Researchers' green Open Access practice: a cross-disciplinary analysis

Introduction

The EC-funded Publishing and the Ecology of European Research (PEER) project (http://www.peerproject.eu/) is an unprecedented collaboration between stakeholders involved in scholarly research and scholarly publishing, looking into specific aspects of the complex relationship between Open Access (OA) and scholarly communication. The project includes publishers and representatives, libraries, open access repositories (OARs) and researchers themselves, both as authors of journal articles and as readers (or consumers) of journal literature.

The behavioural research is one of three strands that form the PEER Observatory, which was set up to investigate the effects of the large scale deposit of stage-two manuscripts (also known as authors' final peer-reviewed and accepted manuscripts) on reader access, journal visibility and viability, and the broader ecology of European research (Shepherd & Wallace, 2009). The PEER behavioural research project was carried out in two phases, between April and September 2009, and from September 2010 to August 2011. Researchers at Loughborough University examined the behavioural aspects affecting self-archiving of stage-two manuscripts in OARs as well as the use of these manuscripts by fellow researchers.

Most discussion of Open Access recognises the two main mechanisms to achieving open access. The gold route, often referred to as the 'author pays' route, involves payment of an article processing charge to publishers enabling the article to be made available to all without subscription or charge barriers. The alternative green route, often referred to as the 'self archiving' route, entails authors submitting manuscripts to traditional journals but maintaining the right to mount a version of their work on an open access repository. Much debate has focussed on the most effective way to achieve Open Access. There are many advocates of the green self archiving route to OA; subject-based repositories containing both stage-two manuscripts and preprints of research articles are a widely accepted development in certain disciplines such as physics and economics. Alongside this, many institutions are developing their own open access repositories and some are mandating deposit into these respositories. ROARMAP (http://roarmap.eprints.org/) and OpenDOAR (http://www.opendoar.org/) outline the extent of these developments worldwide. On the other hand, the recent report by the Working Group on Expanding Access to Published Research Findings, chaired by Professor Dame Janet Finch (Working Group on Expanding Access to Published Research Findings, 2012) recommended developing clear policies in support of publication in Open Access journals. In recommending gold OA, the Finch report requests that repositories carefully consider the balance between the aims of open access and possible risks 'to the sustainability of subscription-based journals during what is likely to be a lengthy transition to open access', however the report does recommend the continual development of the infrastructure surrounding subject and institutional repositories, though primarily for the purposes of research data and grey literature. In the lights of current policy developments in favour of the different routes to Open Access, a cross disciplinary analysis of researcher's views and attitudes towards green (self archiving) OA practice is timely.

Based on Phase 2 of the project, this article extends the preliminary results from phase 1 reported in Creaser et al (2010), and further develops understanding of researchers' green OA experience both as authors and readers of peer-reviewed journal articles by looking in greater detail at their reported use of OARs and the context of that use. The article identifies disciplinary patterns of behaviour at the level of the Medical sciences, Life sciences, Physical sciences & mathematics, and Social sciences, humanities & arts in order to better understand the role of OARs in the scholarly communication landscape.

Background

Levels of awareness of Open Access and open access repositories

Creaser et al. (2010) found that levels of awareness with regard to the concept of OA are growing compared with previous seminal studies conducted by Rowlands et al. (2004) and Swan and Brown (2004, 2005). Creaser et al. (2010) report that two-thirds of respondents indicated that they understood the definition of OA at least in its most simple form, i.e. free electronic access to full text articles. They also reported that researchers from the Social sciences, humanities & arts were more likely to be unsure about the meaning of OA. It is reasonable to anticipate that levels of awareness of OA will grow more rapidly in the future as OA further permeates the policy sphere as reflected in initiatives and debates such as the European Commission's OA Pilot in the Seventh Framework Programme (EC, no date), Harvard academics' reported stance of maintaining copyright in their publications (Harvard University Library, 2010) and, more recently, the UK Government's 'Innovation and research strategy for growth' report (Dep. for BIS, 2011), which adopts a clear position for free access to publicly funded research and gives an indication that UK public funder mandates will be enforced more stringently in the near future (ibid, p.77).

Although there has undeniably been an increase in levels of awareness in recent years, it is particularly difficult to gauge the extent to which researchers are aware of OA. In this respect, Fry et al. (2009) identified a discrepancy between reported levels of awareness amongst their survey respondents (who were invited to choose responses from a selection of possible meanings) and what researchers understood OA to mean when they were prompted to give examples of OA resources, with respondents often showing some confusion between free electronic access via OA resources and seamless desktop access to subscribed journal resources through an institution's single sign-in access management system. Confusion surrounding the implementation of and engagement with OA may well be exacerbated by the combination of an ever-increasing reliance on Internet search engines to fulfil researchers' information needs, and the emergence of information systems operating meta-searches across a wide range of information resources, including those beyond an individual library's subscription package.

Consequently, self-reported levels of awareness can be difficult to interpret as they are an expression of researchers' perceptions and may be based on misunderstanding and misconceptions. Creaser et al. (2010) addressed this issue and found that researchers' awareness may in fact be lower than that reported in the literature. This was supported by the finding that not many researchers associated OA with either self-archiving, also known as Green OA (6% of survey respondents) , or open access publishing, also known as Gold OA (11%) (ibid.). This resonates with previous studies reporting that OARs are generally not on researchers' radar and do not enjoy high visibility (Harnad et al., 2009).

