

Conference Review: 11th International Conference On Open Repositories (#Or2016), Dublin, Ireland, 13th - 16th June 2016

Nick Sheppard – Research Services Advisor

Leeds Beckett University

“a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital material created by the institution and its community members.” (Lynch 2003)

Introduction

The very first International Conference on Open Repositories was convened in Sydney Australia in January 2006, two years on from the definition above and five and a half since the release of EPrints open source repository software in July 2000.

Clifford Lynch's definition certainly still applies. In the UK and elsewhere, however, there is a new imperative to optimise the use of repositories in the form of mandates for open access from research funding bodies. Notably the HEFCE policy in the UK that came into effect in April 2016 requires authors to upload their accepted manuscripts to a repository on acceptance for publication in order to be eligible for the Research Excellence Framework or 'REF', the national research assessment exercise (HEFCE 2014). The 2016 conference addressed the tension between these two facets of open access with 205 presentations, posters and workshops over 4 days.

With so many parallel sessions, social media played a significant role at the conference, enabling a level of virtual participation across the whole conference even for those not physically present. Inevitably this review will focus on selected sessions that its author attended in person while recognising that social media, Twitter in particular, is a powerful tool that can highlight significant conference themes.

Unlike past conferences where the specialised Interest Groups for the three main repository

software platforms (EPrints, DSpace, FEDORA) have been run as separate tracks to the main event, in Dublin they were integrated into the main programme. This approach worked well and better reflects the close relationships in the repository community between open access practitioners and software developers, between technology and policy, university libraries and their infrastructure.

The conference began with a warm welcome by Librarian and College Archivist, Helen Shenton who urged delegates to visit Trinity Library, an 18th Century 'trusted repository' and home to the Book Of Kells. Well over a thousand years old, this 9th Century illuminated manuscript puts into perspective our more recent challenges of digital preservation.

The keynote address from Laura Czerniewicz, Director for the Centre for Innovation in Learning and Teaching at the University of Cape Town in South Africa, explored another major challenge of our own era.

In *Knowledge Inequalities: A Marginal View of the Digital Landscape* Czerniewicz used a series of powerful data visualisations to illustrate the privileges of scholars in the world's most developed countries compared to the 'global south', a region that is partly geographical but also imaginary (Czerniewicz, L., 2016). The point was reinforced by the demographic breakdown of this international conference where, of almost 500 registered delegates, 448 were from Europe and

North America, 11 from Oceania, 9 from Asia and just 3 each from South America and Africa.

Czerniewicz went on to discuss the new opportunities of digital technology to “collapse distance” and “create possibilities for knowledge production and sharing” but stressed that while digital technology affords open, it is not synonymous with it and digital technologies are increasingly implemented to control and restrict access. Several commentators noted the irony of a major sponsor of the event being the academic publisher Elsevier who have been widely criticised for their business practices, seen by many as inimical to open dissemination (COAR 2015).

Finally Czerniewicz discussed search engines in two thought provoking case studies that demonstrated how these technologies are far from neutral and inevitably reflect global societal disparities. Her overall message was that open access is indeed a good thing but challenged an implicit technological determinism that it will inevitably lead to a ‘global village’. She commented that visibility is a “requirement for participation” and online dissemination “adds complexity to the abiding global inequalities of power and resources” (Czerniewicz, L., 2016).

Pre-conference workshops (Day 1)

Sessions attended:

Workshop 15: The Future of Open Access Support and Intelligence Services - Beyond RoMEO, JULIET, FACT and OpenDOAR - Monday, 13/Jun/2016: 4:00pm - 6:00pm

Monday was dedicated to a series of pre-conference workshops which included a workshop on Persistent Identifiers (PIDs) - ORCID records for researchers, DataCite Digital Object Identifiers (DOIs) for datasets and Crossref DOIs for publications. PIDs were a common theme throughout the conference with ample opportunity to explore how they are being embedded into infrastructure and workflow.

