

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Copyright, Fair Use, Scholarly Communication,
etc.

Libraries at University of Nebraska-Lincoln

7-24-2010

CRIS AND INSTITUTIONAL REPOSITORIES

A Asserson

K Jeffery

Follow this and additional works at: <https://digitalcommons.unl.edu/scholcom>



Part of the [Intellectual Property Law Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Copyright, Fair Use, Scholarly Communication, etc. by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

CRIS AND INSTITUTIONAL REPOSITORIES

A Asserson¹ and K Jeffery^{2*}

¹ Department of Research Management, University of Bergen, Nygardsgaten 5, Bergen, Norway

Email: anne.asserson@fa.uib.no

^{2*} Science and Technology Research Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Chilton, Didcot, Oxfordshire OX11 0QX, UK

Email: keith.jeffery@stfc.ac.uk

ABSTRACT

CRIS (Current Research Information Systems) provide researchers, research managers, innovators, and others with a view over the research activity of a domain. IRs (institutional repositories) provide a mechanism for an organisation to showcase through OA (open access) its intellectual property. Increasingly, organizations are mandating that their employed researchers deposit peer-reviewed published material in the IR. Research funders are increasingly mandating that publications be deposited in an open access repository: some mandate a central (or subject-based) repository, some an IR. In parallel, publishers are offering OA but replacing subscription-based access with author (or author institution) payment for publishing. However, many OA repositories have metadata based on DC (Dublin Core) which is inadequate; a CERIF (Common-European Research Information Format) CRIS provides metadata describing publications with formal syntax and declared semantics thus facilitating interoperability or homogeneous access over heterogeneous sources. The formality is essential for research output metrics, which are increasingly being used to determine future funding for research organizations.

1 THE REQUIREMENT

1.1 Introduction

Let us start with that which is required. This is detailed below by type of user (actor) and role, but we can surely agree that the overall aim must be that research output causes wealth creation and / or improvement in the quality of life. It follows therefore that maximising these desirable properties requires maximum access to research output. Provision of maximum access has technical, legalistic, and economic implications. It also requires a broader context to ensure the research output material is understood and used appropriately.

1.2 The Actors

The researcher requires access to find relevant pre-existing research output and to find possible research collaborators. The research manager requires access to check completeness of recorded outputs from her institution, to compare with that of other institutions, and thus to develop strategy for her institution. The funding agency requires access to ensure defined outputs from the funded research proposal are delivered, to compare outputs with those from other funding agencies, and to find appropriate referees. The policymaker requires access to compare outputs produced by different continents, countries, institutions, and research teams. The innovator requires access to find new ideas which are exploitable for wealth creation or improvement in the quality of life. The educator requires access to obtain teaching material. The student requires access to use learning material. The media require access to obtain information that can be recast as 'stories' which popularise research or raise social, ethical, political, or economic issues concerning the research for the public interest.

1.3 The Roles

Any competent researcher before starting a new research idea will review the existing research output. The more complete and accessible this is to her, the better the review will be, nugatory effort will be avoided and a better (novel) idea will be formulated. A researcher working in one topic area may find an applicable and appropriate

technique, such as an experimental protocol or a computer program for simulation or statistical reduction, from another topic area. As a result of one of the above or by an independent search, a researcher may find a potential collaborator or complementary co-worker for a research idea.

One measure of a researcher capability is evaluation of produced output. The more complete and accessible outputs are, the better the quality of the evaluation. The metrics imposed on the raw data (i.e., how one ranks different publication channels such as journals) are a separate issue, but without complete and verifiable raw data, evaluations are worthless. Similarly the performance of an organisational unit can be evaluated based on its outputs. Indeed, one could compare inputs (funding) with outputs as evaluated to obtain some idea of effectiveness and efficiency.

One may wish to evaluate the literature in different topic areas of fields of research. This may inform strategic decisions on research funding or areas of priority in a research institution. The literature provides a source of ideas, usually with associated research to demonstrate their potential use. This is a mine of information for the entrepreneur or innovator who wishes to invest venture capital to create products or services with associated wealth creation (jobs, profits for shareholders).

Today's teaching material is the research output of years ago. As the pace of learning increases and the volume of research output increases, there is a need for faster and easier access to appropriate research literature by educators. Modern learning is more project-based and less 'chalk and talk.' Students are encouraged to utilise technology to find relevant information.