As for whether differences in levels of awareness exist on a disciplinary level, Creaser et al. (2010) found that researchers across different disciplinary groupings shared a similar understanding of OA. Significant disciplinary differences in terms of the associations researchers make between OA and the scholarly communication system were identified, however (Creaser et al., 2010). Researchers in the Life and Medical sciences were more likely to associate OA with the 'author-pays' model of Open Access Journals (OAJ), rather than self-archiving via OARs (*ibid*). This is echoed by findings from recent studies indicating that the 'author-pays' model, or Gold OA, is widespread and well developed in the Life and Medical sciences, whereas self-archiving, or Green OA, is much more developed in disciplines such as Earth sciences, Physics and Astronomy (Björk *et al.*, 2010). The fact that well-known OAJ, such as those from PLoS (Public Library of Science), are prominent and enjoy high visibility amongst research communities in the Life and Medical sciences may reinforce or strengthen the close association of OA with OA publishing within those disciplines. In their study, Dallmeier-Tiessen *et al* (2010) found that STM (Science, Technology and Medicine) represents 66% of the body of pure and hybrid OAJ, and contributes 77% of articles. Within this, biology and life sciences represent 19% of OAJ and 21% of articles, while medicine and health sciences represent 28% of OAJ and 28% of articles (ibid). It is perhaps understandable, therefore, that Creaser et al. (2010) found researchers in the Medical

and Life sciences to be less likely to associate OA with content which is 'not-peer-reviewed', and more likely to associate OA articles with the published version.

In terms of the motivation to self-archive, Creaser et al (2010) found that wide-spread dissemination was the most important reason for depositing articles in an OAR, irrespective of the type of repository. Other important motivations included greater visibility, increased citation rates and reputation building, with researchers in the Social sciences, humanities & arts more likely than those in other disciplines to consider increased citation rates to be very important.

With regards to the differences between article versions, there is evidence that the main distinction between a stage-two manuscript and the final published version, i.e. the differences observed before and after publisher's copy-editing, lies in the correction of spelling, grammar or style mistakes, though a certain number of corrections also deal with mistakes spotted in the references or with missing references. Based on their respective studies, Thatcher (2011), Wates and Campbell (2007) and Goodman et al. (2007) conclude that the quality of stage-two manuscripts is probably good enough for certain types of academic use, such as for general reading, though the question of the propagation of erroneous references may cause subsequent problems. The immediate implication for users of articles found in OARs is that the two versions (stage-two manuscript and final published version) can be difficult to distinguish at first glance, leaving readers reliant on the publisher's branding to identify which version of an article they have accessed. Publisher branding may, on some occasions, be overlooked by readers, or be missing from the final published version itself, which may make it difficult for readers to clearly identify the version. It is also possible that readers never see the repository metadata page (which often outlines the version of the article being accessed) - there is some evidence (reported by repository download statistics) that readers may be directed from an Internet search engine directly to the full text document, despite repositories' efforts to bring their readers to the metadata page first (Organ 2006).

Researchers' scholarly communication practice

Despite novel ways of disseminating and communicating research findings brought about by digital and networked scholarship, there is strong evidence that the published journal article remains the preferred way for researchers to disseminate their research (Research Information Network, 2009). Fry et al. (2009) characterised the central role of the peer-reviewed journal article in scholarly communication and in scholars' research practice, with over 90% of respondents in their survey rating peer-reviewed journals as 'very important' to their research. These findings were in line with studies published by Tenopir et al*.,* (2009) and King et al*.,* (2009) which both found that scholarly articles accounted for over 90% of scholars' information sources, and have been further corroborated by more recent studies (Harley et al., 2010).

Scholarly communication practices vary both within and across disciplines. The culture of a discipline (fast-moving, collaboration with the private/public sector, its research problems and methods etc.) and its norms (or traditions) impact strongly on researchers' scholarly communication practices (Harley et al., 2010). Fry et al. (2009) show that journal articles were found to be rated as 'very important' by 95% of researchers in the Life sciences and 86% in the Social sciences, humanities & arts. On the other hand researchers from the Physical sciences & mathematics and from the Social sciences, humanities & arts also highly valued the use of monographs in their information search process. Conference proceedings were more likely to be used by researchers from Physical sciences & mathematics than by other disciplinary groupings. Furthermore, disciplinary variations in the volume of reading were also highlighted, with Medical and Life sciences researchers reading more articles on average than researchers in the Physical sciences & mathematics or the Social sciences, humanities & arts groups (*ibid)*, which corroborates Tenopir et al. (2011)'s findings.

Fry et al. (2009) also explored other types of research outputs as a means of research dissemination and showed that there were statistically significant differences between disciplines in the relative importance of various other types of output when considered as resources for research. For example, Social sciences, humanities & arts and

Physical scientists & mathematicians considered monographs to have greater importance than did Life scientists and Medical scientists. They also considered conferences to be more important than did researchers from the Social sciences, humanities & arts. Life scientists and Interdisciplinary researchers gave the highest ratings to datasets. Life scientists gave the lowest importance rating to reports, while the Medical scientists rated these more importantly than the other broad disciplinary groups. Social sciences, humanities & arts researchers gave the highest importance ratings to working papers. The importance of professional journals did not differ by discipline; other types of output had insufficient data for formal comparisons to be made.

