



IRUS-UK



IRUS-UK: standardised institutional repository usage statistics

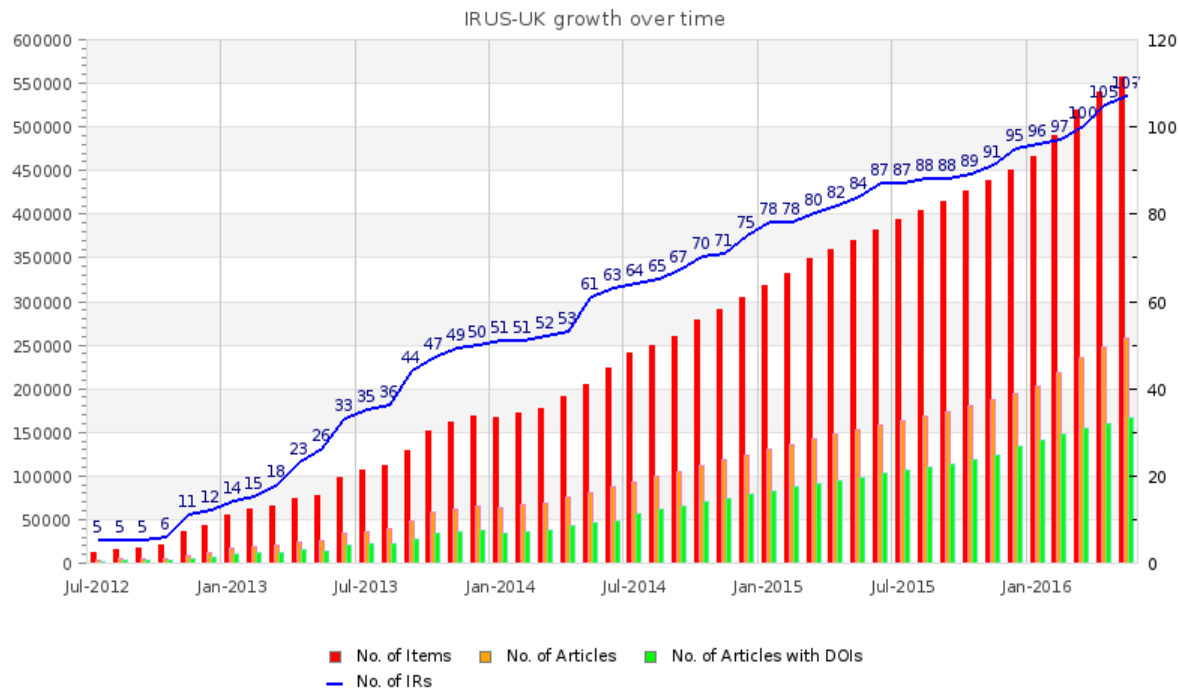
John Salter, White Rose University Consortium,
on behalf of IRUS-UK and COUNTER

Panel 5: Usage Statistics
Open Repositories 2016, Dublin, 16 June

What is IRUS-UK?

A national aggregation service for UK Institutional Repository Usage Statistics:

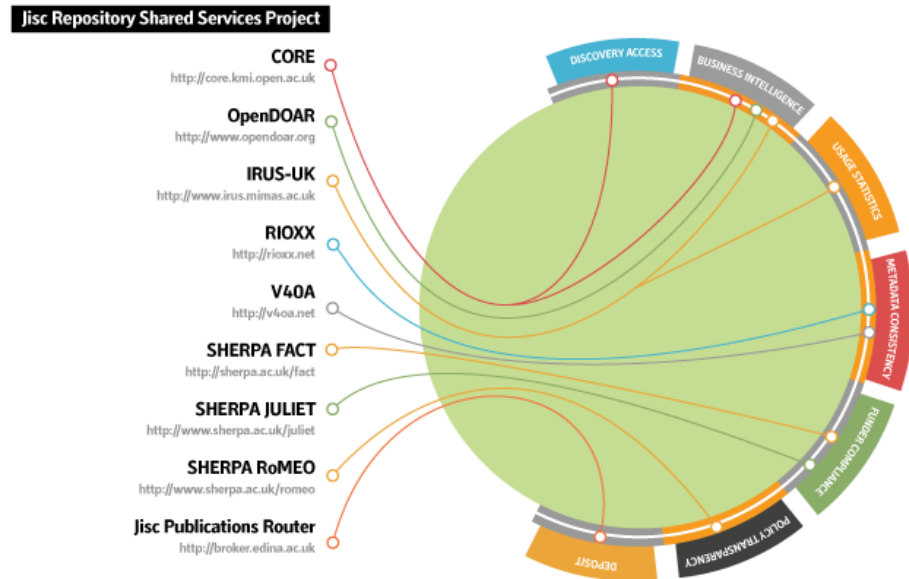
- Collects raw download data from UK IRs for *all item types* within repositories
- Processes raw data into COUNTER-compliant statistics - so produced on the same basis as 'traditional' scholarly publishers



