<td>5,158</td>
<td>210</td>
</tr>
<tr>
<td>Technical Report</td>
<td>107</td>
<td>2,943</td>
<td>54</td>
</tr>
<tr>
<td>Thesis or Dissertation</td>
<td>2,999</td>
<td>126,400</td>
<td>1,968</td>
</tr>
<tr>
<td>Unknown</td>
<td>100</td>
<td>1,223</td>
<td>22</td>
</tr>
<tr>
<td>Working Paper</td>
<td>581</td>
<td>21,527</td>
<td>343</td>
</tr>
</tbody>
</table>
## SEARCH RESULTS

Search for "maur" returned 7 results:

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Title</th>
<th>Author</th>
<th>Overall Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article</td>
<td>Developing the wider role of business in society; the experience of Microsoft in developing training and supporting employability.</td>
<td>Bonfaglione, Elena; Moir, Lance; Ambrosini, Veronique</td>
<td></td>
</tr>
<tr>
<td>Article</td>
<td>Does 'Best Practice' in Setting Executive Pay in the UK Encourage 'Good Behaviour'?</td>
<td>Bender, Ruth; Moir, Lance</td>
<td></td>
</tr>
<tr>
<td>Article</td>
<td>Financial Instruments.</td>
<td>Moir, Lance</td>
<td></td>
</tr>
<tr>
<td>Article</td>
<td>What do we mean by corporate social responsibility?</td>
<td>Moir, Lance</td>
<td></td>
</tr>
<tr>
<td>Conference or Workshop Item</td>
<td>Communicating Philanthropy: Multiple Representations to Multiple Stakeholders?</td>
<td>Moir, Lance</td>
<td></td>
</tr>
<tr>
<td>Conference or Workshop Item</td>
<td>Why do firms give - altruism, advertising or stakeholder management? Development of a framework to understand corporate philanthropy and its application to business involvement with the Arts in the UK.</td>
<td>Moir, Lance</td>
<td></td>
</tr>
<tr>
<td>Thesis or Dissertation</td>
<td>Why does business support the Arts? Philanthropy, marketing or legitimation?</td>
<td>Moir, Lance</td>
<td></td>
</tr>
</tbody>
</table>
### Item Report 1 (IR1): Number of Successful Item Download Requests by Month and Repository Identifier

**Cranfield University CERES**

**Date Run:** 08/11/2012

**Sort report by:** [Downloads] Descending

<table>
<thead>
<tr>
<th>ID</th>
<th>Title</th>
<th>Jul-2012</th>
<th>Aug-2012</th>
<th>Sep-2012</th>
<th>Oct-2012</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>43217</td>
<td>42723</td>
<td>42492</td>
<td>57118</td>
<td>187450</td>
</tr>
<tr>
<td></td>
<td>What do we mean by corporate social responsibility?</td>
<td>550</td>
<td>538</td>
<td>497</td>
<td>829</td>
<td>2414</td>
</tr>
<tr>
<td></td>
<td>Power and energy management of multiple energy storage systems in electric vehicles</td>
<td>420</td>
<td>386</td>
<td>369</td>
<td>473</td>
<td>1660</td>
</tr>
<tr>
<td></td>
<td>Key Account Planning: Benefits, Barriers and Best Practice.</td>
<td>333</td>
<td>354</td>
<td>353</td>
<td>484</td>
<td>1524</td>
</tr>
<tr>
<td></td>
<td>Relationship marketing: The six markets framework</td>
<td>194</td>
<td>249</td>
<td>322</td>
<td>357</td>
<td>1420</td>
</tr>
<tr>
<td></td>
<td>Towards a definition of a business performance measurement system</td>
<td>298</td>
<td>265</td>
<td>328</td>
<td>428</td>
<td>1320</td>
</tr>
<tr>
<td></td>
<td>Airline Jet Fuel Hedging: Theory and practice</td>
<td>258</td>
<td>239</td>
<td>297</td>
<td>509</td>
<td>1293</td>
</tr>
<tr>
<td></td>
<td>Thirty Years of Mergers and Acquisitions Research: Recent Advances and Future Opportunities</td>
<td>313</td>
<td>302</td>
<td>264</td>
<td>357</td>
<td>1236</td>
</tr>
<tr>
<td></td>
<td>Supply chain risk management: outlining an agenda for future research.</td>
<td>208</td>
<td>260</td>
<td>279</td>
<td>409</td>
<td>1213</td>
</tr>
</tbody>
</table>
IRUS-UK: CAR1 Report S/shot

CONSOLIDATED ARTICLE REPORT 1 (CAR1), NUMBER OF SUCCESSFUL MONTHLY ARTICLE DOWNLOAD REQUESTS BY DOI AND REPOSITORY IDENTIFIER

Consolidated Article Report 1 (CAR1), Number of Successful Monthly Article Download Requests by DOI and Repository Identifier

To the Central Clearing House

Date Run: 08/11/2012

Sort report by: Downloads Descending Update

<table>
<thead>
<tr>
<th>DOI</th>
<th>SourceID</th>
<th>Repository</th>
<th>Platform</th>
<th>Oct-2012</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://dx.doi.org/10.1002/tmi.3312">http://dx.doi.org/10.1002/tmi.3312</a></td>
<td><a href="http://hdl.handle.net/1826/3312">http://hdl.handle.net/1826/3312</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>510</td>
<td>510</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://hdl.handle.net/1826/3029">http://hdl.handle.net/1826/3029</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>484</td>
<td>484</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>484</td>
<td>484</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>University of Oxford Eprints</td>
<td>Eprints</td>
<td>403</td>
<td>403</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>428</td>
<td>428</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>393</td>
<td>393</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>371</td>
<td>371</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>357</td>
<td>357</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>359</td>
<td>359</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>342</td>
<td>342</td>
</tr>
<tr>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td><a href="http://dx.doi.org/10.1002/0144156600679524">http://dx.doi.org/10.1002/0144156600679524</a></td>
<td>Cranfield University CERES</td>
<td>DSpace</td>
<td>341</td>
<td>341</td>
</tr>
</tbody>
</table>
IRUS-UK: getting statistics out

- Various reports are/will be available to Institutions
  - For humans, downloads as CSV/Excel spreadsheet files
    - Monthly and daily granularity
  - For machines, harvesting via SUSHI
    - Monthly granularity

- Usage statistics for incorporation into Repositories
  - Via an API/Web Service
    - details yet to be determined
IRUS-UK: where we are now

- Pioneers sending data to IRUS
  - Bournemouth, City, Huddersfield, Reading, Salford (Eprints)
  - Cranfield (Dspace)
  - Others in the pipeline (~10 in testing currently)

- We’re working on:
  - Ingest scripts
  - Portal UI – basic until informed choices can be made
  - Spreading the word about IRUS
IRUS-UK: next steps

• More of what we’re already doing…

• Community Engagement
  • Getting more IRs on-board
  • Defining and evaluating user requirements
    • including surveys, interviews, monitoring feedback, etc.

• Processes to support
  • institutional liaison, data collection, monitoring and responding to future requirements, ensuring interoperability
  • development and production environments including monitoring, back-up, authentication and archiving polices and processes

• Formal application for COUNTER-compliance
Contacts & Information

• For general enquiries, please contact
  • support@repositorynet.ac.uk

• If you are a UK repository wishing to participate in IRUS-UK, please contact
  • irus@mimas.ac.uk

• Project web site:
  • http://www.irus.mimas.ac.uk/

• Thank you!