Repository management: an emerging profession in the information sector.

Keywords: Open Access, repository, scholarly communication, self-archiving, information profession, Repositories Support Project.

Introduction

This paper describes the emerging role of the repository manager in the context of the growth in open access (OA) and the rise of institutional repositories in universities. Traditionally repositories have been a means for the storage and safekeeping of objects but they are also a key means of enabling free and open access to research outputs. It proposes that librarians have played a key role in this field and focuses on the current situation in the United Kingdom. A survey which was conducted in July/August 2010 about the background and skills of repository staff is described. It concludes by considering the importance of support in the workplace, exemplified by the Repositories Support Project and the role of professional networks such as the United Kingdom Council for Research Repositories (UKCoRR).

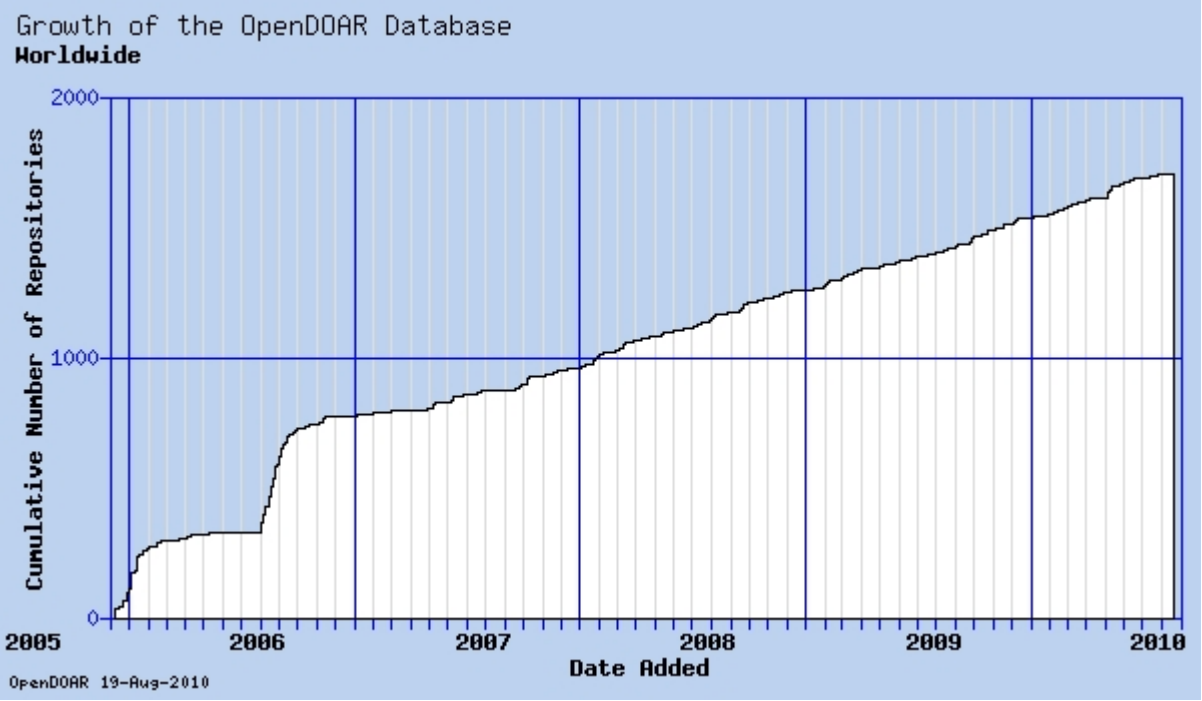
The Open Access landscape

The Open Access movement is bound up with the development of the internet and digital information in the 1990s. The concept of OA has been enshrined in the famous Budapest Open Access Initiative (2002), the Bethesda Statement on Open Access Publishing (2003) and the Berlin Declaration on Open Access to Knowledge in the Science and Humanities (2003). A succinct definition of open access was published by the Research Information Network (2010): "Open Access (OA) means that scholarly literature is made freely available on the internet, so that it can be read, downloaded, copied, distributed, printed, searched, text mined, or used for any other lawful purpose, without financial, legal or technical barriers, subject to proper attribution of authorship" . There are two routes to making research outputs freely available and these are commonly known as the green and the gold routes. The green route, also known as self archiving, describes a process where academics or their mediators deposit the full text of their work in repositories, subject and institutional, which, through indexing by search engines and specialist services, are made available to the world. The gold route follows the model of traditional journal publication with the difference that the publisher recoups costs from the "author" and can therefore make it freely available. This is usually funded by the research grant or the author's institution, although SciELO (Scientific Electronic Library Online in Latin America and the Caribbean) has developed a cooperative model for developing countries. This paper looks at the impact of the growth of institutional repositories (green route) on the information profession and as a catalyst for new roles and skills requirements in the sector.