There is a close interrelationship between dissemination and information behaviours, with choices about how, where and when to disseminate research influencing ways in which researchers seek to locate and access information sources. To obtain a more holistic picture of researchers' perceptions, motivations and behaviours, therefore, it is necessary to view these two processes in relation to one another. Information behaviour theory has highlighted that differences across research cultures also influence ways in which researchers seek, use and avoid information (Hjørland, 2002; Case, 2012; Talja et al, 2007). For example, Hjørland (2002) argues for the importance of adopting a socio-cognitive perspective of information behaviours, whereby the collective social/cultural context in which a researcher is seeking information is taken into account in the development of behavioural models. Hjørland's (*ibid*) typology of epistemic information environments illustrates the relationship between four main epistemic schools; empiricism, rationalism, historicism or pragmatism, and the type of information researchers represented by these schools would consider to be relevant or irrelevant in terms of evaluation criteria. Extending Whitley's (2000) analytic dimensions of the social/cultural and intellectual (cognitive) organization of disciplines Talja et al (2007) found that research group membership, across-fields scattering of literature, and degree of establishment of research area influenced patterns of digital library use. Furthermore, Case (2012) emphasises that information seeking takes place in groups, as well as by individuals, and that sharing of information is common, particularly in certain professional and disciplinary communities such as clinical medicine, thereby providing further evidence of the importance of factors beyond the realms of task, individuals and cognition in influencing patterns of behaviour. Research in the domain-analysis tradition (Hjørland and Albrechtsen, 1995), therefore, highlights the need for sensitivity towards research cultures when developing an understanding of the use, or non-use, of innovation in digital information landscapes, such as OAR.

Fry et al. (2009) highlighted the centrality and prominence of published journal articles in readers' information behaviours. This was further supported by the amount of journal articles read by scholars, with over 95% of their survey respondents reporting that they read, on average, more than 10 peer-reviewed articles per year. Frequency of reading was very high for respondents from the Medical sciences and the Life sciences, with almost half of Medical scientists (49%) and 40% of Life scientists reading more than 100 peer-reviewed articles per year, compared to just over one quarter (28%) of Physical scientists & mathematicians and 16% of researchers from the Social sciences, humanities & arts. Researchers from the Social sciences, humanities & arts were more likely to read between 51 and 100 articles per year. This resonates with the recent study by Tenopir et al. (2011) which shows that access, reading and citing of journal articles are regarded as central to the process of writing grant proposals, grant reports and peer-reviewed journal articles. The authors argue that academics show they value journal articles by the amount of time they spend finding, accessing and reading those journal articles, with the reading aspect taking a greater portion of time whilst the search and access component steps are reported to take less time than before. Looking at researchers' citation practice, Tenopir et al. (2011) argue that citations are central to researchers' work, from grant proposal to journal article, with almost 88% of researchers surveyed reporting that they usually cite 10 or more references when they write a journal article, and that researchers across disciplines read at least 20 articles on average for each article cited. Tenopir et al. (2011) also report that, in line with previous studies (Tenopir, King & Bush, 2004), researchers in the health sciences tend to rely more on journal articles than any other disciplines.

Methodology

Phase 2 of the PEER behavioural research comprised an electronic survey of European authors, conducted between January and March 2011, followed by a one-day participatory workshop drawing European researchers from selected disciplinary areas.

The survey

Invitations to complete the survey were sent out via the participating publishers. The initial intention was to restrict the distribution of the survey to EU-based corresponding authors who had published in the journals included in the PEER Observatory and the control group since it commenced in September 2008. This time scale was set to reduce the potential number of currently inactive researchers who might be approached. Restricting the circulation to EU authors was not always possible, and the survey was distributed more widely by several publishers; non-EU respondents were filtered out prior to analysis.

One of the limitations of the research is that it was impossible for the research team to gauge the representativeness of the survey compared to the actual disciplinary distribution of the population of EU researchers, as such population information is not available. The immediate implication of this limitation is that it is only possible to generalize the overall results to the wider population of EU researchers under the assumption that the initial list of authors contacted was representative of that wider EU population. In order to ensure that the disciplinary distribution of survey responses was an accurate reflection of the initial sample, publishers were asked to supply the total number of authors contacted at the journal level. The broad disciplinary groupings used throughout the research were based on the categorisation of journal titles included in the PEER Observatory (Physical sciences & mathematics; Social sciences, humanities & arts; Life sciences; and Medical sciences)[1]. The disciplinary distribution of the sample could therefore be calculated by the research team, and, by comparing this to the disciplinary distribution of respondents, a weighting scheme was devised. The data reported in this article are weighted according to discipline; the impact on comparisons between disciplines is minimal, however.

Disciplinary differences observed during analysis of the survey responses were tested using the χ^2 statistic. Results are noted in this report as being statistically significant when the probability that the observed differences between the disciplines occurred by pure chance, and that in the wider population no such differences exist, is less than 5% ($p<0.05$).