Probably the workshop that generated the most interest on social media during the morning was on text and data mining (TDM) led by colleagues from the CORE project based at the Open University in the UK and supported as part of the Jisc portfolio of open access services. CORE has aggregated a large corpus of full-text papers from

repositories globally which provides a unique data-set for some exciting developments. The CORE team presented their work throughout the conference including a paper exploring ‘semantometrics’, a potentially seminal technology to leverage TDM for full text-based research evaluation for open repositories. The team also won best poster award for their poster on *Integration of IRUS-UK statistics in the CORE Repositories Dashboard* (Pearce, S. and Pontika, N., 2016)

In the afternoon there was another popular workshop run by the Sherpa Services team based at the Centre for Research Communications at the University of Nottingham.

The Future of Open Access Support and Intelligence Services - Beyond RoMEO, JULIET, FACT and OpenDOAR was particularly welcome from a UK perspective and acknowledged the need to respond to an increasingly complex technical and policy landscape. Adam Field gave an overview of technical developments towards a “single, unified, extensible data model” across all services that will enable more efficient and bespoke data via an API.

Break-out groups discussed ‘user stories’ that were collected by the Sherpa team for later analysis. A beta release of Sherpa RoMEO built on the new API is planned for October 2016.

The immense value of the Sherpa suite of services for open access cannot be overstated, though there was some discussion that a great deal of time and effort is expended on ensuring policy compliance with ‘green’ open access mandates. Meanwhile commercial publishers are potentially moving the sector towards a paid ‘gold’ model, with the increased prevalence of ‘hybrid’ business models for example. Ultimately, however, Sherpa must respond to the contemporary requirements of their user base.

Day 2

Sessions attended:

Papers 1: Cooperative approaches for research data - Tuesday, 14/Jun/2016: 11:00am - 12:30pm. Session Chair: Anthony Ross-Hellauer

EPrints IG 1: Opening Variations - Tuesday, 14/Jun/2016: 2:00pm - 3:30pm. Session Chair: Tomasz Neugebauer

Papers 4: Metrics, Assessment, and Impact - Papers 4: Metrics, Assessment, and Impact - Tuesday, 14/Jun/2016: 4:00pm - 5:30pm. Session Chair: Mark Hahnel

In the morning I attended *Papers 1: Cooperative approaches for research data* while keeping a watchful eye on Twitter discussions emanating from other sessions. In particular *Developer Track 1: Platforms and Analytics* included two papers from the CORE team and though I did not attend the session, I was able to gain some insight from Petr Knoth of CORE later in the conference on a project that is leveraging freely available datasets to evidence impact for a given university based on papers in its repository (Knoth, P. et al., 2016).

With six submissions across the conference, the work of CORE is a central theme in this review and on one level illustrates the power of open data in the hands of talented developers.

In recent years, alongside open access to research papers, there has been an increased emphasis on the underlying research data and its management. Research Data Management (RDM) was correspondingly a major theme of this conference and made the data-driven keynote all the more appropriate.

A Consortial Model for Research Data Services Using Dataverse brought to mind the Research Data Shared Service project from Jisc in the UK, presented as a poster at the conference (Kaye, J., 2016) and due for beta release in 2017. It was fascinating to hear experience from the USA, from a single state larger in area than the whole of the UK, albeit with rather fewer universities and colleges.

Increasing the Availability and Reuse of Earth Science Ontologies Using a Repository was a more specialised paper that nevertheless highlighted important issues for the development of RDM services across disciplines which inevitably require different controlled vocabularies for their practitioners. These issues cropped up again in *Papers 9: What's in here anyway? Metadata and identifiers*, particularly in the context of sustainable development and community uptake which, as I shall discuss, was a key theme of Paul Walk's paper on the RIOXX Application Profile.