The growth of repositories

The first part of the 21st Century has seen the number of institutional repositories grow at a huge rate – at least one per working day over the last three years (Swan, 2011). A graph from the OpenDOAR database of Open Access Repositories demonstrates this:

Figure 1 – Growth of the OpenDOAR database – Worldwide



For UK institutions, they provide an opportunity to showcase their research and can be used to support marketing activities to attract high quality staff and students. There are advantages for academics too, they provide a single location for all their research outputs and can allow a feed to personal web pages, removing the need for the individual to maintain their own web profiles. Moreover, there is a body of evidence that shows a distinct citation advantage to OA publications (Antelman, 2004, Hajjem, Harnad and Gingras, 2005, Swan, 2010). They also have a curation role in providing a secure and stable environment for publications. The data required for the Research Assessment Exercise (RAE) and the forthcoming Research Excellence Framework (REF) can be generated by repositories. In addition, research funders are increasingly requiring institutions to demonstrate the impact of the research, in addition to the usual standards of excellence. Research Councils UK has recently developed its Pathways to Impact as guidance for applicants for research funding (RCUK). At the University of Glasgow (2010), the research information system and the repository (Enlighten) have been integrated in a project called Enrich. One of the benefits of this is that research funders and project information can be seamlessly linked to subsequent publications which in turn can produce reports on downloads and reuse. This data can be used to report on impact. JISC (Joint Systems Information Committee) has invested heavily over the last few years in repository development with a number of programmes and has recently set out its roadmap to the future (Heery, 2009).

An emerging role in the information sector

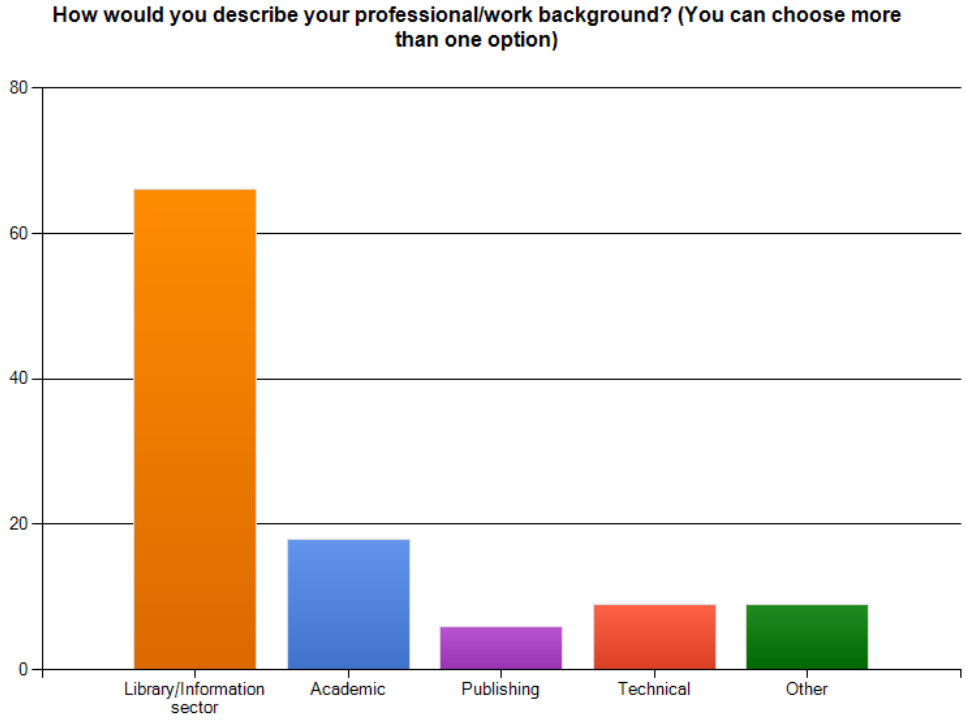
Librarians have played a significant role in leading and managing this growth as they have traditionally been champions of access to knowledge (Ottaviani 2009). There is a growing demand for librarians with skills in managing digital collections and this is changing the role of libraries: “They are no longer passive receivers of information but active disseminators of intellectual output for entire universities” (Walters, 2007). Corral includes repository librarians in her model of the “blended information professional” (Corral, 2010). Currently, the focus is very much on the management of collections of outputs, mostly text based research articles although other media are also included in