Journalists and other media professionals need easy access to research outputs in order to find interesting 'stories' for popularising, to research (verify) the background to 'urban myths' about research, and to find researchers suitable for appearing on TV programmes or writing articles.

1.4 Conclusion

We can conclude that all these actors, in the various example roles discussed, require easy (fast, efficient) access to research output material. Technically this implies the need for excellent descriptive metadata, fast searching of metadata, fast searching of text and multimedia, and well-structured results. Furthermore, access to heterogeneous distributed repositories should appear homogeneous and local to the end-user. This implies reconciliation to a canonical syntax (structure) and semantics (meaning), which in turn is likely to involve translation of character sets, language, and ontological terms. Legalistically it requires unfettered access although restrictive metadata may document, for software to enforce, claimed rights that should be respected (like attribution) and even may define a price for access. Economically it requires a business model where costs are minimised (ideally zero as seen by the end-user), any income lies where the work is done, and costs are borne where benefit is obtained. Furthermore, ideally the actors require the research output material in the context of research project, researchers, organizations involved, facilities and equipment, funding, etc.

2 THE TECHNOLOGIES

2.1 CERIF-CRIS

CRIS (Current Research Information Systems) have been developed over the last 40 years. Currently an EU Recommendation to member states, CERIF (Common European Research Information Format), is being adopted quite widely, and it allows interoperation. A CRIS typically has information on projects, persons, organisational units, funding programmes, research outputs (products, patents, and publications), facilities and equipment, and events. The novelty of CERIF is its formal data structure, its use of linking relations to allow n:m relationships with role and temporal duration, its use of multiple character sets, and provision of multilinguality.

Consider the following case illustrated in Figure 1: A person A is an employee of organisation O and a member of organisations M and N both of which are parts of O. She is author of X in which O claims the IPR (intellectual property right) and project leader of P. In CERIF the following records would be in base tables: Person: A; OrgUnit: O,M,N; Publication: X; Project: P. The link tables would be: Person-OrgUnit: A-employee-O, A-member-M, A-

member-N; OrgUnit-OrgUnit: M-partof-O; N-partof-O; Person-Publication: A-author-X; OrgUnit-Publication: O-IPR-X; Person-Project: A-projectleader-P. In fact, the link tables include, as well as role, the temporal information concerning start and end date-time. In this example it may be that when A authored X she was no longer a member of M. This relatively simple example illustrates the power of CERIF as a data model.

CERIF is maintained by the not-for-profit organisation euroCRIS (www.eurocris.org) from whence details are available. Commercial CRIS offerings are available from [uniCRIS] which is fully CERIF-compatible, [Atira], and [Avedas]. Many funding agencies and research institutions have some form of ‘home-brew’ CRIS; the majority are more-or-less CERIF-compatible. The provision of CRIS in a modern e-infrastructure environment has been discussed in [Je2004].

