



**IRUS-UK**

Making scholarly statistics count in UK repositories

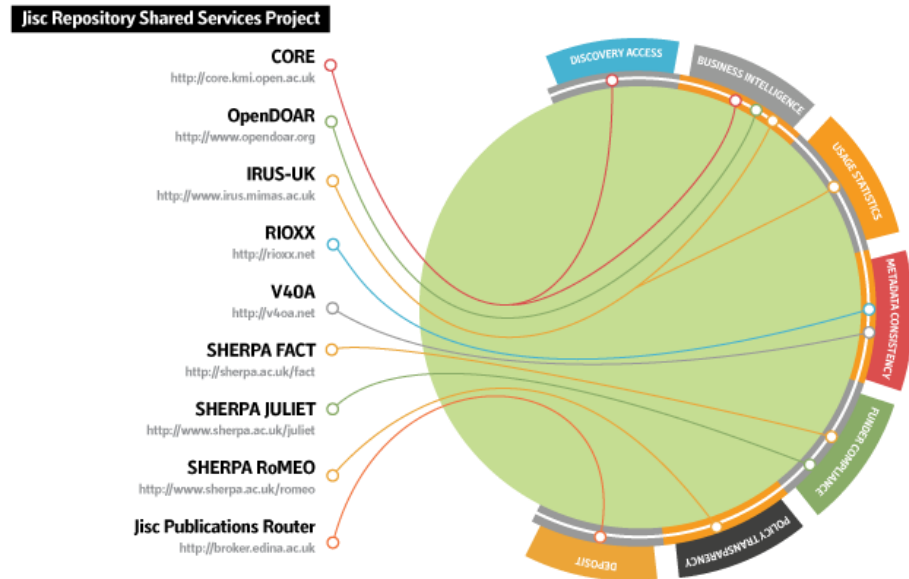
Ross MacIntyre, Mimas

*1:AM Altmetrics Conference, September 2014*

# IRUS-UK

- IRUS-UK: Institutional Repository Usage Statistics – UK
- Project Team Members:
  - Mimas – Project & Service Management & Host
  - Cranfield University - Development
  - Evidence Base, Birmingham City University – User Engagement & Evaluation
- Funded by Jisc