repositories such as videos and images from the Art and Design sector. A more recent development is the involvement of librarians in research data management and it has been identified as a new career path (Swan, 2011). Garritano and Carlson (2009) describe a case study of collaboration on e-Science Projects and envisage a future where librarians are co-Principal Investigators on grant proposals. Familiar skills can be brought to the table, such as categorising data (metadata), management and storage of data and advising on access rights and policies but librarians will need to understand research methods and workflow if they are to play a full part.

Backgrounds of repository staff in the UK

In 2009, Kennan and Kingsley found that 71.1% of staff working in repositories in Australian universities had library backgrounds. The Repositories Support Project conducted a survey in 2010 and found a similar result for UK institutions – 78.6%.

Figure 2 – Professional background of staff



The survey was carried out online from 29th July to 5th September using SurveyMonkey. It was publicised to members of the UK Council for Research Repositories (UKCoRR) only. The rationale being that it was restricted to UK staff and that a good proportion of them are members of the organisation. Recipients were encouraged to distribute the survey to any appropriate person in their institution who was not a member. UKCoRR currently (August 2010) has 215 members and there were 84 respondents.

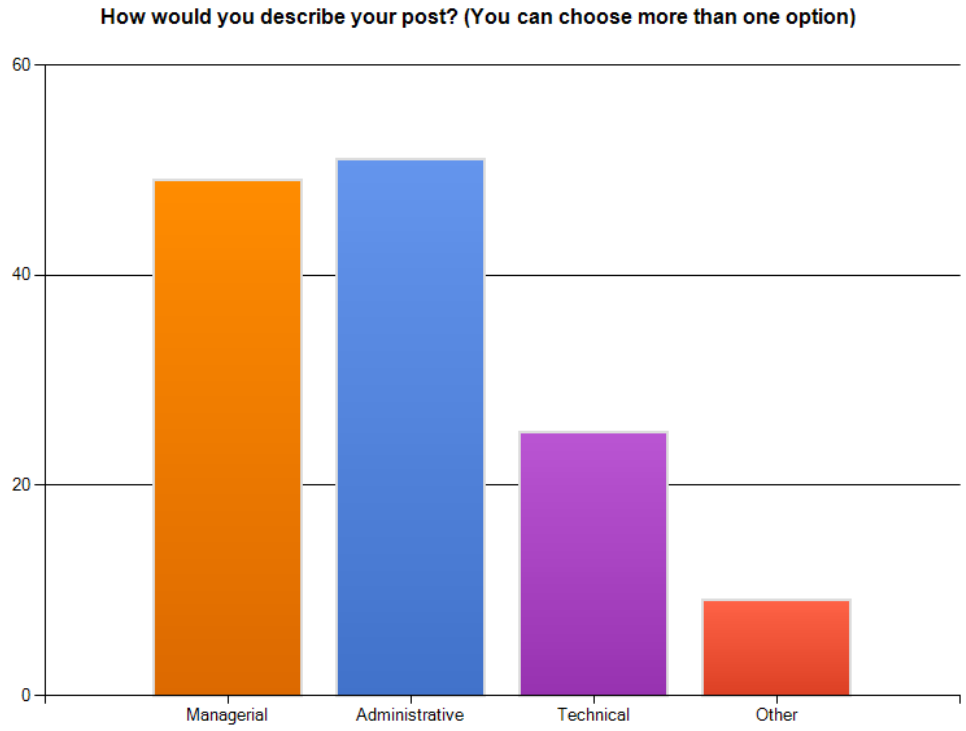
The vast majority of respondents had a first degree (95%) and many had post graduate qualifications (74%). 91.7% (77) were based in higher education institutions and 8.3% (7) worked in Research Institutes. The general picture was that respondents combined repository work with other roles, probably in the same institution (although the survey did not ask this) – 76.2% (64) were part time. For part time staff, the number of hours devoted to the repository varied from “a couple of hours a month” to 30 hours per week. The approximate pattern is shown in the table below (not all respondents provided this information):