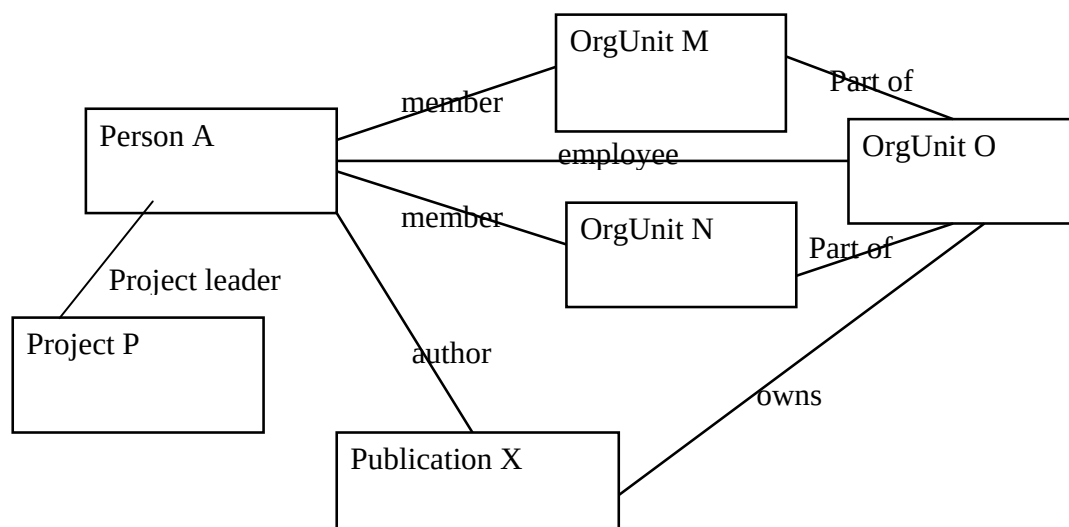


Figure 1: Example of CERIF

2.2 Repositories

Repositories store and provide access to the detailed information. It is usual to separate repositories of research publications from repositories of research datasets and software (e-Science or, better, e-Research repositories) because of their different access patterns and different metadata requirements. The e-Research repositories require much more detailed metadata to control utilisation of the software and datasets in addition to metadata to allow discovery of the resources. At present they tend to be specific to an individual organisation because of their novelty and the differing requirements on metadata imposed by different (commonly international) communities, e.g., in space science, atmospheric physics, materials science, particle physics, humanities, or social science. Publication repositories typically use some form of Dublin Core Metadata [DC] and most are [OAI-PMH] (Open Access Initiative – Protocol for Metadata Harvesting) compliant for interoperability and are indexed by Google Scholar. Example software systems are [ePrints], [DSpace], [Fedora], and [ePubs].

2.3 Metadata

Digitally-created articles rely heavily on both the metadata record and the articles themselves being deposited. International metadata standards and protocols must be applied to repositories so that retrieval may be consistent with appropriate recall (precision) and relevance so that harvesting (or homogeneous retrieval access) across repositories can take place. A model for formalising metadata [Je2000] is required.

The apparently obvious solution is to use conventional repository access and linkage. This technology is using OAI-PMH (Open Access Initiative Protocol for Metadata Harvesting (<http://www.openarchives.org/pmh/>) with DC or

even ORE (Object Re-use and Exchange (<http://www.openarchives.org/ore/>) as packaged metadata. Examples of such an approach have been utilised in the DELOS project and NoE (<http://www.delos.info/>) and DRIVER followed by DRIVER II (<http://www.driverrepository.eu/>). However, it is the experience of the authors that this is insufficient to meet the requirement. The problems concern metadata. The DC 15-element standard does not have a sufficiently formal syntax nor declared semantics for effective processing. Although DC has been extended (Qualified DC) to improve the situation and recent work (2008) has extended DC with domains and ranges, this does not overcome the problem. This may be characterised as the need for machine-understandability as well as machine-readability of the metadata. Furthermore, the research output should be understood in context, that is, the publication or research dataset related to the research projects, persons and their roles, organisational units, funding, research facilities and equipment, etc involved in the research that generated the output. One example should suffice to explain the difficulty using DC. The element contributor is defined at <http://purl.org/dc/elements/1.1/contributor> as:

Label:	Contributor
Definition:	An entity responsible for making contributions to the resource
Comment:	Examples of a Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.

The example illustrates exactly the problem: the ‘type’ of the element contributor is not defined (although with namespaces, domains, and ranges, a limited set of acceptable lexical terms can be defined). The kinds of contributor in the example would likely have different legal status, rights, and responsibilities. There is no concept of the relationship between contributors (except that there is, confusingly, another element named creator, the definition of which has a comment exactly as for contributor). Since 1999 (early in the life of DC) these criticisms have been made by members of euroCRIS and the alternative approach based on formal syntax and defined semantics (described below) proposed.

In fact, the DRIVER consortium itself in a public paper http://www.driversupport.eu/documents/DRIVER_Review_of_Technical_Standards.pdf criticises DC as unsuitable, criticises other formats (such as MODS), and even mentions CRIS and CERIF. A later (and very recent) paper http://www.driversupport.eu/documents/D4%203_Tech_Watch.pdf is more specific on the need for CERIF-like metadata and lifts much information from the euroCRIS website in section 2.3. Despite this later paper recommending integration of CRIS and OAR, there is no technical proposal of how this should be done. Thus, unfortunately these papers misunderstand and misrepresent the CERIF form and function, basically not comprehending the concept of semantically typed n:m relationships between base entities (objects) such as persons and publications recorded using a formal syntax.

Similarly the Knowledge Exchange <http://www.knowledge-exchange.info/> (which has an intersection of members with DRIVER) has considered the relationship between CRIS and repositories and even initiated a project (2008-2010) on this. This project was instigated following a meeting where a euroCRIS member presented the case for integrating CRIS and repositories. Unfortunately again this project misunderstands and misrepresents the form and function of CERIF-CRIS in the same way as DRIVER. Both DRIVER and Knowledge Exchange have claimed there is no method for interoperability between a CRIS and OA repository; in fact within the euroCRIS community there are several working examples, and the solution proposed below is based on the euroCRIS members’ experience of this. In the later driver paper http://www.driversupport.eu/documents/D4%203_Tech_Watch.pdf, some case studies indicate in overview that such a linkage is possible thus contradicting the earlier statement.