Demographics of survey respondents

A total of 1,427 valid responses were received and analysed. The 34 disciplines reported by survey respondents were mapped onto one of four broad disciplinary groupings derived from the PEER Observatory categorisation of journal titles as shown in table 1, including an interdisciplinary group for those respondents who had selected disciplines from two or more broad disciplinary groupings. (The numbers of respondents after weighting are shown in brackets; in broad terms, Medical scientists were under-represented in the sample, with Life scientists over-represented)

Table 1: mapping of disciplines onto the four PEER journal categories

Broad disciplinary grouping	Disciplines included:	No. of respondents
Medical sciences	Clinical medicine; Clinical dentistry; Anatomy & physiology; Nursing & paramedical studies; Health & community studies; Pharmacy & pharmacology	194 (258)
Life sciences	Biosciences; Psychology & behavioural sciences; Earth, marine & environmental sciences; Veterinary science; Agriculture & forestry	311 (257)

[1] Details of the selection and categorisation of journal titles can be found in D9.4 PEER Annual Report – Year 1, page 8. Available at http://www.peerproject.eu/fileadmin/media/reports/20090928_PEER_D9_4_annual_public_report_final.pdf <accessed 9th July 2012>

Physical sciences & mathematics	Chemistry; Physics; Mathematics; General engineering; Chemical engineering; Mineral, metallurgy & materials engineering; Civil engineering; Electrical, electronic & computer engineering; Mechanical, aero & production engineering; Information technology & systems sciences & computer software engineering	454 (424)
Social sciences, humanities & arts	Architecture, built environment & planning; Catering & hospitality management; Business & management studies; Economics; Geography; Social studies; Media studies; Humanities & language based studies; History; Archaeology Modern languages; Design & creative arts; Education & Sports	167 (172)
Interdisciplinary	Researchers with interests in disciplines spanning two or more of the above broad groupings.	300 (306)
Total[2]		1,426 (1,416)

Respondents were also asked to indicate the type of institutions they were affiliated to (Table 2) and for how long they had been doing research (Table 3). Again, the numbers of respondents after weighting are given in brackets.

Table 2: Distribution of survey respondents by type of institution

Institution type		
University or College	922	(895)
Hospital or medical school	145	(174)
Research Institute	265	(249)
Industry or commercial	29	(36)
Government	36	(30)
Other	30	(32)
Total	1,427	(1,416)

Table 3: For how long have you been conducting research?

Fewer than 3 years	46	(48)
3-5 years	212	(212)
6-9 years	251	(251)
10-14 years	277	(277)
15-24 years	327	(325)
25 years or longer	307	(297)
Not applicable	4	(4)
Total	1,424	(1,413)

Participatory workshop

Twenty two participants were recruited from across Europe. Invitations to participate in the workshop were initially sent to over 700 survey respondents from selected disciplines who had expressed an interest in being contacted further in relation to this research. The final composition of the workshop was: six participants from Earth, marine & environmental sciences; five from Chemistry; and six from Engineering. Within the Social sciences, humanities & arts, there were insufficient potential participants from any single discipline, so a more general approach was taken to

[2] One respondent did not provide discipline information.

ensure this broad disciplinary group was represented, and five participants were recruited. There were insufficient volunteers from disciplines in the Medical sciences to include this broad disciplinary group in the workshop. Full details of workshop participants are given in the PEER behavioural research Final Report (Fry et al., 2011).

Findings

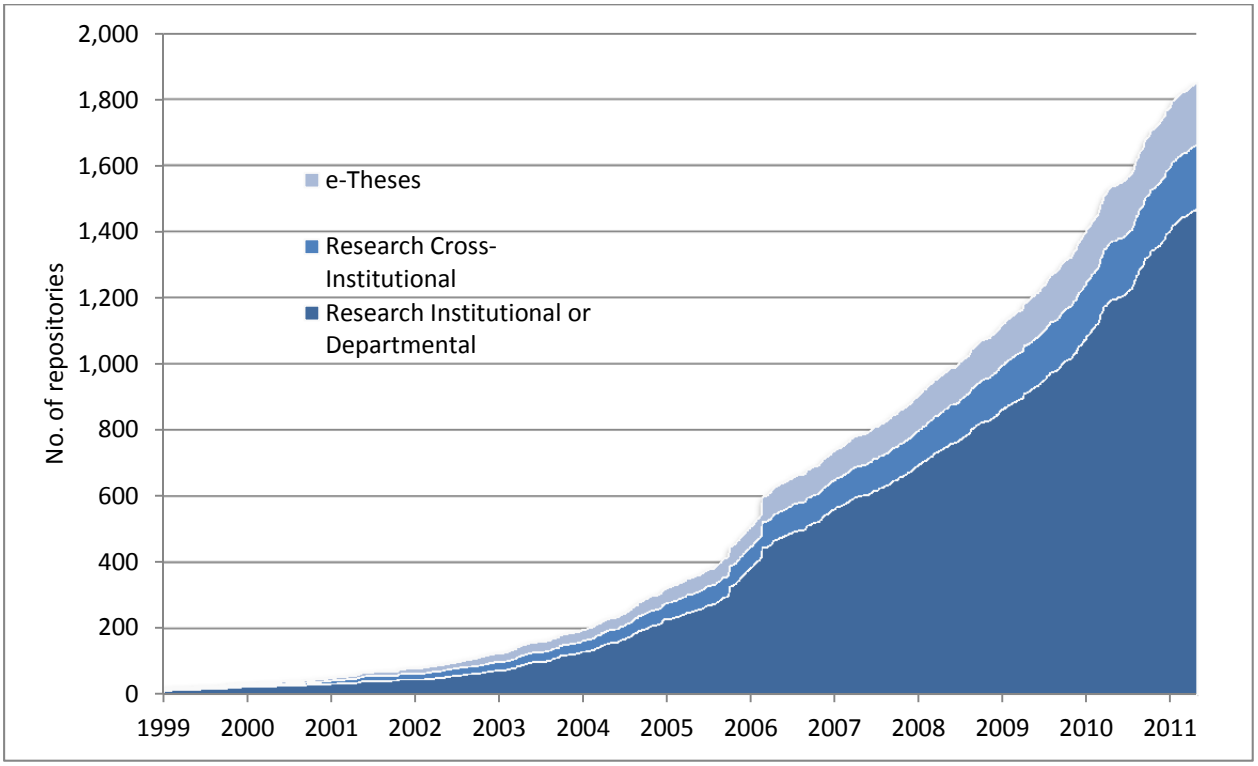
This section reports a disciplinary focussed analysis of the findings from Phase 2 of the PEER behavioural research. It builds on initial findings, outlined above, reported in Creaser et al (2010) and the Baseline Report (Fry et al, 2009).