Figure 3 – Work patterns

Days worked per week	Number	%
>half	13	23.6
Half - 1	7	12.7
1 – 2.5	8	14.5
2.5	11	20
2.5 - 3	8	14.5
3 - 4	8	14.5

73.8% of all staff worked as part of a team. This seems to indicate that the majority of people do not work solely on the repository but do work alongside other repository colleagues. People were also asked to describe their posts:

Figure 4 Types of role



Role and skills of repository staff

As a result in the growth of repositories, there have been some national initiatives to scope the skills and knowledge requirements primarily to assist in the recruitment of staff (SHERPA, 2008, JISC

2009). There are many different permutations in staffing structures across UK universities but there are primarily three roles associated with repository delivery:

- Repository management – strategic and financial management, advocacy and communication, staff and project management, expert advice to the institution.
- Technical – knowledge and experience of software platforms and the main repository software, deployment, testing, upgrading and development of software.
- Administrative – adding records, checking metadata and copyright.

The table below provides more detail on these roles. It is an updated version of the original guidance developed by SHERPA with input from the repository community in 2008. In addition, in 2009 JISC commissioned a Recruitment Toolkit which lists key competencies. There will be many aspects listed in the table which are familiar to the library and information world but repository staff need to extend their expertise in order to deliver their services. One example of this is the subject of metadata standards. Readers will not be surprised to see that familiarity with Dublin Core, MARC, OAI-PMH etc is listed as a requirement. However, it is also essential for repository staff to understand CERIF (the Common European Research Information Format). This is a data model which enables the storage and interchange of information between Current Research Information Systems (CRIS). Many universities are adopting CRIS as the hub for managing their research information – they interact with other university systems to link data about research funding, the individual (Human Resources), projects and finance. Universities recognise the need to also integrate the actual research outputs, and even the research data, which are housed in the repository. Staff need to have the knowledge to enable interoperability with these systems.

Figure 5 Staff and skills set

Management
Manage the budget and cost for future development of the repository. Source funding opportunities for repository projects where appropriate Collaborate with research staff in the university to prepare research proposals (advising on data management and standards) Manage the repository service by identifying goals and future strategies for improvement in the repository service based on new developments, usage statistics and feedback from users. Develop workflows to manage the capture, description and preservation etc. of research outputs Manage the day-to-day running of the repository including any mediated-deposit service (if required or possible) or self-archiving by authors Coordinate and manage activities of repository personnel and coordinate repository development with associated departments Monitor the quality of the service – handling feedback and complaints and managing user expectations
Software
Familiarity with: Standard web-based software systems including (but not limited to) Unix, Linux, SQL Server, MySQL, SGML, XML, PHP, JAVA, PERL At least one major repository software including (but not limited to) EPrints, DSpace, Fedora. Current Research Information Systems e.g. Symplectic, PURE *Ability to:*