The third paper *Research Data Management in Austria – A manageable task for Austrian research institutions?* was the most relevant to our current

work at Leeds Beckett University and presented the results of a survey conducted across Austria in 2015. It reflected many similar issues to current experience in the UK with concerns around “data types and formats, data archiving, backup and loss, ethical and legal aspects, accessibility and reuse as well as infrastructure and services”. The main expectations of academics emerged as “qualified personnel, guidelines and policy” and the project emphasised that “solid research data management is the foundation for cooperative, open research and thus for their comprehensibility and verifiability.” (Ganguly, R., 2016). One question arising from the audience was just what academics might understand by reusable data with the suspicion that it might merely be available, without proper consideration of type, format and appropriate metadata.

In *EPrints IG 1: Opening Variations* Les Carr gave an overview of EPrints 3.4, now available in beta for test installation and which comes in 3 ‘flavours’ optimised to support publications, research data or educational resources respectively.

Other customisations include a ‘tweepository’ to run and archive searches on twitter for later analysis, to inform sociological research for example, particularly in the context of the ‘new’ discipline of ‘Web Science’, pioneered at the University of Southampton and of which Les Carr is Professor.

Kelly Terrell, also of Southampton, spoke on *Repositories for open education: a reflection and look forward for EPrints* which was of particular interest given a history of managing Open Educational Resources or ‘OER’ at my own institution.

EdShare has long been an exemplar of a successful OER repository and it was reassuring to hear that it continues to be successful institutionally, in contrast to Jorum, the national OER repository recently retired by Jisc (Jisc 2015). EdShare underpins the open education ‘flavour’ of EPrints 3.4, and was influenced by social media sharing sites while placing particular emphasis on guiding users to apply appropriate Creative Commons licences.

Creative Commons is also gaining popularity in the contexts of OA research papers and research data. It is useful to consider what lessons from

open education and ‘learning object’ repositories might be applicable to RDM which, like OER, perhaps “require more emphasis on well described content which can be easily navigated and discovered by the diverse audience they are released for”. (Terrell, K., 2016)

Engaging Arts Researchers: EPrints Research Repositories in a Post-Kultur Landscape considered the fact that institutional repositories can often be ill-equipped to showcase the diverse range of atypical research outputs associated with the arts and related disciplines. These can comprise potentially complex file types, once again bearing comparison with RDM and OER. The ‘Kultur’ plugin provides a metadata profile adapted to practice-based outputs and “preview-oriented abstract pages and slideshows for multimedia EPrints repositories” (Brody, T., et al. 2012). These are issues that we are currently considering at Leeds Beckett University with colleagues from our faculty of Arts, Environment & Technology.

Institutional repositories have long been recognised as an important component of the scholarly and institutional infrastructure and thanks to HEFCE have now become truly business critical in UK Higher Education.

As the most widely used repository software in the UK, it is crucial that EPrints continues to develop. There is clearly a vibrant community of developers working on the software, developing software extensions or ‘plugins’ which can be added from the EPrints Bazaar, and it was reassuring to learn about ongoing development of the EPrints core at Southampton. Looking ahead to EPrints v.4.0, Carr asked what a modern repository should look like; “Repositories of the Future” was the subject of a panel discussion on the final day of the conference.

In the afternoon I attended *Papers 4: Metrics, Assessment, and Impact* which included the fascinating paper from the CORE team mentioned previously *Exploring Semantometrics: full text-based research evaluation for open repositories*.

It is difficult to summarise semantometrics and it is worth reading the full report which introduces the hypothesis that the added value of a given publication (p) can be estimated based on the “semantic distance” from publications cited

by p to publications citing p (Herrmannova, D., and Knoth, P., 2016). In other words, computer analysis and comparison of papers’ content (the actual words used) can provide an insight into its relationship with the broader research corpus. Such an approach, the authors claim, is superior to both established citation-based metrics and newer article level measures (e.g. repository download count and ‘altmetric’ scores) which are merely proxies for quality; all such measures are dependent on the number of interactions in a scholarly communication network and can be easily manipulated or ‘gamed’.

