

Institution: University of Southampton
Unit of Assessment: 11 Computer Science and Informatics
Title of case study: 11-06 Open Access and Digital Archiving
<p>1. Summary of the impact</p> <p>Free and open access (OA) to publicly funded research offers significant benefits, but it also requires complex new systems to underpin it. University of Southampton research has resulted in software products enabling large numbers of research institutions to implement their own digital research repositories. Studies on the viability and impact of OA have steered institutions towards a more cost-effective and impactful model for disseminating research, and UK public policy has been directly influenced by the Southampton team's advocacy work. The research also led to economic benefits through two spin-outs and the development of digital archiving techniques, which have been widely used by broadcast and film institutions.</p>
<p>2. Underpinning research</p> <p>In recent years, the OA model has been playing an increasingly important role in scientific publishing. It has been argued that there is a strong moral case for publicly-funded research to be freely accessible. Research carried out at the University of Southampton has produced institutional repository and digital archiving software, tackled the technical challenges of implementation and driven the OA agenda forward by proving its impact.</p> <p>Key researchers were Professor Leslie Carr (Web and Internet Science Research Group, at Southampton since 1994); Professor Stevan Harnad (Web and Internet Science Research Group, since 1994); and Dr Matthew Addis (IT Innovation Centre manager, at Southampton since 1996).</p> <p>A seminal paper published by Harnad in 1994 [3.1] sparked international debate on systematically disseminating academic research online to address the limitations of traditional publishing: limited readership, costs and delays. In 1995 Southampton launched the Open Journal project (funded by the Joint Information Systems Committee) and in 1997 it created one of the first central OA repositories, Cogprints. In 1999 Carr and Harnad participated in the Open Archives Initiative to design the metadata-harvesting protocol for creating interoperable institutional repositories, and funding was granted for the "OpCit" collaboration [Grant 1] between ECS, Cornell University in the USA and arXiv.org (then Los Alamos National Laboratory's physics repository) to link references of all refereed research worldwide. Citebase, the first citation impact-measuring search engine was developed and used by the Southampton team for bibliometric analyses of the correlation between OA and citation rates (the "OA advantage"); the study found a significant correlation of approximately 0.4 between citation and download impact of articles [3.2].</p> <p>From 2000, the Southampton team developed <i>ePrints</i>, a generic, open source institutional repository system which they also used as a platform to research the impact of OA and implementation practicalities [3.3].</p> <p>In 2003, the University of Bath partnered with Southampton on "EBank UK" [Grant 2] to develop digital repositories of crystallographic data using ePrints. The principal pilot with the International Union of Crystallography generated a common standard for data exchange adopted by experimental labs internationally.</p> <p>Under the 2004 "PRESERV" project [Grant 3], the Southampton team worked with the British Library to explore cost-effective methods for long-term digital preservation, developing service models extending ePrints' capabilities to media assets and opening up techniques previously the preserve of national libraries.</p> <p>"EdSpace", a 2007 Southampton-led e-learning project [Grant 4] focused on innovations in the dissemination of learning materials and was closely aligned with the Massive Open Online Courses movement, which extends public access to HEI educational materials [3.4].</p> <p>In 2009, Southampton researchers examined the costs and risks of archiving and restoration for the audiovisual industry [Grant 5, 6, 7] and produced <i>A-Stor</i>, a complement to ePrints, providing storage services for repositories, such as data movement [3.5]. This work was extended to the engineering sector, including a project with BAESystems and EuroStep (a Swedish company</p>

Impact case study (REF3b)

focusing on the application of product data standards for product life cycle management) on data retention and access services in aerospace supply chains [**Grant 8**].

In a 2010 paper [**3.6**] Southampton and University of Quebec researchers set out to establish whether the OA advantage was down to authors preferentially selecting the highest quality work to make OA, by comparing self-selective and mandatory self-archiving for a sample of 27,197 articles published 2002–2006. The OA advantage proved just as high for both models, but was correlated with quality, establishing that it is in fact readers who select for quality.