Customise, deploy and manage repository and associated software Arrange and carry out testing of the system and evaluate results Design and develop repository interface and tools
Metadata
Familiarity with: Relevant metadata standards including (but not limited to) Dublin Core, MARC, METS, MODS, OAI-PMH. This includes standards developed for the interoperability of research information such as CERIF. *Ability to:* Identify or develop appropriate metadata and other standards Liaise and test implementation with cataloguing team where appropriate Ensure compliance and monitor metadata quality on an ongoing basis
Storage & Preservation
Familiarity with: Current best practice procedures and external advice and resources *Ability to:* Work with IT Services on the use of their network storage and on backup requirements in the medium and long term Work with institutional personnel including (but not limited to) University Records Manager, Archivist and IT services, as well as external organisations in order to identify best practice and establish requirements for preservation and develop a policy for how different materials should be preserved (or not)
Content
Familiarity with: IPR Issues in order to develop policies and guidelines, to provide advice to users and check deposits from academics.
Ability to: Identify suitable material for the repository and set up automatic downloads of content from bibliographic databases such as Web of Science and from subject repositories such as UKPubMed Central Develop a content policy to include, for example, types of material, handling embargoed material.
Liaison (internal & external)
Ability to: Liaise with a wide variety of departments and interest groups in order to: • Develop strategic plans to ensure ongoing support for and relevance of the repository • Embed the repository in the research processes of the institution and with the library services • To achieve buy-in by IT services into the repository so that they understand the needs of the repository and to ensure the repository is integrated and aligned with other university systems to deliver services • Promote the repository outside the institution as a showcase of the institution's work

• Liaise with external stakeholders in open access and repository development, including (but not limited to) funding agencies; publishers; repository groups or federations; service providers; learned societies; international peers and related organisations
Advocacy, training & Support
Ability to: Develop an advocacy programme to address the full spectrum of stakeholders to create a broad culture of engagement within the institution, including training sessions, departmental presentations Develop advocacy and publicity materials for use within the institution e.g. webpages, guides, FAQs and presentations Be proactive in publicising repository developments via institutional newsletters, seminars and email alerts etc
Current Awareness & Professional Development
Familiarity with: Current trends in the repository community, particularly with respect to events within the UK, through attendance at relevant conferences, meeting and reading relevant mail lists and professional literature Developments within the general research community and the UK higher education system to identify potential implications for the repository. For example research evaluation processes such as the REF are likely to significantly impact the repository. Technical and repository developments through attendance at relevant workshops and training courses
Ability to: Participate (where appropriate) in new developments, best practice, and relevant projects within the repository community Support the repository community through sharing experiences and contributing to community discussions as appropriate

The RSP survey outlined earlier also asked respondents which were the three most important skills for their post. The results do not differ significantly from the 2008 SHERPA guidance.

Fig 6 Key skills – survey results

The text of responses was analysed in Wordle and it is hardly surprising that communication is the most dominant term given that the main role of repository staff is to radically alter the process of scholarly communication, with their own interpersonal skills being the key weapon in their armoury. Many of the other skills listed such as the delivery of training and presentations and liaison with other departments in the university are linked to this. Strategic planning, project management and prioritisation were also common themes as well as accuracy and attention to detail. Many people highlighted the need for perseverance also using terms such as determination, patience and persistence. One respondent summed it up as “aka bloody minded obsession”!.

Professional knowledge was also highlighted in areas such as cataloguing and metadata and IPR (intellectual property rights). Staff needed to be technically literate with a keen awareness of emerging technologies. They also needed in depth knowledge of research communication processes and open access issues, in order to promote their service and engage with academics.