Self-archiving behaviours

Levels of awareness

General levels of awareness of OARs have not really increased in the 18-month period between the two PEER surveys, despite a growing presence of the OA debate within research communities at an international level, notably with research funders increasingly requesting that the outputs of funded research be made available through OA, and a growing number of institutions setting up institutional repositories (Figure 1).

Figure 1 Growth of institutional repositories internationally

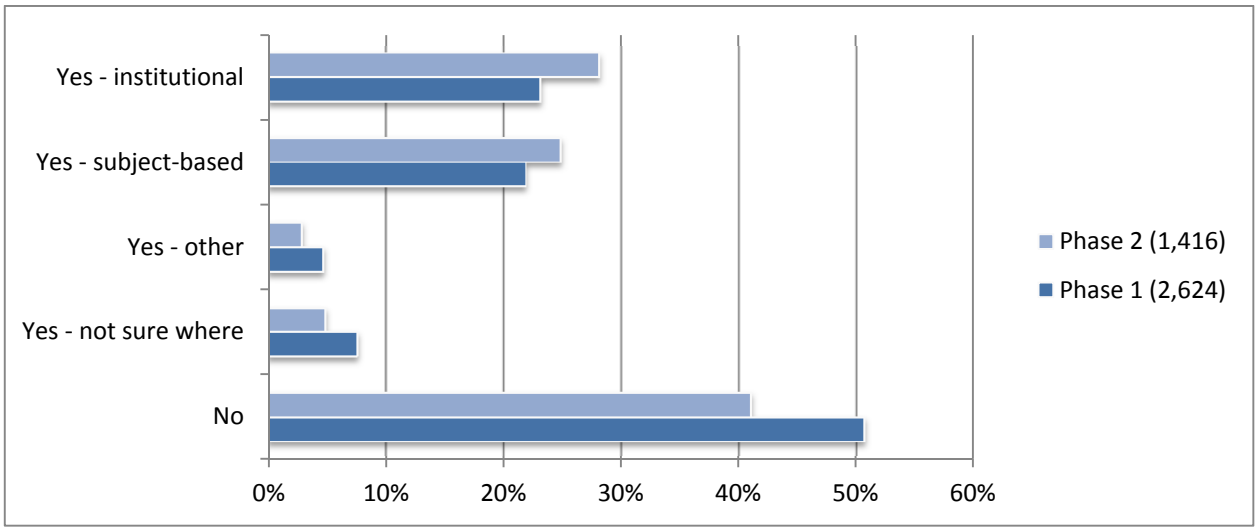


Source: data supplied by ROAR (July 2011)

Deposit rates

Over half of the authors surveyed reported having placed, or having had placed on their behalf, a version of their journal article in an OAR in the last five years, with an increase in the self-reported deposit rates from 53% in Phase 1 to 59% in Phase 2, as shown in Figure 2. Throughout this article, results relating to the self-archiving of journal articles include both authors who made the deposit themselves and those who had someone else make the deposit on their behalf, unless stated otherwise.

Figure 2 Have your peer-reviewed articles been self-archived?