Thus, the current DC metadata standards (DC) and (OAI-PMH) for interoperability are insufficient for scalable, automated retrieval with appropriate relevance (precision) and recall. DC is machine-readable but not machine-understandable. One basic problem is that a formalised syntax and semantics (vocabulary) for each relevant DC element was not specified in ‘simple DC’ and has only partially been overcome by the use of namespaces in ‘qualified DC’ as illustrated above. A second problem concerns the element set tags ‘contributor,’ ‘creator,’ and ‘publisher,’ which are actually roles of a person or organisational unit and should be represented by a relationship (between the article and the person or organisational unit) where the role belongs to a namespace and is temporally limited. A third problem is the tag ‘relation,’ which is extremely general; the real world is much better modelled through typed relations with role and temporal validity. Other problems include the tag ‘coverage,’ which only recently has been separated into temporal and spatial aspects yet these are fundamental retrieval criteria for much

material. A formalised version of DC overcoming these limitations has been suggested [Je99b] and defined [AsJe2005] to form also part of the CERIF model allowing tight integration with CRIS. Recently the DC community has recognised these problems and with more recent work [DCAM2007] [DCRDF2007] is attempting to address them.

To ensure that research output material is available for future generations, curation and preservation issues must be addressed. There is current work to define metadata standards to achieve this [OAIS] but a major problem concerns maintaining the articles on current (i.e., usable) media.

2.4 Linking CRIS and institutional repositories

The linking together at an institution of a ‘green’ OA repository of articles, a CRIS (to provide contextual information), and an OA repository of research datasets and software [JeAs2006a] {Figure 2Figure 2} ensures that an institution can manage its IP for benefit whether that benefit is in innovation and investment, in educational resources, in stimulation of future research, or in publicity. Furthermore, the formalised structure of the CRIS allows a reliable workflow to be engineered, which in turn encourages deposit of research outputs. Such a system is being implemented progressively at STFC Rutherford Appleton Laboratory where the CERIF-CRIS is named the Corporate Data Repository, the OA repository is ePubs, and the e-research repository is the e-Science repository.

Linking together these institutional CRIS systems (which have a formal structure and hence can be interoperated reliably and in a scalable way), [Je2005] provides a network of access to institutional OA repositories (of articles) or e-research repositories via the CERIF-CRIS gateways enhancing and controlling the access using the CERIF-CRIS information as formalised, structured, and contextual metadata, which is more detailed than DC and suitable for intelligent (machine-understandable) interoperation {Figure 3}. Interoperation of CERIF-CRIS has been demonstrated, most recently for euroHORCS (European Heads of Research Councils) in October 2006. However, as yet, the whole architecture has not been demonstrated.

The key point is that the metadata for a publication (or dataset) is stored in the CRIS in formal syntax and with defined semantics, and the repository just acts as a deposit space. In this way management information and analysis can be done using the (formal) CRIS while retrieval of the individual publication (or dataset) is done through the repository sanctioned by the CRIS. The repository may or may not also store metadata (usually in DC, for OAI-PMH interoperation and OAISTER retrieval), but this metadata is best generated from the CRIS. This is because the CRIS in a research institution is intimately linked to the researcher workbench, and organizational workflow and much of the metadata required for a publication (author, institution, rights) is already stored in the CRIS and does not need the author to re-input.