It is suggested that possibilities of semantometrics include detecting the quality of research practices and analysing the context and sentiment of a citation. Semantometrics therefore offers the possibility that an evolved metric based on the technology would encourage a focus on quality rather than quantity and that is comparable across disciplines.

The crucial aspect in the context of open repositories is the potential of TDM which is still largely untapped. Though copyright law has allowed for TDM since 2014, copyright ownership by commercial publishers and associated restricted access still presents a barrier that might be mitigated by a move to gold OA under a Creative Commons Attribution licence (CC-BY). The potential cost of a fully gold scholarly communication ecosystem, however, is considerable. Moreover there is ongoing uncertainty around copyright ownership of accepted manuscripts that may carry legal implications for the CORE model of mining aggregated green OA papers. One initiative in particular, led by Imperial College London and discussed by Torsten Reimer on Day 3 is seeking to address this with the UK Scholarly Communications Licence (UK-SCL) discussed in more detail below.

Scaling Usage Statistics across Repositories as an OpenAIRE Analytics Service focussed on article level metrics from repositories, views and downloads, to complement more traditional bibliometrics and was comparable to the work of IRUS-UK to provide COUNTER compliant download data to UK based repositories.

Usage statistics is an area that we have explored at Leeds Beckett University, contributing an

IRUS-UK case study for example and were discussed in more detail, including limitations, in a panel session on Day 4.

Day 3

Sessions attended:

24x7-2: Second 24x7 session - Wednesday, 15/Jun/2016: 9:00am - 10:30am. Session Chair: Jon Stroop

Papers 9: What's in here anyway? Metadata and identifiers - Wednesday, 15/Jun/2016: 11:00am - 12:30pm. Session Chair: Jenn Riley

EPrints IG 2: Research Management - Wednesday, 15/Jun/2016: 2:00pm - 3:30pm. Session Chair: Tomasz Neugebauer

Papers 12: We know what you do: Integrations for Faculty Research Systems/CRIS - Wednesday, 15/Jun/2016: 4:00pm - 5:30pm. Session Chair: David Minor

Day 3 began with ten seven minute presentations covering everything from linked data (*Linked Data for Libraries (LD4L): Data, Tools, and Discovery*) to long-term storage of research data (*Developing a Data Vault*) to an embedded repository brand at the University of Glasgow (*Let there be (En)Light(en): Building a repository brand at the University of Glasgow*). Glasgow's is a mature repository many look towards for inspiration and this was a taster to a couple of more in depth papers from the Enlighten team later in the day.

First came *Papers 9: What's in here anyway? Metadata and identifiers* and one of the most important papers of the day, at least from a UK perspective, by Paul Walk of EDINA. *RIOXX: a Modern Metadata Application Profile* focussed on the requirement for institutions to supply metadata about their research projects, including open access status, in unambiguous terms. Existing metadata profiles were found not to be suitable, leading to the development of RIOXX, a list of key metadata fields that can be added to a

repository by means of a plugin. The project utilised innovative development approaches to maximise community engagement and in its first year RIOXX has been adopted by more than 40 UK institutional repositories.

N.B. very few of these repositories are currently returning valid RIOXX. There is an important caveat in that there may be an issue in the way EPrints returns records from an OAI-PMH request. This is a somewhat technical point but which has considerable import for a widely used open source platform that is relied upon by over 100 institutions in the UK to meet their funder obligations for open access¹.

EPrints IG 2: Research Management included 2 papers from the Enlighten team at the University of Glasgow with Michael Eadie providing insight to how they have effectively built a fully-featured 'CRIS' around EPrints and William Nixon looking ahead to the next REF with particular focus on the EPrints REF Compliance Checker plugin.