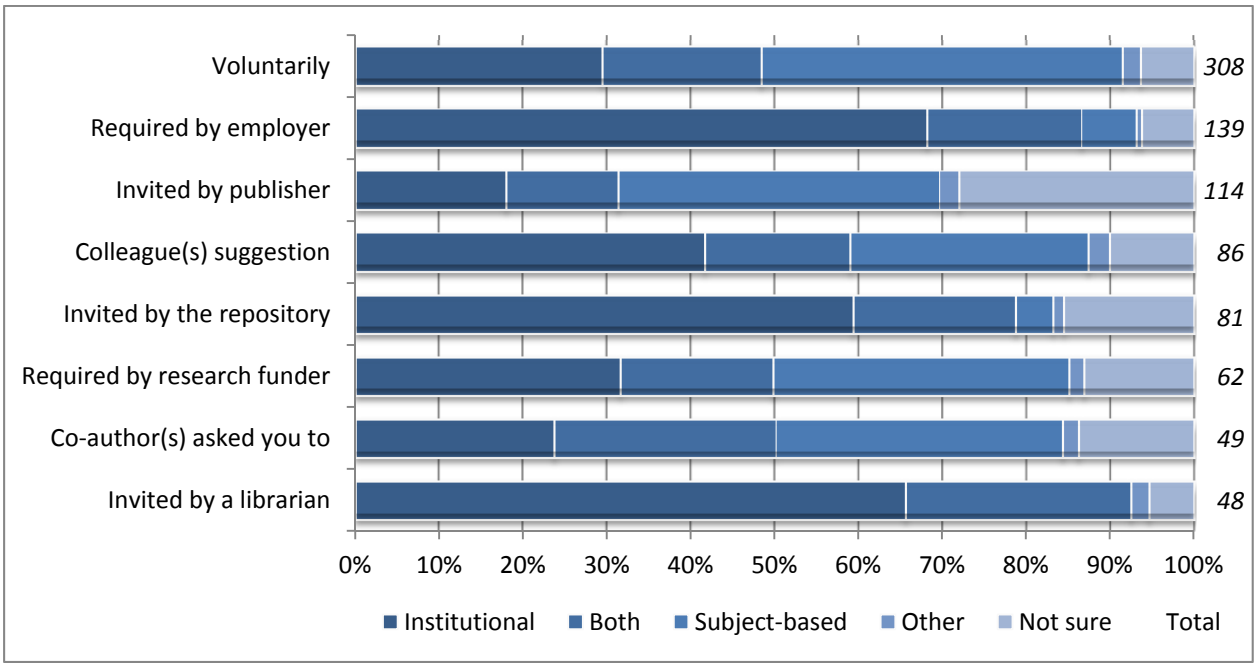


(Weighted numbers of responses given in brackets)

Overall, researchers who reported having deposited one of their journal articles in an OAR in the last five years showed a slight preference for institutional repositories over subject-based repositories (Figure 2). There were disciplinary differences, with respondents from the Physical sciences & mathematics most likely to deposit material into subject-based repositories, whilst those from the other broad disciplinary groups reported being more likely to deposit material into institutional repositories.

The predominant motivation for self-archiving was reaching the target audience(s) and wide-spread dissemination of research, but the research also aimed to gauge the influence of peers as well as that of disciplinary research culture, on researchers' self-archiving behaviour. Motivating factors can be broadly distinguished between self-initiated and externally initiated. Of those researchers who had self-archived a version of their article(s) in an OAR, 46% had done so voluntarily, 20% were responding to an employer's requirement, 16% were responding to an invitation from a publisher, 12% had responded to a suggestion from a colleague, and 11% had responded to an invitation from the repository itself. A higher percentage (70%) of researchers reported having self-archived voluntarily in the phase 1 survey, than in the phase 2 survey. It is unclear why this percentage has fallen; in both surveys respondents could indicate multiple motivations, so one possible explanation may be that other motivations have become more prominent in authors' minds, reducing the likelihood of considering the deposit to have been made voluntarily. The nature of the motivation to self-archive appears to be related to the type of OAR into which authors are likely to place a version of their article(s). Figure 3 shows that if the motivation is voluntary, requested by a co-author, or an invitation from a publisher, then authors are more likely to choose a subject-based repository. If, on the other hand, authors are invited by repository/library staff, or were mandated by their employer, they are more likely to choose an institutional repository. Respondents to the phase 2 survey considered funder and institutional mandates to be relatively unimportant as motivators for self-archiving. The participatory workshop threw some light on this, with researchers reporting that where funder or institutional mandate policies did exist they tended to be insufficiently enforced.

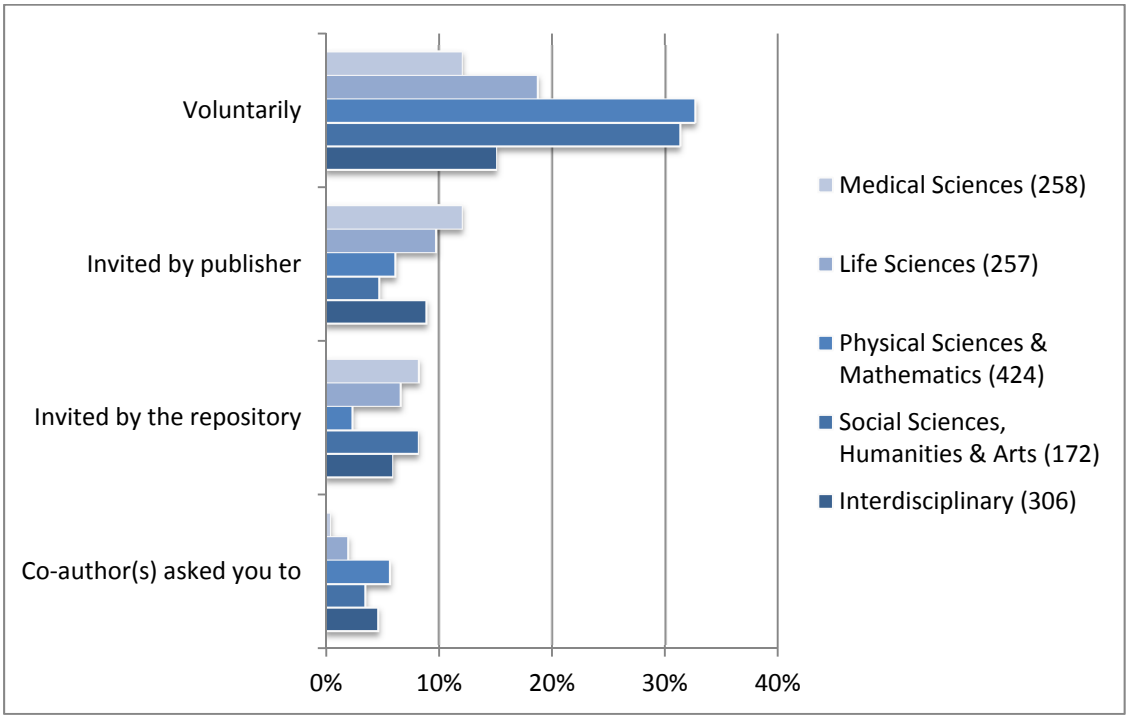
Figure 3 Motivations for repository deposit, by type of repository chosen