3. References to the research (best three are starred)

Research outputs

- 3.1 Harnad, S (1995) A Subversive Proposal. In, Okerson, A. and O'Donnell, J. (eds.) *Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing*. Association of Research Libraries.
- 3.2 * Brody, T., Harnad, S. and Carr, L. (2006) Earlier Web Usage Statistics as Predictors of Later Citation Impact. *Journal of the American Association for Information Science and Technology (JASIST)*, 57 (8). pp. 1060-1072.
- 3.3 * Hitchcock, S, Carr, L., Jiao, Z., and Bergmark, D., Hall, W., Lagoze, C. and Harnad, S. (2000) Developing services for open eprint archives: globalisation, integration and the impact of links, 5th ACM conference on Digital libraries, ACM, 143-151, 2000.
- 3.4 Davis, H., Carr, L., Hey, J., Howard, Y., Millard, D., Morris, D. and White, S. (2010) Bootstrapping a Culture of Sharing to Facilitate Open Educational Resources. *IEEE Transactions on Learning Technologies*, 3 (2). pp. 96-109.
- 3.5 Addis, M., Wright, R. and Miller, A. (2009). The Significance of Storage in the "Cost of Risk" of Digital Preservation. *International Journal of Digital Curation (IJDC)*, 4 (3).
- 3.6 * Gargouri, Y., Hajjem, C., Lariviere, V., Gingras, Y., Brody, T., Carr, L. and Harnad, S. (2010) Self-Selected or Mandated, OA Increases Citation Impact for Higher Quality Research. *PLOS ONE*, 5 (10). e13636.

Projects

- Grant 1 "OpCit: Open Citations", NSF / JISC International Digital Libraries Programme, 1999-2002. £300k. PI Stevan Harnad. <http://opcit.eprints.org/>
- Grant 2 "EBank UK" JISC Support for E-Research Programme, 2003 –2004 . £120k. PI LA Carr.
- Grant 3 "PRESERV" (phase I & II) JISC Digital Preservation Programme, 2004 – 2008. £293k. PI LA Carr
- Grant 4 "EdSpace," JISC Institutional Innovation Programme, 2007 –2009, PI HC Davis, Co I LA Carr, £300K <http://www.edspace.ecs.Southampton.ac.uk>
- Grant 5 "PrestoSpace", EC FP6, 2004 – 2008. €380k PI Matthew Addis. <http://www.prestospace.org/>
- Grant 6 AVATAR-m UK (TSB) Jan 2006 – Nov 2010 £630k www.avatar-m.org.uk
- Grant 7 PrestoPRIME EC (FP7) Jan 2009 – Nov 2012 €700k www.prestoprime.org
- Grant 8 "RASSC", UK TSB, 2006 – 2010 £85k PI Matthew Addis.

4. Details of the impact

While other bodies have been working on OA, the University of Southampton has been the only UK institution carrying out research and development while lobbying for policy change.

Impact on Policy

"Gold" OA mandates require authors to publish their work in available OA journals, whereas under "Green" they publish in toll-access journals and their institution's OA repository. Southampton's recommendation to the Select Committee on Science and Technology in 2004, that all UK institutions and funders adopt Green OA mandates, was accepted by the committee. However, in

Impact case study (REF3b)

2012 the Government accepted the Finch report's recommendation to prefer Gold, which then informed Research Councils UK's policy. Southampton has been increasingly successful in lobbying for Green, as evidenced by the new HEFCE proposed policy advocating Green and Gold (consultation launched July 2013) [5.1]. Harnad submitted written evidence on the two types of OA to the House of Lords Science and Technology Select Committee and the House of Commons Business, Innovation and Skills Select Committee (both February 2013) whose recommendations were critical of the original Finch Gold strategy and highly supportive of the role of open repositories [5.2].

Impact on Practitioners

The original 2000 release of ePrints was the first "general purpose" institutional repository software (other repositories had been discipline-specific) and pioneered the notion that individual HEIs could implement repositories across all their departments. Since 2008, 46 UK institutions have implemented ePrints (total to date 113); as at July 2013, these institutions host 2.2 million publication records and sample figures indicate each paper is downloaded around 50 times a year [5.3]. ePrints installations are used at all forms of research-involved organisations including:

- Health: UK NHS Trusts (Heart of England NHS Trust) Ifakara Health Institute in Tanzania, Fred Hutchinson Cancer Research Centre in the US, Vinzenz Gruppe in Austria and the Swiss pharmaceuticals company, Novartis.
- Charitable organisations: National Museums of Scotland, British Library
- Governmental: DEFRA in the UK, Queensland Department of Agriculture, Fisheries and Forestry in Australia, National Oceanic and Atmospheric Administration (NOAA) in the USA.
- NGOs: ICOMOS (Conseil International des Monuments et des Sites) UNESCO, Global HIV Vaccine Enterprise.

The Southampton team was instrumental to the formation of the United Kingdom Council of Research Repositories, to drive the recognition of repository administration and management as a professional function and sharing of best practice. In the six years since its inception in 2007, UKCoRR has grown to 257 members.