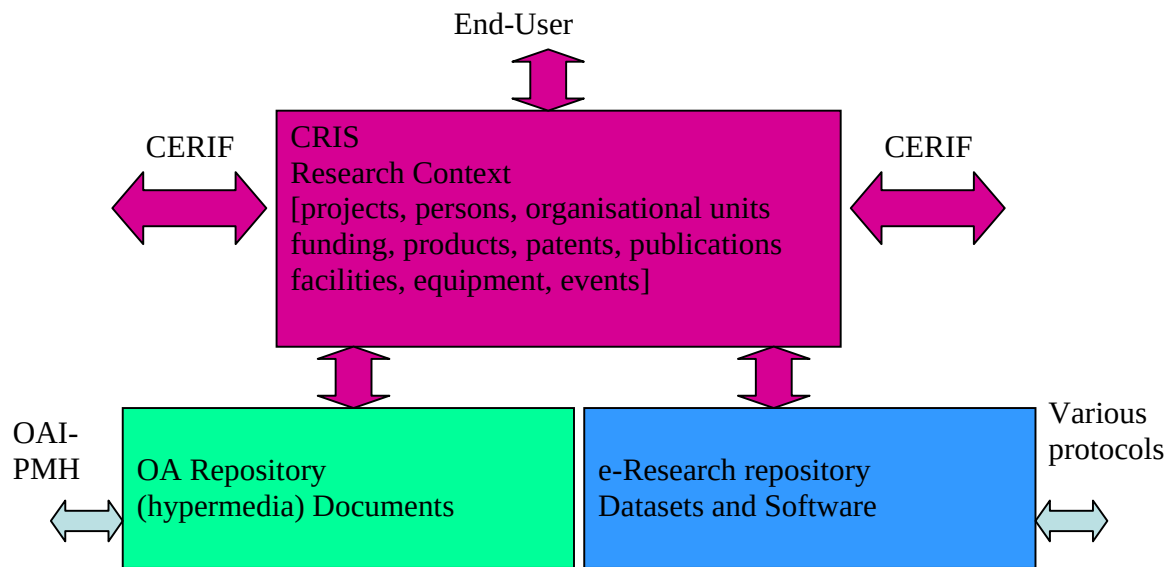


Figure 2. Architecture for an Institution

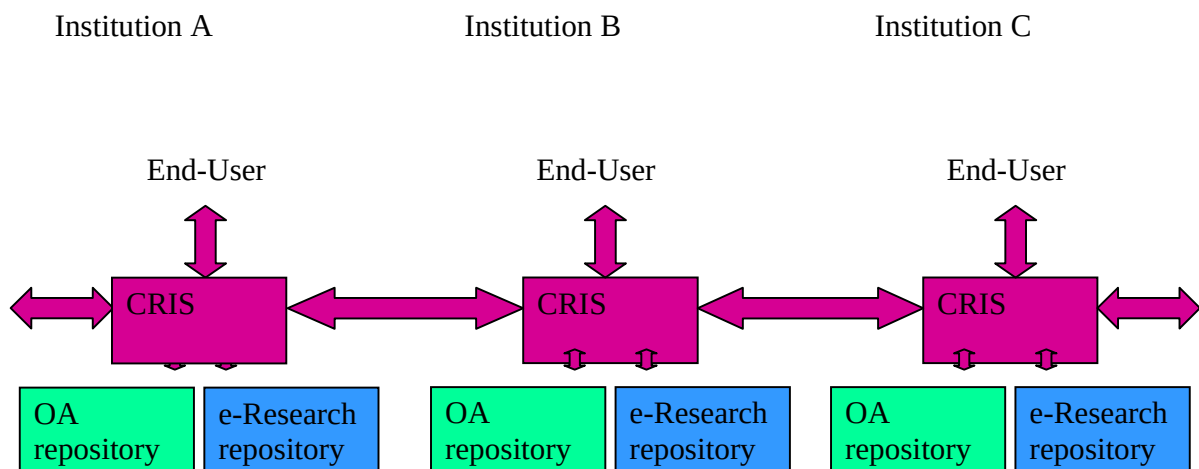


Figure 3. Architecture for OA

3 FUTURE

Looking to the future speculatively, it is possible to imagine ‘green’ OA repositories becoming commonplace and used heavily. At that point, some argue, one could change the business model so that an author deposits in an open access ‘green’ repository but instead of submitting in parallel to a journal or conference peer-review process, the peer-review is done either by:

a) a learned society managing a ‘college’ of experts and the reviewing process, for a fee paid by the institution of the author or the author or

b) allowing annotation by any reader (with digital signature to ensure identification / authentication); in both cases being alerted by ‘push technology’ that a new article matching their interest profile has been deposited.

The former peer-review mechanism would maintain learned societies in business, would still cost the institution of the author or the author, but would probably be less expensive than publisher subscriptions or ‘gold’ (author or author institution pays) open access. The latter is much more adventurous and in the spirit of the internet; in a charming way it somehow recaptures the scholarly process of two centuries ago (initial draft, open discussion, revision, and publication) in a modern world context. It is this possible future that is feared by commercial publishers.