The term CRIS (Current Research Information System) is typically associated with commercial platforms like Pure, Converis and Symplectic Elements owned by Elsevier, Thomson Reuters and Digital Science respectively. Enlighten demonstrates how, with sufficient development, many of the more sophisticated features associated with a CRIS can be built around an open source platform; this is important in the light of concerns around the potential monopoly of data about researchers and their research by commercial companies. Features of Enlighten include publications and awards management, research analytics, reporting and assessment, evidence of impact and esteem. It is integrated with various institutional systems including HR databases and pushes information out to public facing websites such as Staff Profile pages (Eadie, M., 2016).

William Nixon began by emphasising the "Byzantine" open access landscape in the UK which requires vast amounts of support and technical development. The REF Compliance

¹ For full discussion see

<http://www.rioxx.net/2015/11/25/rioxx-and-oai-pmh/>

Checker is designed to facilitate the process of establishing whether an article is eligible for REF under the terms of HEFCE's policy and integrates with the RIOXX2 and Dates, Dates, Dates plugins for EPrints, further illustrating the immense value of a flexible open source platform that can be customised by its community of technical developers.

The late afternoon session focussed on CRIS integration and I was particularly interested to see the broad range of data sources integrated into Symplectic Elements at Virginia Tech in the US which could inform our own institutional challenges using similar software. Their linked repository is DSpace rather than EPrints and also incorporates a 'semantic layer' with VIVO, the "open source software and an ontology for representing scholarship" (VIVO, n.d.).

The key paper of the session for me, however, and arguably pivotal for the future of green open access in the UK was presented by Torsten Reimer as part of his presentation *Imperial College London - journey to open scholarship*.

Imperial College also use Symplectic Elements linked to a DSpace repository and Torsten described the evolution of policy and infrastructure at the college since 2012, including an RDM policy. Torsten highlighted a study that found that papers published open access in 2011 were cited twice as much as paywalled papers. The potential silver bullet, however, is the UK Scholarly Communications Licence (UK-SCL), inspired by the OA policy at Harvard and adapted to a UK legal and policy context. The aim is to enable academics to grant their university a non-exclusive licence to their scholarly output which would take precedence over any subsequent contract with a publisher and enable the University to make accepted manuscripts available under Creative Commons. The project was initiated by Torsten and his colleague Chris Banks, Director of Library Services at Imperial College, who writes for The UK Serials Group (UKSG) about the impetus at Imperial College to develop a policy "which might transcend the myriad funder and publisher policies, and which might include a licence that would enable single-step compliance with multiple funder and publisher policies whilst preserving academics' freedom to publish where most appropriate." (Banks, C., 2016)

Day 4

Sessions attended:

EPrints IG 3: Innovation and Use - Thursday, 16/Jun/2016: 9:00am - 10:30am. Session Chair: Tomasz Neugebauer

Panel 5: Usage Statistics – Thursday, 16/Jun/2016: 11:00am - 12:30pm. Session Chair: Stephanie Taylor

I delivered our paper as part of *EPrints IG 3: Innovation and Use*. The other two papers were *Discovery of Data Gems in EPrints Repositories* from the University of Zurich which presented interesting methods to visualise data in the EPrints' native interface and *IST DataRep - implementing an institutional repository* which was a useful presentation exploring the development of a data repository using EPrints at a small institution in Austria.

Our paper examined how an increasingly integrated scholarly communications infrastructure, locally and on the wider web, can be leveraged to support open dissemination through a variety of complementary channels. Leeds Beckett University has sought to integrate and display data from third party services - Scopus, Web of Science, PubMed, Twitter, YouTube, IRUS-UK, altmetric.com, ORCID - via a suite of software including Symplectic Elements, EPrints, Sitecore and LibGuide Content Management Systems. The paper built on a workshop at IWMW2013 which explored the challenges of integrating the institutional web site and the institutional repository (Sheppard, N. and Taylor, S., 2013) as well as exploring how we utilise social media to disseminate institutional research.