Who is responsible for IRUS-UK?

- Funded by Jisc
- Team Members:
 - Jisc – Service Management & Host
 - Cranfield University – Service Development & Maintenance
 - Evidence Base, Birmingham City University – User Engagement & Evaluation

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



About COUNTER

- **Counting Online Usage of Networked Electronic Resources**
 - an international initiative serving librarians, publishers and intermediaries
 - setting standards that facilitate the recording and reporting of online usage statistics
 - consistent, credible and comparable
- **COUNTER Codes of Practice:**
 - [Release 4 of the COUNTER Code of Practice for e-Resources](#)
 - Books, databases , journals and multimedia content
 - [Release 1 of the COUNTER Code of Practice for Articles](#)
 - Journal articles
 - [Release 1 of the COUNTER Code of Practice for Usage Factors](#)
 - Alternative metric to the Journal Impact Factor
- **The Codes specify:**
 - How raw data should be processed into statistics
 - How statistics reports should be formatted and delivered

The Tracker Protocol

- To enable repositories to participate in IRUS-UK a small piece of code is added to repository software, which employs the ‘Tracker Protocol’
 - <http://www.irus.mimas.ac.uk/help/toolbox/TrackerProtocol-V3-2014-04-22.pdf>
 - Patches for DSpace (1.8.x, 3.x, 4.x) and Plug-ins for Eprints (3.2-3.3.x), implementation guidelines for Fedora
 - In development for PURE portal and figshare
- The Tracker Protocol was devised in collaboration with and is endorsed by COUNTER
- Gathers basic data for each raw download and sends to IRUS-UK server
 - Raw download data is stored in daily logfiles before being ingested into IRUS-UK database