4 REFERENCES AND BIBLIOGRAPHY

CRIS Strategy and Development

[AsJe2004] Asserson, A; Jeffery, K.G.; ‘Research Output Publications and CRIS’ in A Nase, G van Grootel (Eds) Proceedings CRIS2004 Conference, Leuven University Press ISBN 90 5867 3839 May 2004 pp 29-40

[Je2004] Jeffery, K.G.; ‘The New Technologies: can CRISs Benefit’ in A Nase, G van Grootel (Eds) Proceedings CRIS2004 Conference, Leuven University Press ISBN 90 5867 3839 May 2004 pp 77-88

[Je2005] K G Jeffery CRISs, Architectures and CERIF CCLRC-RAL Technical Report RAL-TR-2005-003 (2005)

[DiBaHoArvaMe2006] NARCIS: The Gateway to Dutch Scientific Information ELPUB2006. Digital Spectrum: Integrating Technology and Culture - Proceedings of the 10th International Conference on Electronic Publishing held in Bansko, Bulgaria 14-16 June 2006 / Edited by: Bob Martens, Milena Dobrev. ISBN 978-954-16-0040-5, 2006, pp. 49-58

[JeAs2006a] Keith G Jeffery, Anne Asserson: ‘CRIS Central Relating Information System’ in Anne Gams Steine Asserson, Eduard J Simons (Eds) ‘Enabling Interaction and Quality: Beyond the Hanseatic League’; Proceedings 8th International Conference on Current Research Information Systems CRIS2006 Conference, Bergen, May 2006 pp109-120 Leuven University Press ISBN 978 90 5867 536 1

[JeAs2006b] Keith G Jeffery, Anne Asserson: ‘Supporting the Research Process with a CRIS’ in Anne Gams Steine Asserson, Eduard J Simons (Eds) ‘Enabling Interaction and Quality: Beyond the Hanseatic League’; Proceedings 8th International Conference on Current Research Information Systems CRIS2006 Conference, Bergen, May 2006 pp 121-130 Leuven University Press ISBN 978 90 5867 536 1

CRIS Systems

[uniCRIS]: Retrieved from the WWW, April 10, 2010: www.unicris.com

[Atira] (PURE System): Retrieved from the WWW April 20, 2010: <http://www.atira.dk/en/>

[Avedas]: Retrieved from the WWW April 20, 2010: <http://www.avedas.com/>

CERIF

CERIF data model documentation: Retrieved from the WWW, April 21, 2010:

<http://www.eurocris.org/cerif/introduction/>

[AsJeLo2000] Asserson,A; Jeffery,K.G; Lopatenko,A: ‘CERIF: Past, Present and Future’ in Adamczak,W & Nase,A (Eds): Proceedings CRIS2002 6th International Conference on Current Research Information Systems; Kassel University Press ISBN 3-0331146-844 pp 33-40

[LoAsJe2000] Lopatenko,A Asserson,A; Jeffery,K.G;: ‘CERIF: Information Retrieval of Research Information in a Distributed Heterogeneous Environment’ in Adamczak,W & Nase,A (Eds): Proceedings CRIS2002 6th International Conference on Current Research Information Systems; Kassel University Press ISBN 3-0331146-844 pp 59-68

Metadata

[Je99b] Jeffery, K G: ‘An Architecture for Grey Literature in a R&D Context’ Proceedings GL’99 (Grey Literature) Conference Washington DC October 1999 Retrieved from the WWW April 20 2010
<http://www.emeraldinsight.com/Insight/viewContentItem.do?contentType=Article&contentId=861814>

[Je2000] Jeffery, K G: 'Metadata': in Brinkkemper,J; Lindencrona,E; Solvberg,A (Eds): 'Information Systems Engineering' Springer Verlag, London 2000. ISBN 1-85233-317-0.