(Weighted numbers of responses given in italics)

There were some notable differences in the reported motivations for deposit by discipline, as illustrated in Figure 4, with respondents from the Physical sciences & mathematics and the Social sciences, humanities & arts more likely to have self-archived voluntarily, compared to respondents from the Medical sciences. Respondents from the Medical and Life sciences, on the other hand, were more likely to have been invited by a publisher to self-archive than respondents from the Physical sciences & mathematics or the Social sciences, humanities & arts. Overall, respondents from the Physical sciences & mathematics were least likely to have been invited by a repository to self-archive. A request from co-authors was less likely to prompt respondents from the Medical sciences to self-archive in an OAR than respondents from the Physical sciences & mathematics or Interdisciplinary researchers.

Figure 4 Motivations for repository deposit, by subject



(Weighted numbers of responses given in brackets)

Who deposits?

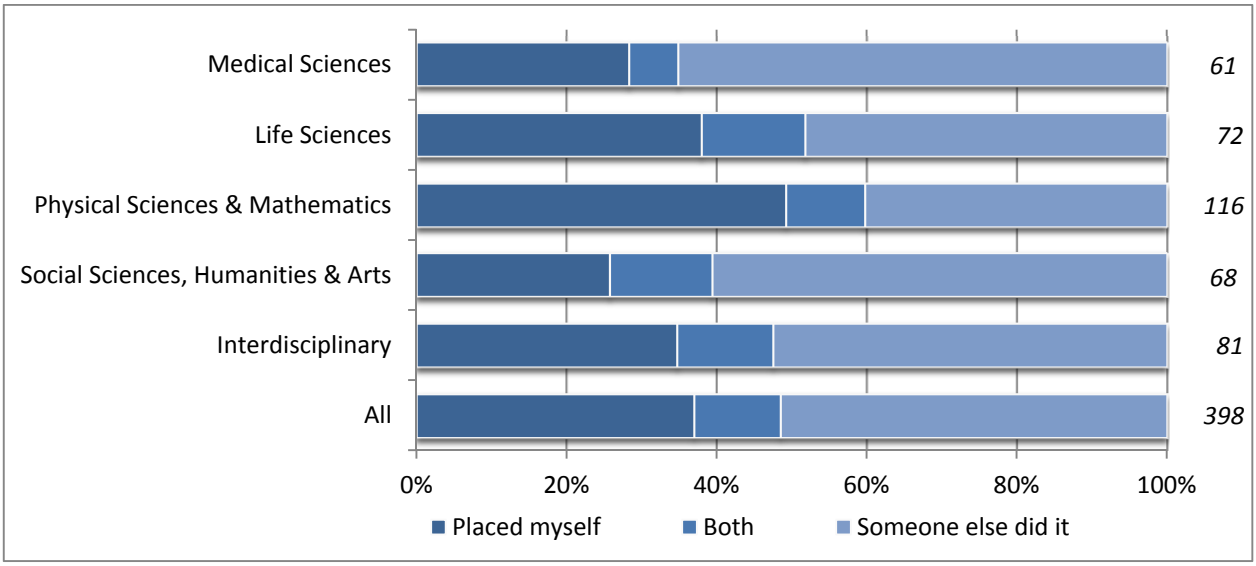
With regard to whether the deposit is made by the author(s) or mediated by somebody else, 75% of researchers who had reported that any version of their published journal article(s) had been self-archived in an OAR in the last five years, reported that it had been placed by someone else, compared to 63% who reported that they had placed it themselves[3].

There were notable disciplinary differences in this pattern, with authors from the Physical sciences & mathematics and Life sciences more likely to place a version of their journal article in an OAR themselves, whilst authors from the Medical sciences and Social sciences, humanities & arts were more likely to have a version of their journal article deposited by someone else. For those who indicated that someone else had placed a version of their article(s) in a repository on their behalf, this was usually done by library/repository staff (36%), colleague or co-author (18%), administrative staff or secretary (16%).

Furthermore, it was found that the choice of repository had only a marginal influence on the deposit process, i.e. whether articles are deposited by the authors themselves or by somebody else. Indeed, amongst the 28% of authors who had deposited an article in an institutional repository, 48% of these reported that they had done so themselves, and amongst the 24% of authors who indicated that they had deposited in a subject-based repository, 51% of these indicated that they had placed an article themselves.

For institutional repositories there were some disciplinary differences in terms of the deposit process being mediated; researchers from the Physical sciences & mathematics and the Life sciences were more likely to have placed the article themselves than respondents from the Medical sciences or the Social sciences, humanities & arts. Researchers from the Medical sciences and the Social sciences, humanities & arts were more likely to have had someone else place their article in an institutional repository than researchers from the Physical sciences & mathematics *(Figure 5).* Disciplinary differences were also apparent in who made the deposit of articles to subject-based repositories, with those respondents from the Physical sciences & mathematics more likely to have placed an article themselves, whilst those from the Life sciences and Interdisciplinary researchers were more likely to have had an article placed in a subject-based repository by someone else *(Figure 6).*