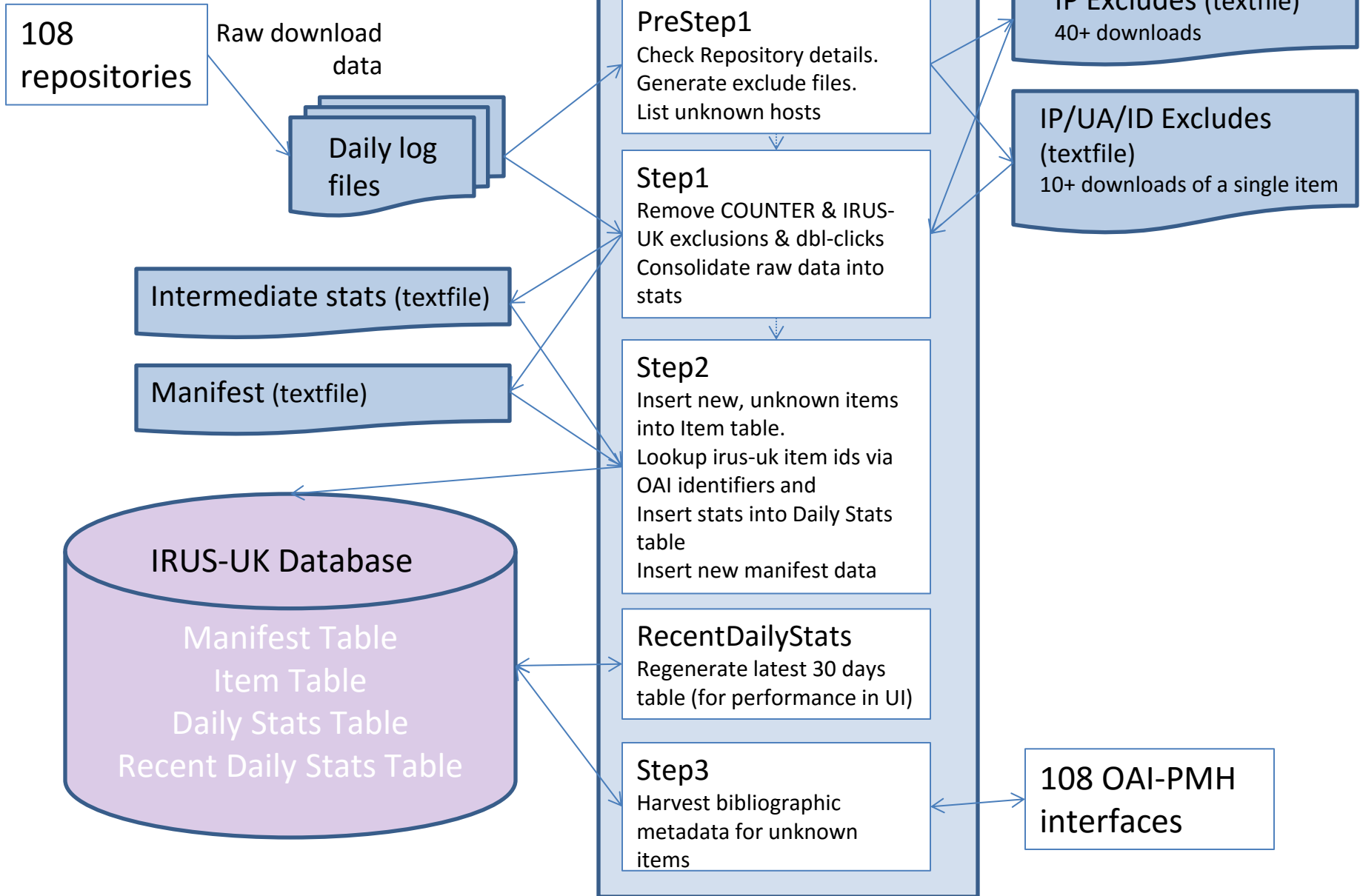
IRUS-UK ingest procedures

- The IRUS-UK daily ingest follows 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](#)
- Processes entries from recognised IRs
- Sorts and filters entries following COUNTER rules to remove robot entries using the [COUNTER User Agent Exclusion List](#)
 - regarded as a minimum requirement for excluding bots, etc.
- Filters entries using additional IRUS-UK filters
- Filters entries by applying the COUNTER 30 second double-click rule
- Consolidates raw usage data for each item into daily statistics
- At the end of each month, daily statistics are consolidated into monthly statistics

Robots & rogue usage – IRUS-UK current approach

- As stated, the COUNTER Robot Exclusion list is a minimum requirement for eliminating robots and rogue usage
 - It worked well for traditional scholarly publishers with resources behind a subscription barrier
 - As we move to an OA world and newer players – repositories – more needs to be done
- So in addition IRUS-UK:
 - excludes several IP ranges that are associated with some known robots and spiders, which don't identify themselves in their User Agent string
 - And based on several years of empirical evidence:
 - exclude downloads from IP addresses with 40+ downloads in a day
 - Exclude downloads from IP address/User Agent where a single item is downloaded 10+ times a day
 - Though we are aware that more can still be done
- More on robots later, but next a diagrammatic summary of the ingest . . .