I also attended *Panel 5: Usage Statistics* before setting off for home and following the afternoon via social media and free Wi-Fi at Dublin airport, including the closing keynote from Dr Rufus Pollock.

According to data presented by Joseph Greene from University College Dublin, 85% of unfiltered downloads come from web 'robots'. This is not necessarily undesirable and robots are essential to ensure that a repository or website is properly indexed and discoverable via the major search engines. Nevertheless, 'credible' download data to

promote the benefits of repositories and Open Access is tacitly understood to mean ‘real’ downloads by real people and IRUS-UK follows the COUNTER Code of practice to “facilitate the recording, exchange and interpretation of online usage data [...] that are consistent, credible and compatible.” (COUNTER, n.d.).

One provocative contribution from the floor was that “web traffic analysis is the worst thing that ever happened on the internet” and it is difficult to argue that raw numbers are particularly meaningful. Academia is as vulnerable to the phenomenon of ‘click-bait’ as any other area of online discourse which is a criticism that has arisen in the related area of altmetrics. Some commentators are dismissive with Colquhoun and Plested discussing “Why you should ignore altmetrics and other bibliometric nightmares” in their blog-post (Colquhoun, D. and Plested, A., 2014.)

Like any debate, the value and drawbacks of usage statistics is non-binary and in July 2015 an independent report into the role of metrics in research assessment emphasised that repositories are amongst “a range of websites hosting free general scholarly databases [...] some of which form new sources of citation or usage data” (Wilson, J. et al., 2015).

Significantly, altmetrics offer the possibility of visualising the scholarly network which can provide context for repository downloads. Data from altmetric.com is available for individual records from IRUS-UK and our paper included a case study demonstrating that repository downloads recorded by IRUS-UK were directly associated with dissemination to an established research network via Twitter (Sheppard et al., 2016).

The other panel session *Panel 6: Repositories of the Future* emphasised that while repository systems are nearly ubiquitous in Higher education, they are still using technologies and protocols from the early days of the web. The Confederation of Open Access Repositories used the session to launch a technical working group to “identify the architectures, functionalities and technologies required for next generation repositories and to make recommendations about how these can be adopted in the context of repository platforms” (Rodrigues, E. et al., 2016).

One theme that emerged was that commercial publishers are increasingly moving into the repository space and it is difficult for repositories to compete with commercial software in terms of usability. Paul Walk emphasised that the distributed nature of repositories is a vital safeguard against the monopoly control of research. There were also parallels with the discussion above with a recognition that while resources (papers, datasets) are obviously crucial, the researcher network is also important and already being leveraged in the form of open peer review as well as social media.

The ‘Ideas Challenge’, always a fertile strand at Open Repositories, this year had invited developers to team up with colleagues to think of an idea to make researchers lives easier including underlying technology. Ten ideas were presented just before the final keynote and ranged from the sublime to the ridiculous. The winning entry was the Magic Place™, cheekily trademarked by the team which presented a potential ‘one drop’ solution to enable academics to transfer metadata and linked file(s) to their “faculty profile, in the repository, in the CRIS system, through the publication office, research data channels, and wherever else it is needed - consistently, magically, wonderfully.” (Triggs, G., et al., 2016)

The closing keynote from Rufus Pollock, founder and President of the Open Knowledge Foundation. *Making an Open Information Age* argued that control and access to information is the single biggest social & political issue of the 21st century and nicely bookended Czerniewicz’s opening address by emphasising that knowledge inequality equates with material inequality and that “knowledge is power, openness is empowerment” (Pollock, R., 2016)

Conclusion

Open Repositories is ostensibly about scholarly infrastructure and the technology underpinning a web-based network of research outputs. However, that network itself depends on people - researchers, publishers, developers and repository managers.

A significant theme of the conference was the perceived threat of commercial interests to the ‘open’ space, discussed by both Czerniewicz and Pollock in their respective keynotes.