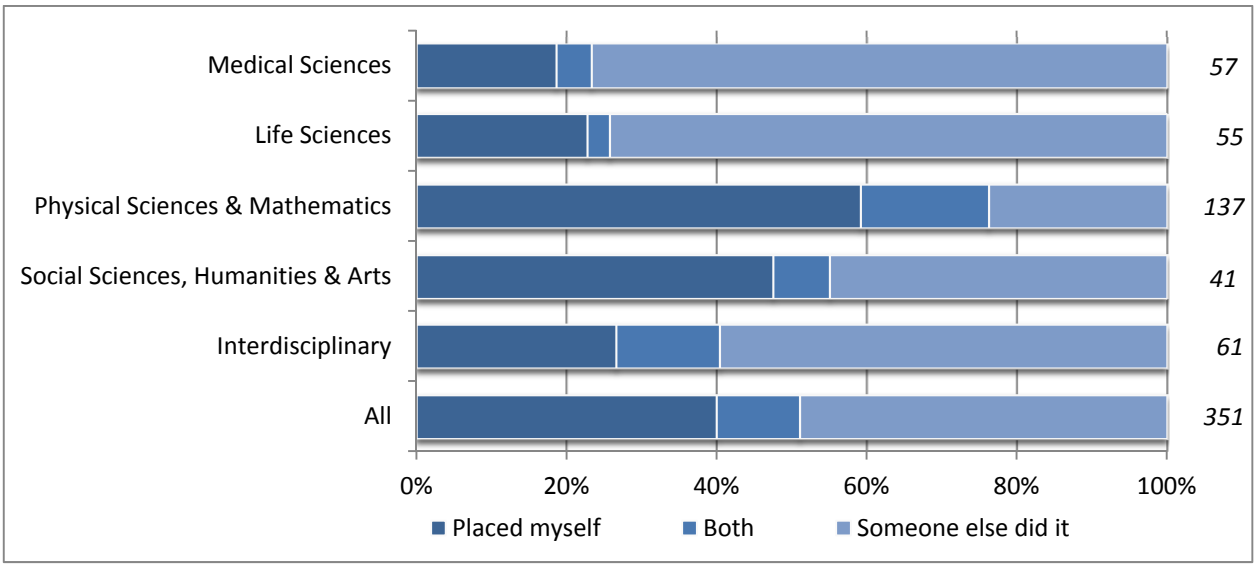
Figure 5 Method of deposit in institutional repositories, by broad discipline group



(Weighted total numbers of responses given in italics.)

[3] Respondents could select multiple answers to this question.

Figure 6 Method of deposit in subject-based repositories, by broad discipline group

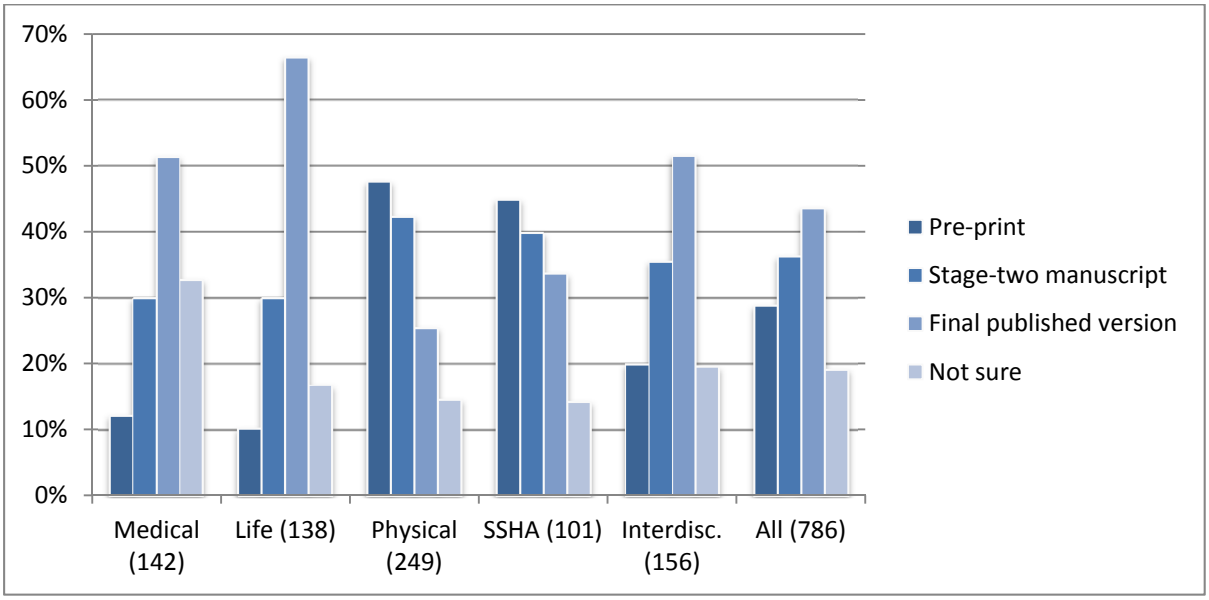


(Weighted total numbers of responses given in italics)

Which versions of articles are self-archived?

The phase 2 survey findings indicated that the final published version was the most frequently deposited (44%), followed by the stage-two manuscript (36%). Preferences relating to the version deposited varied according to broad discipline group (Figure 7); authors from the Life sciences were more likely to have self-archived the final published version than those from the Physical sciences & mathematics or the Social sciences, humanities & arts, whilst authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to have deposited a pre-print or stage-two manuscript, than authors from the Life sciences. