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%2F9B208+Safari%2F7534.48.3&rft.artnum=oai%3Aeprints.hud.ac.uk%3A8795&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
IRUS-UK: Overall Summary S/shot

OVERALL SUMMARY STATISTICS

OVERALL SUMMARY AS AT 7TH NOV 2012

At a glance
- Repositories: 6
- Items: 21,032
- Downloads: 5,873,558
- Downloads up to Oct-2012: 504,628
- 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></td>
<td>Items</td>
<td>Downloads</td>
<td>Items</td>
</tr>
<tr>
<td>Bournemouth University BURO</td>
<td>4,505</td>
<td>94,038</td>
<td>2,366</td>
</tr>
<tr>
<td>City Research Online</td>
<td>1,101</td>
<td>16,405</td>
<td>466</td>
</tr>
<tr>
<td>Cranfield University CERES</td>
<td>4,895</td>
<td>187,450</td>
<td>2,896</td>
</tr>
<tr>
<td>University of Huddersfield Repository</td>
<td>3,815</td>
<td>70,351</td>
<td>1,928</td>
</tr>
<tr>
<td>University of Reading CronUR</td>
<td>802</td>
<td>13,727</td>
<td>495</td>
</tr>
<tr>
<td>University of Salford USR</td>
<td>6,145</td>
<td>126,257</td>
<td>2,454</td>
</tr>
</tbody>
</table>
IRUS-UK: Item Types S/shot

<table>
<thead>
<tr>
<th>ITEM TYPE SUMMARY STATISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>From Jul 2012 to Oct 2012</td>
</tr>
<tr>
<td>This month, so far...</td>
</tr>
<tr>
<td>Total Downloads</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td><strong>IRUS Item Type</strong></td>
</tr>
<tr>
<td>Article</td>
</tr>
<tr>
<td>Book</td>
</tr>
<tr>
<td>Book Chapter</td>
</tr>
<tr>
<td>Book Section</td>
</tr>
<tr>
<td>Conference or Workshop Item</td>
</tr>
<tr>
<td>Dataset</td>
</tr>
<tr>
<td>Exam Paper</td>
</tr>
<tr>
<td>Manual</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Presentation</td>
</tr>
<tr>
<td>Report</td>
</tr>
<tr>
<td>Technical Report</td>
</tr>
<tr>
<td>Thesis or Dissertation</td>
</tr>
<tr>
<td>Unspecified</td>
</tr>
<tr>
<td>Working Paper</td>
</tr>
</tbody>
</table>
IRUS-UK: (Author) Search S/shot
### 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] [Update]

<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>What do we mean by corporate social responsibility?</td>
<td>43217</td>
<td>42723</td>
<td>44292</td>
<td>57118</td>
<td>187450</td>
</tr>
<tr>
<td></td>
<td>Power and energy management of multiple energy storage systems in electric vehicles</td>
<td>550</td>
<td>538</td>
<td>497</td>
<td>839</td>
<td>2414</td>
</tr>
<tr>
<td></td>
<td>Key Account Planning: Benefits, Barriers and Best Practice.</td>
<td>420</td>
<td>398</td>
<td>399</td>
<td>473</td>
<td>1660</td>
</tr>
<tr>
<td></td>
<td>Relationship marketing: The six markes framework</td>
<td>333</td>
<td>354</td>
<td>353</td>
<td>484</td>
<td>1524</td>
</tr>
<tr>
<td></td>
<td>Towards a definition of a business performance measurement system</td>
<td>194</td>
<td>243</td>
<td>322</td>
<td>357</td>
<td>1420</td>
</tr>
<tr>
<td></td>
<td>Airline jet fuel hedging: Theory and practice</td>
<td>298</td>
<td>265</td>
<td>328</td>
<td>428</td>
<td>1320</td>
</tr>
<tr>
<td></td>
<td>Corporate Social Responsibility in Supply Chains of Global Brands: A Boundaryless Responsibility</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>245</td>
<td>260</td>
<td>260</td>
<td>510</td>
<td>1284</td>
</tr>
<tr>
<td></td>
<td>Supply chain risk management: outlining an agenda for future research.</td>
<td>313</td>
<td>302</td>
<td>264</td>
<td>357</td>
<td>1236</td>
</tr>
</tbody>
</table>

---

[IRUS-UK: IR1 Report S/shot](https://ir-us.uk/secure/iris/1417)

[IRUS-UK](https://ir-us.uk)

[irus.mimas.ac.uk](https://irus.mimas.ac.uk)
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: 09/11/2012

Sort report by: Downloads Descending

[Table of DOI, SourceID, Repository, Platform, Oct-2012, Downloads]
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 website:
  - http://www.irus.mimas.ac.uk/

- Thank you!