IRUS-UK Daily Ingest



IRUS-UK Daily Files – example snippets

IP Excludes (textfile)

IP address	Downloads
139.222.81.231	205
14.139.186.178	52
143.159.5.143	42
144.76.153.198	1506
144.76.71.83	1546
148.251.133.219	386
148.251.54.44	1524
155.198.8.192	103
163.44.165.228	59
172.16.20.21	1301
173.44.129.103	47

IP/UA/ID Excludes (textfile)

IP address~Useragent~oai identifier	Downloads
49.51.8.24~Mozilla/5.0+(Windows+NT+6.1;+WOW64;+rv:31.0)+Gecko/20100101+Firefox/31.0~oai:sure.sunderland.ac.uk:3880	10
49.51.8.24~Mozilla/5.0+(Windows+NT+6.1;+WOW64;+rv:31.0)+Gecko/20100101+Firefox/31.0~oai:sure.sunderland.ac.uk:3904	32
49.51.8.24~Mozilla/5.0+(Windows+NT+6.1;+WOW64;+rv:31.0)+Gecko/20100101+Firefox/31.0~oai:sure.sunderland.ac.uk:4542	10
49.51.8.24~Mozilla/5.0+(Windows+NT+6.1;+WOW64;+rv:31.0)+Gecko/20100101+Firefox/31.0~oai:sure.sunderland.ac.uk:4575	25
49.51.8.24~Mozilla/5.0+(Windows+NT+6.1;+WOW64;+rv:31.0)+Gecko/20100101+Firefox/31.0~oai:sure.sunderland.ac.uk:4682	10
5.120.14.26~Mozilla/5.0+(compatible;+MSIE+10.0;+Windows+NT+6.2;+WOW64;+Trident/6.0)~oai:dspace.lib.cranfield.ac.uk:1826/8508	31
95.145.89.122~Mozilla/5.0+(Windows+NT+10.0;+WOW64;+Trident/7.0;+rv:11.0)+like+Gecko~oai:bradscholars.brad.ac.uk:10454/4308	27
95.149.102.145~Mozilla/5.0+(Macintosh;+Intel+Mac+OS+X)+Word/14.48.0~oai:open.ac.uk.OAI2:17219	12
98.139.14.250~AHC/2.0~oai:dspace.lib.cranfield.ac.uk:1826/3570	11
98.139.14.250~PSEARCH/2.0~oai:dspace.lib.cranfield.ac.uk:1826/3570	10
98.139.14.251~PSEARCH/2.0~oai:dspace.lib.cranfield.ac.uk:1826/3570	10

IRUS-UK Daily Files – example snippets

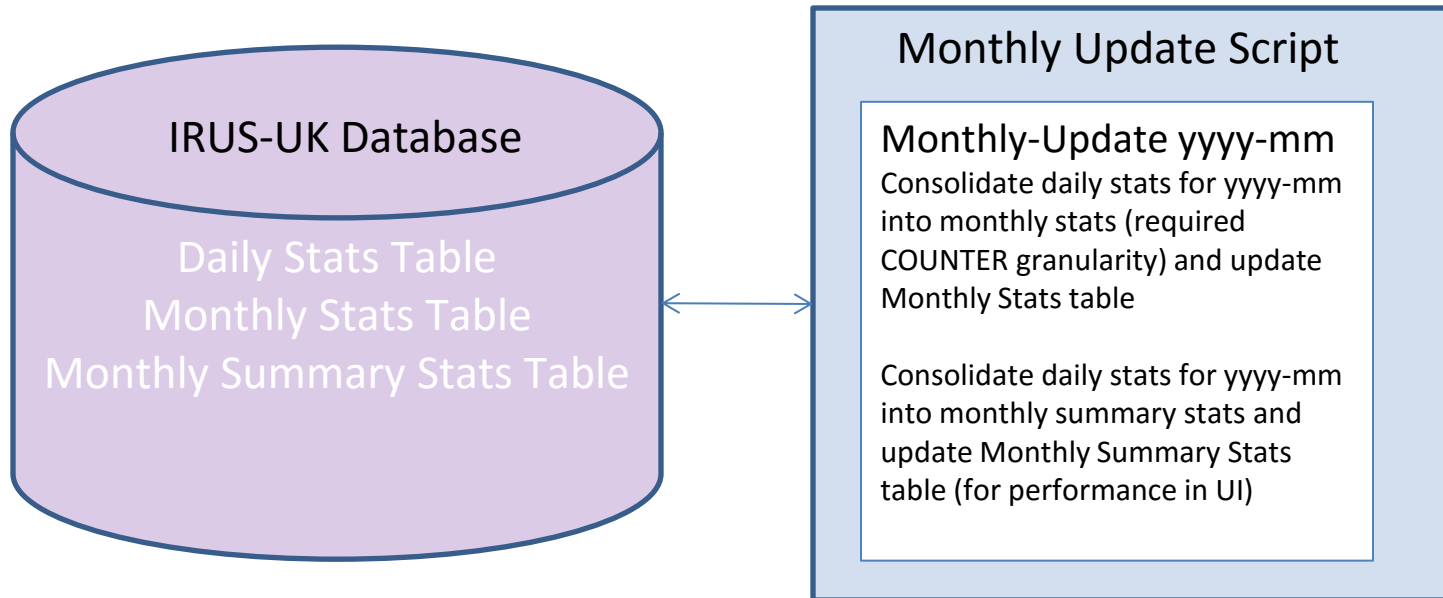
Manifest (textfile)