Training and support for repository staff

In 2008, Zuccala, Oppenheim and Dhiensa noted that while the curricula of most schools of library and information science cover the basics of digital library management, “none, to our knowledge, focuses on the particular needs and requirements of repository managers. “. They also asked managers if they had participated in any training before setting up their repositories but few had. More recently, the University of Sheffield has introduced specific project work in their ‘Academic and Research Libraries’ module including a case study on institutional repository developments (Corral, 2010). Support for practice in the workplace is well developed and has been funded by JISC in England and Wales and the Scottish Funding Council in Scotland. The Repositories Support Project began in October 2006 with a remit to increase the number of repositories – there are now 122 in the UK. The second phase began in March 2009 and is due to complete in March 2012. The Welsh Repository Network was set up in April 2007 and has established repositories in all Welsh institutions. ERIS (Enhancing Repository Infrastructure in Scotland) is focused on motivating researchers to deposit their work in repositories. The rest of this section will focus on the work of the Repositories Support Project.

As noted in the already, the number of institutional repositories has grown phenomenally since the beginning of 21st Century and the UK is considered a world leader in this area. The Repositories Support Project is based in the Centre for Research Communication at the University of Nottingham. It consists of a team of three people: a co-ordinator, an events organiser and a technical officer. The objectives are as follows:

- more repositories in higher education institutions in England, Wales and Northern Ireland
- more content in existing repositories
- more types of content in existing repositories
- closer integration of repositories into institutional information systems
- promotion of best practice and standards
- investigation of the new role of institutions in research output curation and access

The methodology for achieving this is summarised in the list below, some of these will be explained in more detail:

- **outreach** activities, including consultancy visits;
- **events**, including workshops and seminars

- **information** gathering, information services and information provision through the website, briefing papers, web advisory documents, podcasts, etc;
- direct **support** through the RSP enquiry service;
- **liaison** with key stakeholders, including researchers and higher education institutions
- **bridging** between repository managers and technical administrators -- and repository developers, standards developers, policy developers and the wider global repository community.

The project offers consultancy to individual institutions. This, by its nature, is tailor made to individual requirements. One typical example might involve a member of the team attending a project steering group, meeting with Directors of Research, advising project sub-groups on IPR and copyright or promotion and advocacy. There have been 25 consultancy visits since January this year (8 month period), mainly to institutions who are setting up repositories. Feedback for these is generally very positive.

Regular free events are arranged for the target community (although a charge is made for residential events). Figure 7 shows the programme for April 2010 to March 2011. The reach (actual and anticipated delegate numbers) for this period is 557.

Figure 7 - RSP events 2010 - 2011

Date	**Event Title**	**Type of event**	**Numbers** (either actual or anticipated)
14th May 2010	Communication skills for Effective Advocacy	Workshop	27
26th May 2010	Preservation for Repository Practitioners, Edinburgh	Workshop	13
27th May 2010	Preservation for Repository Practitioners, Birmingham	Workshop	13
2nd – 4th Jun 2010	RSP Summer School	Residential school	31
11th Aug 2010	Workshop for Repository Administrators	Workshop	33
1st Sep 2010	RoMEO API Workshop	Technical workshop	10
27th Oct 2010	Doing it differently	Conference	70
Nov 2010	ePrints training	Technical workshop	20
10th Dec 2010	Open Access: the impact for university libraries and librarians	Conference	90
Jan 2011	DSpace training	Technical workshop	20
Jan 2011	e-Theses workshop	Workshop	80
9th – 11th Feb 2011	Winter School	Residential school	30
March 2011	Softwares conference and exhibition	Conference	120

As the repository scene matures in the UK, the focus for events has correspondingly changed. Originally, it was very much about technical set up and establishing a presence in the institution. Now issues such as increasing content and forging integration with other institutional initiatives and systems are more prevalent. Increasing content in repositories is one of the major challenges facing managers – estimates of the proportion of full text records vary from 15%-30%. (Harnad, 2009). (Most repositories will accept metadata only records so as to get a complete picture of the institutions research output). Metadata records can be harvested from various sources e.g. Web of Science, the institution's own research database but the full text can usually only be obtained by persuading the