Economic Benefit

Demand for ePrints led to the creation in 2006 of a successful service business, ePrints Services, which since 2008 has provided repository hosting services, training and consultancy to over 100 client organisations. ePrints Services employs seven staff and has generated over £3 million in revenues [5.4]. ePrints Services continue to develop ePrints in response to the needs of industry as well as the public sector: the latest version, 3.3.12, features "ePrints Bazaar", which allows repository administrators to increase functionality by installing plug-ins and extensions.

A successful new business was also spun out in 2011 for the A-Stor software: Arkivum, a contractually-guaranteed archiving service for digital assets. Arkivum, which secured £2.3 million in venture capital funding, employs ten staff and has fifteen current client contracts [5.5].

In addition to the direct impact on industry demonstrated by the launch of ePrints Services and Arkivum, the work of the Southampton team continued to influence the preservation and audiovisual communities throughout the REF period. Partners of the "AVATAR-m", "PrestoSpace". "PrestoPRIME" and "RASSC" projects adopted best practice developed by Southampton research [5.6]. From 2012, risk management and cost modelling techniques developed in PrestoSPACE and RASSC have been applied by Arkivum to deliver data assurance services to 15 customers that include Tate Galleries, MRC National Institute of Medical Research and the Financial Conduct Authority. Other research outcomes such as open source tools for preservation planning have been used by digital archiving and audiovisual organisations like the National Archives (2010) [5.7].

One significant commercial impact from the research has been to encourage organisations to adopt a more cost-effective model for archiving and disseminating research. The 2009 Houghton report (on the economic implications of alternative scholarly publishing models) estimated that OA publishing for journal articles could deliver system savings of around £215 million a year in the UK (at 2007 prices and publishing activity) of which £50 million would accrue to research institutions outside of HE [5.8].

Public Engagement

Southampton is active in OA policy design and awareness-raising worldwide. Harnad has been a board member of EnablingOpenScholarship, which promotes open education initiatives, since its launch in 2009 [5.9]. He has also lectured on OA at events open to the public such as the 2008 International Science Fair in Trieste, Italy (which attracts around 30,000 visitors in total). Addis has been active in outreach on the curation and preservation of audiovisual collections, e.g. through the Screening the Future series of events for the audiovisual community, as organised by the PrestoCentre (www.prestocentre.org) advisory board member for the 4C project, (<http://www.4cproject.eu/>) and as a member of the Digital Preservation Coalition. (<http://www.dpconline.org/>).

Finally, Harnad and colleagues have done a considerable amount of work to raise awareness and understanding of the benefits of OA among the general public. Some examples among many include Harnad's authored article in The Guardian (September 2012); his three 2009 YouTube videos, the first of which has attracted over 700 views by July 2013; and an extensive February 2010 interview on Information Today Inc, which targets librarians and the information industry.

Section 5

[5.1] HEFCE Consultation on open access in the post-2014 Research Excellence Framework, <http://www.hefce.ac.uk/pubs/year/2013/201316/>

[5.2]

- a) Harnad Evidence to House of Lords Science and Technology Select Committee on Open Access: <http://eprints.soton.ac.uk/348479/>
- b) Harnad Evidence to BIS Select Committee Inquiry on Open Access: <http://eprints.soton.ac.uk/348483/>
- c) BIS Select Committee on Open Access Final Report: <http://www.publications.parliament.uk/pa/cm201314/cmselect/cmbis/99/9902.htm>

[5.3] The Registry of OA Repositories (roar.eprints.org) monitors the impact of OA repository software across the world. For data on the 46 institutions which have implemented EPrints since 2008 see <http://tinyurl.com/psb9bjt>

[5.4] EPrints Services Business Manager can provide a list of customers on request and confirm other impact claims related to the company.

[5.5] Details of the 45.8% holding IP group has in Arkivum can be found here: <http://www.ipgroupplc.com/portfolio/arkivum> and from Arkivum Chief Technology Officer.

[5.6] Principal Scientist, BAE systems Advanced Technology Centre.

[5.7] Development of the National Archives Linked Data PRONOM Service: <http://labs.nationalarchives.gov.uk/wordpress/index.php/2011/01/linked-data-and-pronom/>.

[5.8] Houghton report on Economic Implications of Alternative Scholarly Publishing Models 2009 <http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf>

[5.9] http://www.openscholarship.org/jcms/c_6095/en/people?hiText=Harnad