[JeLoAs2000] Jeffery,K.G; Lopatenko,A; Asserson,A: 'Comparative Study of Metadata for Scientific Information: The Place of CERIF in CRISs and Scientific Repositories' in Adamczak,W & Nase,A (Eds): Proceedings CRIS2002 6th International Conference on Current Research Information Systems; Kassel University Press ISBN 3-0331146-844 pp 77-86

[AsJe2005] Asserson, A; Jeffery, K.G.; 'Research Output Publications and CRIS' The Grey Journal volume 1 number 1: Spring 2005 TextRelease/Greynet ISSN 1574-1796 pp5-8

[DC] Retrieved from the WWW, April 10, 2010: <http://dublincore.org/>

[OAI-PMH] Retrieved from the WWW, April 10, 2010:
<http://www.openarchives.org/OAI/openarchivesprotocol.html>

[MARC] Retrieved from the WWW, April 10, 2010 <http://www.loc.gov/marc/>

Curation

[OAIS] Retrieved from the WWW, April 10, 2010: <http://ssdoo.gsfc.nasa.gov/nost/isoas/>

Background Documents on OA

Self-archiving FAQ: Retrieved from the WWW, April 10, 2010: <http://www.eprints.org/self-faq/>

Bibliography of OA Advantage Retrieved from the WWW, April 10, 2010: <http://opcit.eprints.org/oacitation-biblio.html>

American Scientist Open Access Forum Retrieved from the WWW April 20, 2010: <http://amsci-forum.amsci.org/archives/American-Scientist-Open-Access-Forum.html>

[Je2006] Retrieved from the WWW, April 10, 2010:
http://www.ercim.org/publication/Ercim_News/enw64/jeffery.html

Need for OA (problems with traditional access)

[Ha1995] Harnad, S. (1995) A Subversive Proposal. In: Ann Okerson & James O'Donnell (Eds.) Scholarly Journals at the Crossroads; A Subversive Proposal for Electronic Publishing. Washington, DC., Association of Research Libraries, June 1995. Retrieved from the WWW, April 10, 2010: <http://www.arl.org/scomm/subversive/toc.html>

[RoNi2005] Rowlands, I. and Nicholas, D. (2005) /New Journal Publishing

[Sp2005] Sparks, S. (2005) /JISC Disciplinary Differences Report,/ Rightscom, London.
http://www.jisc.ac.uk/uploaded_documents/Disciplinary%20Differences%20and%20Needs.doc.

[EPS2006] EPS (2006) UK scholarly journals: 2006 baseline report An evidence-based analysis of data concerning scholarly journal publishing, RIN, RCUK and DTI. Retrieved from the WWW, April 10, 2010:
<http://www.rin.ac.uk/data-scholarly-journals>.

OA Policy

[Budapest]: Retrieved from the WWW, April 10, 2010: <http://www.soros.org/openaccess/read.shtml>

[Bethesda]: Retrieved from the WWW, April 10, 2010: <http://www.earlham.edu/~peters/fos/bethesda.htm>

[Berlin]: Retrieved from the WWW, April 10, 2010: <http://www.zim.mpg.de/openaccess-berlin/signatories.html>

[OECD]: Retrieved from the WWW, April 10, 2010:
http://www.oecd.org/document/15/0,2340,en_2649_201185_25998799_1_1_1_1,00.html

[UKHCST] UK House of Commons Science and Technology Select Committee Retrieved from the WWW April 20 2010 <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf>

[RCUK] Retrieved from the WWW April 20, 2010: <http://www.rcuk.ac.uk/access/default.htm>

[FRPAA] Retrieved from the WWW April 20, 2010: <http://www.taxpayeraccess.org/issues/frpaa/index.shtml>

[EURAB] Retrieved from the WWW April 20, 2010:
http://ec.europa.eu/research/eurab/pdf/eurab_sciplib_report_recomm_dec06_en.pdf

[STM] Retrieved from the WWW April 20, 2010: http://www.stm-assoc.org/public_affairs_brussels_declaration.php

OA Policies, Registries and Directories

Overview: Retrieved from the WWW, April 10, 2010: <http://www.openarchives.org>

Institutional Archives Registry: Retrieved from the WWW, April 10, 2010:
<http://archives.eprints.org/eprints.php?action=browse>

Institutional Self-Archiving Policy Registry. Retrieved from the WWW, April 10, 2010:
<http://www.eprints.org/signup/fulllist.php>

Journal Self-Archiving Policy Directory. Retrieved from the WWW, April 10, 2010: <http://romeo.eprints.org/>

‘Gold’ open-access publication channels. Retrieved from the WWW, April 10, 2010: <http://www.doaj.org/>

In all cases self-archive a version in your institutional repository. Retrieved from the WWW, April 10, 2010:

<http://www.eprints.org/self-faq/> <http://archives.eprints.org/> <http://openaccess.eprints.org/>

[Po2007] Retrieved from the WWW, April 10, 2010: <http://poynder.blogspot.com/2007/03/open-access-war-in-europe.html>