Logfile name	Repo id	Platform id	Raw data in	Counter exclusions	IRUS exclusions	Double clicks	Filtered data out
2016-06-01.txt	1	2	3463	1309	239	24	1891
2016-06-01.txt	2	2	3570	1966	318	9	1277
2016-06-01.txt	3	1	6352	3597	613	202	1940
2016-06-01.txt	4	2	3246	1363	634	45	1204
2016-06-01.txt	5	2	8589	5885	427	130	2147
2016-06-01.txt	6	2	6758	5805	180	28	745
2016-06-01.txt	7	2	12997	7639	929	90	4339

Intermediate stats (textfile)

Repo id	Platform id	OAI id	Date	Downloads
1	2	oai:eprints.bournemouth.ac.uk:9773	2016-06-01	1
1	2	oai:eprints.bournemouth.ac.uk:9808	2016-06-01	1
1	2	oai:eprints.bournemouth.ac.uk:9836	2016-06-01	1
1	2	oai:eprints.bournemouth.ac.uk:9849	2016-06-01	1
1	2	oai:eprints.bournemouth.ac.uk:990	2016-06-01	1
10	2	oai:gala.gre.ac.uk:10259	2016-06-01	4
10	2	oai:gala.gre.ac.uk:10506	2016-06-01	2
10	2	oai:gala.gre.ac.uk:10689	2016-06-01	3
10	2	oai:gala.gre.ac.uk:10708	2016-06-01	3
10	2	oai:gala.gre.ac.uk:10728	2016-06-01	1
10	2	oai:gala.gre.ac.uk:10738	2016-06-01	3

IRUS-UK Monthly Update



Viewing/retrieving IRUS-UK statistics

- Web User Interface - The IRUS-UK Portal
 - Access currently behind Shibboleth authentication/authorisation
 - Wide range of views, slicing and dicing stats from the IRUS-UK database
 - Reports available for download as TSV/CSV/Excel spreadsheet files
- SUSHI service
 - XML SOAP service
 - for more info see: <http://www.niso.org/workrooms/sushi/>
- SUSHI Lite API
 - Under development by the NISO SUSHI Lite Technical Report Working Group (http://www.niso.org/workrooms/sushi/sushi_lite/)
 - Allows retrieval of stats reports and snippets as JSON to be embedded into Repository (and other) web pages, or ingested into other services
 - Example client: http://www.irus.mimas.ac.uk/api/sushilite/v1_7/Client/

IRUS-UK: Home

IRUS-UK

Last update: 8th Jun 2016

[Home](#) [Statistics Views](#) [Statistics Reports](#) [Metadata Views](#) [Metadata Reports](#) [Admin](#)

[More search options](#)

Home

Overall Summary

These tables provide summaries of data in IRUS-UK including: number of participating repositories, number of items downloaded from these repositories since they joined, numbers of downloads (in total, to the end of the previous month, and during the current month) and also the latest date for which data is available in IRUS-UK.