As open access becomes the new norm in scholarly communication, and as traditional subscription based business models are eroded, it is inevitable that commercial publishers will aim to diversify, develop and acquire underlying technology in order to capitalise on the new reality. Nevertheless, publishers are but one player in the developing landscape.

Open Repositories highlights a significant portion of the broader network, notable for their enthusiasm, expertise and willingness to share. And that can only be a good thing for the future of scholarly communication.

Open Repositories 2017

The 12th International Conference on Open Repositories will be jointly hosted by The University of Queensland (UQ), Queensland University of Technology (QUT) and Griffith University in Brisbane, Australia, 26th – 30th June 2017.

Glossary

Altmetrics – “alternative metrics” are a variety of article level metrics derived from download counts and social media activity

API – Application Programming Interface. Set of protocols to enable data transfer between computer systems/software

Bibliometrics - quantitative measures used to assess research output i.e. publication and citation data analysis

COUNTER – a ratified standard that enables information professionals to count the use of electronic resources.

Creative Commons – free copyright licenses provide a standardised way to give permission to share and use creative work

CRIS – Current Research Information System. Comprehensive software system to manage the full range of research information including staff, publications, projects and proposals, post-graduate research, impact, ethics and key performance indicators

DSpace - open source repository software from Massachusetts Institute of Technology (MIT)

EPrints – open source repository software from the School of Electronics and Computer Science at the University of Southampton

EPrints Bazaar - provides ‘plug-ins’, software extensions for the EPrints repository platform developed by members of the EPrints community

FEDORA - Flexible and Extensible Digital Object and Repository Architecture

HEFCE – The Higher Education Funding Council of England has been responsible for the distribution of funding to universities and Colleges of Higher and Further Education in England since 1992.

IRUS-UK - enables UK Institutional Repositories (IRs) to share and expose statistics based on the COUNTER standard

Metadata – descriptive data and information to describe and contextualise other data

Open Educational Resources (OER) – openly licensed digital resources for teaching, learning, and assessment

ORCID - a persistent digital identifier to disambiguate researchers. ORCID is increasingly

integrated into the scholarly communication infrastructure to support automated linkages between researchers and their professional activities

PID – Persistent Identifier. A long-lasting reference to a document, file, web page, or other object usually used in the context of digital objects that are accessible over the Internet

PubMed - PubMed comprises more than 23 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites

REF – Research Excellence Framework. The regular national research assessment exercise in the UK overseen by HEFCE

Research Data Management (RDM) - the organisation of data, from its entry to the research cycle through to dissemination and archiving of valuable results

RIOXX – a set of metadata fields designed to help repositories comply with the RCUK policy on open access

Scopus - the world's largest abstract and citation database of peer-reviewed literature with more than 2,1000 titles from 5,000 international publishers and 5.5 million conference papers going back as far as 1823

Semantometrics - a new experimental class of metrics for evaluating research based on the premise that full-text is needed to assess the value of a publication

TDM – text and data mining. Refers to the process of deriving high-quality information from text

VIVO – open source software and an ontology for representing scholarship that supports recording, editing, searching, browsing, and visualizing scholarly activity according to the principles of the semantic web

Web of Science - multidisciplinary content covering over 12,000 journals worldwide, including Open Access journals and over 150,000 conference proceedings. Includes current and retrospective coverage in the sciences, social sciences, arts, and humanities, with coverage to 1900