Effectiveness of OA

[HaCaBrOp2003] Harnad, S., Carr, L., Brody, T. & Oppenheim, C. (2003) Mandated online RAE CVs Linked to University Eprint Archives: Improving the UK Research Assessment Exercise whilst making it cheaper and easier. Ariadne 35. Retrieved from the WWW, April 10, 2010: <http://www.ecs.soton.ac.uk/~harnad/Temp/Ariadne-RAE.htm>

[Ha2007] Harnad, S. (2007) Open Access Scientometrics and the UK Research Assessment Exercise. Proceedings of the 11th Annual Meeting of the International Society for Scientometrics and Informetrics. Madrid, Spain, 25 June 2007. Retrieved from the WWW, April 10, 2010: <http://arxiv.org/abs/cs.IR/0703131>

[SwBr2004] Swan, A. and Brown, S. (2004) ISC/OSI JOURNAL AUTHORS SURVEY Report. Technical Report, JISC, HEFCE. Retrieved from the WWW, April 10, 2010: <http://eprints.ecs.soton.ac.uk/11002/>

[SwBr2005] Swan, A. and Brown, S. (2005) Open access self-archiving: An author study. Technical Report, External Collaborators, JISC, HEFCE. Retrieved from the WWW, April 10, 2010:
<http://eprints.ecs.soton.ac.uk/10999/>

[SwNePrMuOpOBHaRo2005] Swan, A., Needham, P., Proberts, S., Muir, A., Oppenheim, C., O'Brien, A., Hardy, R., and Rowland, F. (2005) Delivery, Management and Access Model for E-prints and Open Access Journals within Further and Higher Education. Technical Report, JISC, HEFCE. Retrieved from the WWW, April 10, 2010:
<http://eprints.ecs.soton.ac.uk/11001/>

Retrieved from the WWW, April 10, 2010: <http://opcit.eprints.org/oacitation-biblio.html>

Retrieved from the WWW, April 10, 2010: http://citebase.eprints.org/isi_study/

[HoStSh2006] Houghton, J., Steele, C. & Sheehan, P. (2006) Research Communication Costs in Australia: Emerging Opportunities and Benefits. A report to the Department of Education, Science and Training. Retrieved from the WWW, April 10, 2010: http://www.dest.gov.au/NR/rdonlyres/0ACB271F-EA7D-4FAF-B3F7-0381F441B175/13935/DEST_Research_Communications_Cost_Report_Sept2006.pdf

[HoSh2006] Houghton, J. & Sheehan, P. (2006) The Economic Impact of Enhanced Access to Research Findings. Centre for Strategic Economic Studies Victoria University. Retrieved from the WWW April 20 2010
<http://www.cfses.com/documents/wp23.pdf>

[Chawki2005] Retrieved from the WWW, April 10, 2010:
<http://www.crsc.uqam.ca/lab/chawki/graphes/EtudeImpact.htm>

[Sa207] Sale, A. (2007) The Patchwork Mandate D-Lib Magazine 13 1/2 January/February Retrieved from the WWW, April 10, 2010: <http://www.dlib.org/dlib/january07/sale/01sale.html> doi:10.1045/january2007-sale.

Copyright status of publishers

[Romeo] Retrieved from the WWW, April 10, 2010: <http://romeo.eprints.org/stats.php>

Effect of OA on Journal Subscriptions

[SwBr2005] Swan, A. and Brown, S (2005) Open access self-archiving: an author study.

Retrieved from the WWW April 20, 2010:
http://www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

[BLDRHaSh2005] Berners-Lee, T., De Roure, D., Harnad, S., and Shadbolt, N. (2005) Journal publishing and author self-archiving: Peaceful Co-Existence and Fruitful Collaboration. Retrieved from the WWW, April 10, 2010: <http://eprints.ecs.soton.ac.uk/11160/>

OA Repositories / systems

[arXiv] Retrieved from the WWW, April 10, 2010: <http://arxiv.org/>

[ePubs] Retrieved from the WWW, April 10, 2010: <http://epubs.cclrc.ac.uk/>

[eprints] Retrieved from the WWW, April 10, 2010: <http://www.eprints.org/>

[DSpace] Retrieved from the WWW, April 10, 2010: <http://www.dspace.org/>

[Fedora] Retrieved from the WWW, April 10, 2010: <http://www.fedora.info/>

[Citebase] Retrieved from the WWW, April 10, 2010: <http://citebase.eprints.org/>