Overall Summary

Headlines at a glance as of 8th Jun 2016	
Repositories	107
Items	565,511
Downloads up to May-2016	61,309,195
Downloads in Jun-2016, so far...	596,067
Total Downloads	61,905,262

Yearly summaries

Year	No. of Repositories	No. of Articles	Article Downloads	No. of Items	Total Downloads
2012	13	11,507	373,601	42,546	1,038,181
2013	50	65,390	3,321,848	167,691	9,036,143
2014	75	123,759	6,075,981	304,716	16,570,691
2015	95	193,491	8,154,275	450,082	22,373,006
2016	107	257,482	4,964,104	556,304	12,291,174

IRUS-UK: Repository Report 1

IRUS-UK

Last update: 8th Jun 2016

[Home](#) [Statistics Views](#) [Statistics Reports](#) [Metadata Views](#) [Metadata Reports](#) [Admin](#)


[More search options](#)

Repository Report 1 (RR1), Number of Successful Item Download Requests by Month and Repository

Repository Report 1 (RR1), Number of Successful Item Download Requests by Month and Repository

To whom it may concern

Period covered by Report: 2016-01-01 to 2016-05-31

Filters:

Date Run: 2016-06-10

Number of results: 107

Repository	Total	Jan-2016	Feb-2016	Mar-2016	Apr-2016	May-2016
Totals	12,291,174	2,294,939	2,374,001	2,677,154	2,548,657	2,396,423
Aberdeen University Research Archive (AURA)	71,496	13,888	17,154	13,748	16,600	10,106
Abertay Research Collections (ARC)	9,620	1,512	1,788	2,480	2,116	1,724
Aberystwyth University - CADAIR	58,128	10,224	11,218	13,059	12,789	10,838
Anglia Ruskin Research Online, ARRO	6,888	1,103	1,195	837	1,528	2,225
Apollo - University of Cambridge Repository	633	0	0	0	0	633
BIROn - Birkbeck Institutional Research Online	15,530	13,295	2,235	0	0	0
Bath Spa University - ResearchSPAcE	8,230	1,480	1,533	1,681	1,931	1,605
Bournemouth University BURO	205,813	35,246	39,563	31,737	44,237	55,030
Bradford Scholars	49,194	3,823	10,611	12,270	11,586	10,904



IRUS-UK: Item Report 1



[More search options](#)

Item Report 1 (IR1), Number of Successful Item Download Requests by Month and Repository Identifier

Item Report 1 (IR1), Number of Successful Item Download Requests by Month and Repository Identifier	
Period covered by Report:	2016-01-01 to 2016-05-31
Filters	Repository: White Rose Research Online;
Date Run	10/06/2016
Number of results	20,721

[Previous page](#)
001
002
003
004
005
006
007
008
009
010
 ...
 415
[Next page](#)

ItemURL	OAISourceID	Title	Author	Itemtype	Total	Jan-2016	Feb-2016	Mar-2016	Apr-2016	May-2016
Totals					491,856	96,526	98,318	104,101	103,281	89,630
http://eprints.whiterose.ac.uk/8904/	oai:eprints.whiterose.ac.uk:8904	"Who are you?" - Learning person specific classifiers from video"	Sivic, J.; Everingham, M.; Zisserman, A.	Conference Proceedings	19	5	7	3	1	3
http://eprints.whiterose.ac.uk/83849/	oai:eprints.whiterose.ac.uk:83849	"A canvas of endless extent": Granville Bantock, Robert Southey and the curse of Kehama project	Allis, MJ	Article	1	0	0	0	1	0
http://eprints.whiterose.ac.uk/85285/	oai:eprints.whiterose.ac.uk:85285	"A petrification of one's own humanity"? Nonattachment and ethics in yoga traditions	Burley, M	Article	107	18	26	34	10	19
http://eprints.whiterose.ac.uk/9295/	oai:eprints.whiterose.ac.uk:9295	"A wealth of knowledge": A survey of the employment experiences of older nurses and midwives in the NHS	Wray, J.; Apsland, J.; Gibson, H.; Stimpson, A.; Watson, R.	Article	10	3	0	3	3	1
		"An Impediment to Living Life".	Mackie, S.; Hughes							