**Bringing together key repository services to deliver a connected national infrastructure to support OA**

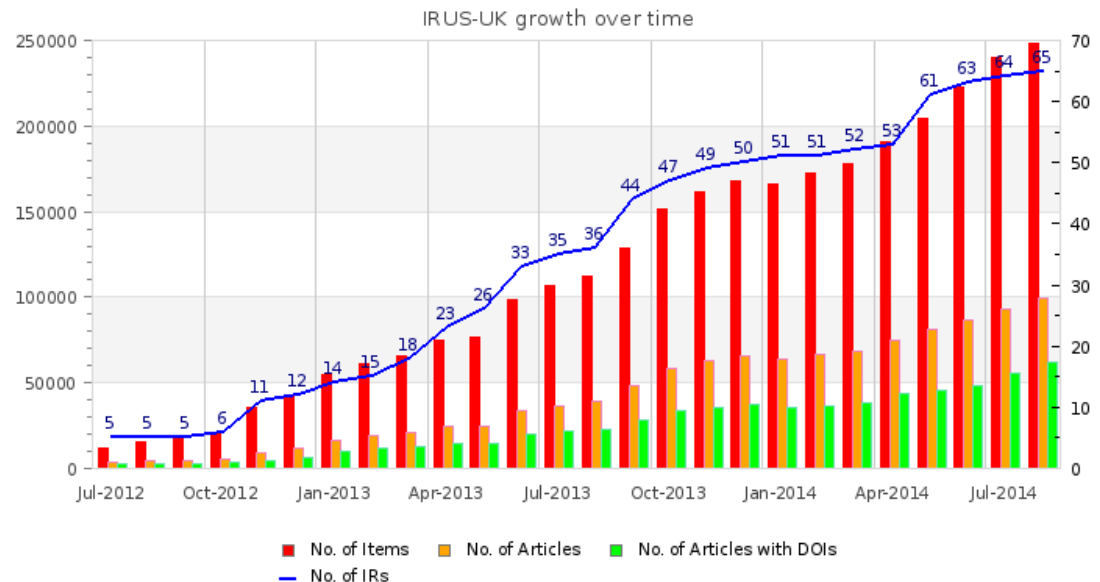


# IRUS-UK: background

- PIRUS2 (Publisher and Institutional Repository Usage Statistics)
  - Aimed to develop a global standard to enable the recording, reporting and consolidation of online usage statistics for individual journal articles hosted by IRs, Publishers and others
  - Proved it was \*technically feasible\*, but (initially) easier without ‘P’
  - <http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/>
- Main outcomes:
  - IRUS-UK
    - <http://www.irus.mimas.ac.uk/>
  - Release 1 of the COUNTER Code of Practice for Articles
    - <http://www.projectcounter.org/counterarticles.html>

# IRUS-UK: aims and objectives

- Provide a national aggregation service, enabling (a growing number of) UK Institutional Repositories to share/expose usage statistics at the individual item level, based on a global standard – COUNTER
  - Collect raw download data from UK IRs for \*all item types\* within repositories
  - Process those raw data into COUNTER-compliant statistics
- Facilitate comparable, standards-based measurements
- Provide an evidence base for repositories to develop policies and initiatives to help support their objectives
- Provide consistent and comprehensive statistics, presenting opportunities for benchmarking at a national level



# IRUS-UK: gathering data

- The method we use to gather raw download data is simple:
  - Whenever a file is downloaded from a participating repository, it sends a message to the IRUS-UK server with some details about the download
- Accomplished by adding a small piece of code to repository software, which employs the ‘Tracker Protocol’
  - <http://www.irus.mimas.ac.uk/help/toolbox/TrackerProtocol-V3-2014-04-22.pdf>
  - Pushes minimal raw download metadata to a third-party server as OpenURL (like) Key/Value strings
  - Patches for DSpace (1.8.x, 3.x, 4.1) and Plug-in for Eprints (3.2-3.3.x)
  - Implementation guidelines for Fedora
- Not in IRUS-UK scope, but also successfully deployed by:
  - OAPEN Library - freely accessible academic books, ARNO software
  - CORE - millions of scholarly articles aggregated from many Open Access repositories

# IRUS-UK: processing data

- The ingest process has been described in detail previously, see earlier presentations and webinars, available from the IRUS-UK News page: <http://www.irus.mimas.ac.uk/news/>
- The key point is that we adhere to the processing rules specified in:
  - Release 4 of the COUNTER Code of Practice for e-Resources
  - Release 1 of the COUNTER Code of Practice for Articles
- i.e. we filter out robot accesses and double clicks on the same basis as scholarly publishers
- The COUNTER Robot Exclusion list is specified only as a \*minimum requirement\*
  - So we've supplemented the COUNTER CoPs by adding additional filters to
    - Remove more user agents
    - Apply a simple threshold for 'overactive' IP addresses
- However, there's still more can be done!

# IRUS-UK: robots and unusual usage

- We commissioned *Information Power* to:
  - Analyse raw data we've collected since July 2012
  - Test the feasibility of devising a set of algorithms that would 'dynamically' identify and filter out unusual usage/robot activity
  - A report on that work is available from <http://www.irus.mimas.ac.uk/news/>
- Key findings from the work are
  - Suspicious behaviour can't necessarily be judged on the basis of one day's usage records or a month's.
  - At certain levels of activity machine/non-genuine usage is practically indistinguishable from genuine human activity.
- Taking this forward
  - We're involved in the recently formed COUNTER Working Group on Robots
  - Devising more sophisticated - but practical - algorithms to filter out 'rogue' usage
  - Outcomes will eventually become incorporated into the COUNTER standard
  - And, of course, adopted by IRUS-UK!