References

- About VIVO (n.d.) [online] Available from: <<http://www.vivoweb.org/info/about-vivo>>
- Banks, C., (2016). Focusing upstream: supporting scholarly communication by academics. *Insights*. 29(1), pp.37-44. DOI: <http://doi.org/10.1629/uksg.292>
- Bauer, B., Ferus, A., Gorraiz, J., Gründhammer, V., Gumpenberger, C., Maly, N., Mühlegger, J.M., Preza, J.L., Sánchez, S.B., Schmidt, N., Steineder, C., Frasnelli, E., Seissl, M., Ernst, K., Ganguly, R., Meyer, T., Thorp, J. (2015) Researchers and Their Data. Results of an Austrian Survey – Report Available from: <https://zenodo.org/record/34005#.V3Zrw_krKHu>
- Brody, T., Gray, A., Tarrant, D., Francois, S. and Miles-Board, T. (2012) Kultur EPrints Extension [online] Available from: <<http://files.eprints.org/773/>>
- COAR (2015) Statement against Elsevier's sharing policy [online] Available from: <<https://www.coar-repositories.org/activities/advocacy-leadership/statements-and-guidelines/petition-against-elseviers-sharing-policy/>>
- Colquhoun, D. and Pledsted, A. (2014). Why you should ignore altmetrics and other bibliometric nightmares. *DC's Improbable Science* [online blog] Available online: <<http://www.dscience.net/2014/01/16/why-you-should-ignore-altmetrics-and-other-bibliometric-nightmares/>>
- COUNTER Code of Practice (n.d.) [online] Available from: <<https://www.projectcounter.org/code-of-practice-sections/general-information/>>
- Czerniewicz, L. (2016) Knowledge Inequalities: A Marginal View of the Digital Landscape. In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland. (keynote)
- Available from: <http://www.slideshare.net/laura_Cz/laura-czerniewicz-open-repositories-conference-2016-dublin>
- Eadie, M. (2016) CRISPrints: can a repository be a CRIS? In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland.
- Ganguly, R. (2016) Research Data Management in Austria – A manageable task for Austrian research institutions? In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland.
- HEFCE (2014) Policy for open access in the post-2014 Research Excellence Framework: Updated July 2015 [online] Available from: <<http://www.hefce.ac.uk/pubs/year/2014/201407/>>
- Herrmannova, D., and Knoth, P., (2016) Towards full-text based research metrics: Exploring semantometrics – Report of experiments Available from: <<http://repository.jisc.ac.uk/6376/>>
- Jisc (2015) Jisc to retire Jorum and refresh its open educational resources offer [online] Available from: <<https://jisc.ac.uk/news/jisc-to-retire-jorum-and-refresh-its-open-educational-resources-offer-24-jun-2015>>
- Kaye, J. (2016) Research Data Shared Service – OR2016. Research Data and Related Topics [online blog] Available from: <<https://researchdata.jiscinvolve.org/wp/2016/06/14/jisc-research-data-shared-service-or2016/>>
- Kerrell, T. (2016) Repositories for open education: a reflection and look forward for EPrints. In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland.
- Knoth, P., Cancellieri, M. and Herrmannova, D. (2016) Oxford vs Cambridge Contest: Collecting Open Research Evaluation Metrics for University Ranking. In: 11th International Conference on

Open Repositories, 13-16 June 2016, Dublin, Ireland

Lynch, C. (2003) ARL: A Bimonthly Report, no. 226, Association of Research Libraries [online] Available from: <<http://old.arl.org/resources/pubs/br/br226/br226ir~print.shtml>>

Pearce, S. and Pontika, N. (2016) Integration of the IRUS-UK Statistics in the CORE Repositories Dashboard (Poster) In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland. Available from: <https://figshare.com/articles/Integration_of_the_IR_US-UK_Statistics_in_the_CORE_Repositories_Dashboard/3407548>

Pollock, R. (2016) Making an Open Information Age In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland.

Rodrigues, E., Walk, P. and Shearer, K. (2016) Next generation repositories: building the repository of the future In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland.

Triggs, G., Cramer, T., Simpson, J., Kuncar, J. and Conlon, M. (2016) The Magic Place™ In: 11th International Conference on Open Repositories, 13-16 June 2016, Dublin, Ireland. Available from: <<http://or2016.net/ideas-challenge-results/>>

Wilsdon, J., et al. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. DOI: 10.13140/RG.2.1.4929.1363 Available from: <<http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title.104463.en.html>>