IRUS-UK: Item Statistics

IRUS-UK

Last update: 9th Jul 2015

[Home](#) [Statistics Views](#) [Statistics Reports](#) [Metadata Views](#) [Metadata Reports](#)

[More search options](#)

Item Statistics

The Distance Between Mars and Venus: Measuring Global Sex Differences in Personality



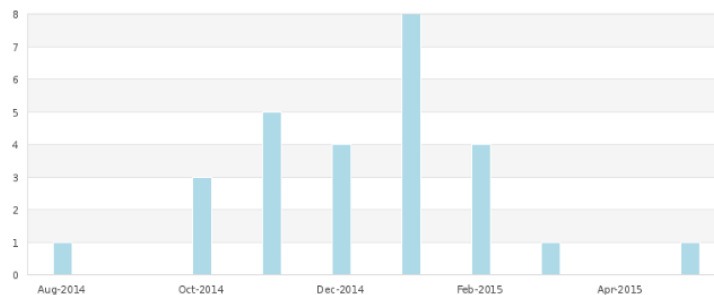
- Picked up by 3 news outlets
- Blogged by 8
- Tweeted by 172
- On 12 Facebook pages
- Mentioned in 12 Google+ posts
- Reddited by 9
- On 4 videos
- 122 readers on Mendeley
- 1 readers on CiteULike

[See more details](#)

Author(s)	Del Giudice, M., Booth, T. ; & Irwing, P. (2012).
IRUS item type	Article
Source item type	Journal article
URL	http://www.manchester.ac.uk/escholar/uk-ac-man-scw:151752
DOI	http://dx.doi.org/10.1371/journal.pone.0029265
Repository	Manchester eScholar
Platform	Fedora

Monthly downloads from Manchester eScholar (27)

Aug-2014	Sep-2014	Oct-2014	Nov-2014	Dec-2014	Jan-2015	Feb-2015	Mar-2015	Apr-2015	May-2015
1	0	3	5	4	8	4	1	0	1



What next – COUNTER and IRUS-UK?

- IRUS-UK follows the COUNTER Codes of Practice for its ingest . . .
- But reports we produce are not covered by the existing releases
 - So, they are COUNTER-like rather than official reports
- We're working with COUNTER and the COUNTER auditors to define requirements for additional reports to cater for
 - Repositories and other new services
 - An OA landscape
- COUNTER and the NISO SUSHI Committee are developing a new, comprehensive R5 Code of Practice
 - Simplifying and harmonising reporting requirements
 - Fewer reports with increased flexibility
 - Integrating requirements for repositories
 - Incorporating procedures and statistics for Research Data

What next – COUNTER and robots?

- Members of this panel are in broad agreement:
 - Static lists of IPs and UAs, while useful to a certain extent, are not sufficient to eliminate robots and rogue usage
 - A better approach is to look at dynamic filtering based on observed behaviour of IPs and UAs
- Panel members, COUNTER and others interested parties are
 - Working to identify the most practical and generally applicable robot detection techniques
 - Attempting to determine and agree various thresholds for elimination of non-human usage (with a justifiable statistical basis)
- COUNTER intends to incorporate the results of this work into the R5 Code of Practice

What next – COUNTER and robots?

- Members of this panel are in broad agreement:
 - Static lists of IPs and UAs, while useful to a certain extent, are not sufficient to eliminate robots and rogue usage
 - A better approach is to look at dynamic filtering based on observed behaviour of IPs and UAs
- Panel members, COUNTER and others interested parties are
 - Working to identify the most practical and generally applicable robot detection techniques
 - Attempting to determine and agree various thresholds for elimination of non-human usage (with a justifiable statistical basis)
- COUNTER intends to incorporate the results of this work into the R5 Code of Practice

Final caveat!

At certain levels of activity machine/non-genuine usage is practically indistinguishable from genuine human activity.

Contacts & Information

IRUS-UK web site:

- <http://irus.mimas.ac.uk/>

Contact IRUS-UK:

- irus@mimas.ac.uk
- [@IRUSNEWS](#)

COUNTER web site:

- <http://www.projectcounter.org/>

Contact COUNTER:

- <http://www.projectcounter.org/contact.html>
- [@projectCOUNTER](#)