author to deposit. Often this is mediated by the repository team (e.g. checking copyright compliance). The success of this depends very much on the advocacy skills of the repository staff. Many librarians are skilled in dealing with apathy – for example when faced with an information literacy session with new undergraduates. However, the repository staff may well face outright hostility particularly as there are many misconceptions about open access e.g. it will destroy the bedrock of scholarly communication, peer review. Even when the institution makes it a requirement for academics to deposit their work, there is still a need for advocacy. Staff need to be knowledgeable about the evidence and arguments and adept at influencing their audience through understanding their motivations and fears. The Advocacy Workshop referred to in Figure 7 addressed the processes of communication so that delegates became more skilled in dealing with its emotional aspects, investigated the sources of power when influencing people and learned tactics for dealing with resistance. An open discussion session explored the objections to OA commonly faced by participants and discussed the evidence to address these. As a result of the workshop, the RSP set up a closed wiki which captured these objections with corresponding suggestions for dealing with them, backed up with references to the research evidence.

Feedback from event attendees is generally very positive – all events are evaluated systematically and suggestions sought for future topics. Overall rating as Very Good and Good so far this year have achieved the following percentages: Communication skills for Effective Advocacy 100%, Preservation for Repository Practitioners 75%, Summer School 100% and Workshop for Repository Administrators 96%.

Reactive support is a feature of the RSP service with a telephone helpline and an e-mail helpdesk. More recently a Buddy scheme has been launched in June 2010, This scheme enables the RSP to put colleagues in touch with nominated 'buddies' across the country. The buddies on the register have agreed that they would be happy to allow colleagues from other institutions to pay them a visit and talk to them about their work. So far there are nineteen registered buddies and there have been six requests for support.

The RSP website itself is a source of information developed for use by UK higher education institutions in establishing and developing repositories but freely available for anyone to download. It has three main content sections for users: Building Repositories, Expanding Content and Increasing Usage and is also a vehicle for promoting RSP events and news. At the time of writing, it is in the process of being redesigned – a blog has already been developed. The new site will include less content as much of this has been transferred to the JISC Digital Repositories infoKit and more dynamic Web 2.0 type content. The RSP Briefing papers, which are housed on the site are concise introductions to key topics of interest to the community e.g. Handling Version Information, Key Stakeholders and Benefits, which are periodically revised. There are also Planning Checklists, including one for Expanding Content, and Podcasts.

The RSP has a key "bridging" role between staff in individual repositories and the wider community such as software and standards developers. A recent example is the issue of the OpenAire guidelines which outline the Open Access requirements of those in receipt of European Union FP7 funding who are part of the OA pilot. The RSP is analysing the requirements and liaising with repository software providers in order to provide UK specific guidance to individual institutions.

Professional networks

The UK has a well established professional network for repository staff: UKCoRR (United Kingdom Council of Research Repositories. This was founded in May 2007 and is an independent body which:

- Promotes repository management as a recognised and respected profession
- Provides a forum for discussion and exchange of experience
- Represents the views and concerns of those who work with repositories in organisational, policy and strategic development

It has gained a reputation for expertise and excellence and the committee is often asked to comment on key issues in research communications. Membership, which is limited to those working in UK repositories, currently stands at 215 and it has a lively mailing list including robust discussion and

requests for advice. In addition, meetings are held twice a year with speakers on current issues. At present, there is no subscription charge and the committee works on a voluntary basis.

Conclusion

The open access (OA) movement in scholarly communication has grown considerably over the last ten years. More recently there has been an impetus to share online teaching and learning resources which has led to the Open Educational Resources (OER) programme in the United Kingdom. These two developments have driven an increase in the establishment of institutional repositories in UK universities and a subsequent need for their effective management. Many information professionals have moved into these roles and are facing new challenges and developing appropriate skills.