# IRUS-UK: Exposing statistics

- Web User Interface - The IRUS-UK Portal
  - Access currently behind Shibboleth authentication/authorisation, though with community agreement we hope to make it openly accessible:  
*Open Access, Open Data, Open Metrics!*
  - The portal offers:
    - a wide range of views – slicing and dicing stats from the IRUS-UK database
    - Reports available for download as CSV/Excel spreadsheet files
    - Altmetric donuts for individual items 😊
- SUSHI service
  - standard client/server web service utilizing a SOAP request/response to retrieve the XML version of COUNTER or COUNTER-like reports
- SUSHI Lite API
  - A new, simpler '21<sup>st</sup> century' approach, under development by the NISO SUSHI Lite Technical Report Working Group ([http://www.niso.org/workrooms/sushi/sushi\\_lite/](http://www.niso.org/workrooms/sushi/sushi_lite/))
  - RESTful: uses standard HTTP GET returning JSON
  - Allows retrieval of stats snippets to be embedded into Repository (and other) web pages



# IRUS-UK: Overall Summary

The screenshot shows a web browser window displaying the IRUS-UK portal. The browser's address bar shows the URL <http://www.irus.mimas.ac.uk/portal/>. The page features a navigation menu with links for HOME, ABOUT, HELP, PARTICIPANTS, NEWS, and PORTAL. A search bar is visible with the text "RESTful" and a search button. The main content area is titled "IRUS-UK" and "Home". Below the title, there is a section for "Overall Summary as at 22nd Sep 2014". This section includes a paragraph explaining that the table provides a summary of data in IRUS-UK, including the number of participating repositories, items downloaded, and download statistics. A link is provided to download the table as a PDF. Below this, there is a table titled "Headlines at a glance" with the following data:

Headlines at a glance	
Repositories	67
Items	260,421
Downloads up to Aug-2014	20,165,963
Downloads this month	863,499
Total Downloads	21,029,462

The left sidebar contains a navigation menu with the following items:

- Home
- Search
- Statistics Views
  - Repository stats
  - Country stats
  - Platform stats
  - ItemType stats
  - Ingest Stats
- Metadata Views
  - IRUS-UK Itemtype mappings
  - DOI summary
  - Article DOI summary by IR
- Statistics Reports
  - Article Report 4 (AR4) **NEW!**
  - Item Report 1 (IR1)
  - Item Report 2 (IR2)
  - ETD Report 1 (ETD1)
  - Repository Report 1 (RR1) **NEW!**

# IRUS-UK: Best features (our survey said)

- Reliable, authoritative statistics
  - COUNTER compliant statistics
  - Filtering of robots
  - IR1 used for reporting to SCONUL
  - Can repurpose for other reporting mechanisms and different audiences
- Ability to benchmark against others
- Comparison of download statistics across participating IRs
- Number (and range across the sector) of participating institutions
- Easy to use
  - Easy to setup and use
  - User friendly way to get stats
  - 96% find the current user interface clear
  - 96% find the current functionality clear to understand

# IRUS-UK: Value

- An altmetric that no-one else is providing - yet!
- Demonstrates the importance of repositories in disseminating scholarly outputs
- Uniquely positioned to act as an intermediary between UK repositories and other actors:
  - Funders
  - Publishers
  - National shared services
  - Etc.
- 2014 IRUS-UK user survey:
  - 68% reported that IRUS-UK has improved statistical reporting
  - 66% reported that IRUS-UK saves time collecting statistics
  - 66% reported that IRUS-UK enables reporting previously unable to do
  - 83% hope to use IRUS-UK for benchmarking

# Contacts & Information

- If you wish to contact IRUS-UK:
  - [irus@mimas.ac.uk](mailto:irus@mimas.ac.uk)
  - [@IRUSNEWS](https://twitter.com/IRUSNEWS)
- Project web site:
  - <http://irus.mimas.ac.uk/>

*“The set up was quick and painless, which is always a delight!”*

*“Consistent collection of statistics without me having to do it!”*