References

Antelman, K. (2004). Do Open-Access articles have a greater research impact? *College and Research Libraries, 65*(5), 372-382. http://eprints.rclis.org/2309/

Bethesda Statement on Open Access Publishing. (2003). http://www.earlham.edu/~peters/fos/bethesda.htm

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. (2003). http://oa.mpg.de/openaccess-berlin/berlindeclaration.html

Budapest Open Access Initiative. (2002). http://www.soros.org/openaccess/read.shtml

CERIF (Common European Research Information Format). http://www.eurocris.org/cerif/introduction/

Corral, S. (2010). Educating the academic librarian as a blended professional: a review and case study *Academic Librarian 2 : Singing in the Rain, ALSR 2010, Conference towards Future Possibilities, Hong Kong, 11-12 March 2010, conference proceedings, session 1A, p. [1-24]*. http://repository.lib.polyu.edu.hk/jspui/bitstream/10397/1731/1/Session1A_Corrall.pdf

Enhancing Repository Infrastructure in Scotland (ERIS). http://www.jisc.ac.uk/whatwedo/programmes/inf11/sue2/eris

Garritano, J.R. and Carlson, J.R. (2009) A subject librarian's guide to collaborating on e-science projects. *Issues in Science and Technology Librarianship, 57 (Spring 2009).* http://www.istl.org/09-spring/refereed2.html

Hajjem, C., Harnad, S. and Gingras, Y. (2005). Ten-year cross-disciplinary comparison of the growth of Open Access and how it increases research citation impact. *IEEE Data Engineering Bulletin, 28*(4). http://sites.computer.org/debull/A05dec/hajjem.pdf

Harnad, S. (2009) DEBATE: Institutional repository success is dependent upon mandates. Bulletin of the American Society for Information Science and Technology, 5 (4). http://eprints.ecs.soton.ac.uk/18035/

Heery, R. (2009). Digital Repositories Roadmap review: towards a vision for research and learning in 2013. *JISC 2009* http://www.jisc.ac.uk/whatwedo/themes/informationenvironment/reproadmaprev.aspx

JISC Digital Repositories infoKit. http://www.jiscinfonet.ac.uk/infokits/repositories/index_html

JISC Recruitment Toolkit.
http://www.jisc.ac.uk/whatwedo/themes/informationenvironment/recruitment.aspx

Kennan, M.A. and Kingsley, D.A. (2009). The state of the nation: a snapshot of Australian institutional repositories. *First Monday 14(2).*
http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2282/2092#p3

OpenAIRE Open Access Infrastructure for Research in Europe. http://www.openaire.eu/

OpenDOAR Directory of Open Access Repositories http://www.opendoar.org

Ottaviani, J. And Hank, C. (2009). Libraries should lead the institutional repository initiative and development at their institutions. *Bulletin of the American Society for Information Science and Technology, 35(4).* http://onlinelibrary.wiley.com/doi/10.1002/bult.2009.1720350408/full

Repositories Support Project. www.rsp.ac.uk

Repositories Support Project Blog. http://rspproject.wordpress.com

Research Councils UK. Pathways to Impact. http://impacts.rcuk.ac.uk/default.htm

Research Information Network. (2010). Getting your feet wet: an introduction to Open Access.
http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/introduction-open-access

Scientific Electronic Library Online. (SciELO). http://www.scielo.org/php/index.php?lang=en

SHERPA (2008). Staff and skills set. http://www.sherpa.ac.uk/news/staffandskillsrevision.html

Swan, A. (2010). The Open Access citation advantage: studies and results to date. Technical Report, School of Electronics & Computer Science, University of Southampton.
http://eprints.ecs.soton.ac.uk/18516/

Swan, A. (2011) Institutional repositories - now and next. In: *University Libraries and Digital Learning Environments (eds Penny Dale, Jill Beard and Matt Holland)*, Ashgate Publishing. (In Press)
http://eprints.ecs.soton.ac.uk/21471/

United Kingdom Council of Research Repositories. http://www.ukcorr.org/

University of Glasgow. Enrich project. http://www.gla.ac.uk/enrich/

Walters, T.O. (2007). Reinventing the library – How repositories are causing librarians to rethink their professional roles. *Portal: Libraries and the Academy, 7(2).*
http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v007/7.2walters.html

Welsh Repository Network. http://www.wrn.aber.ac.uk/en/

Zuccala, A, Oppenheim, C and Dhiensa, R. (2008). Managing and evaluating digital repositories. *Information Research 13(1).* http://informationr.net/ir/13-1/paper333.html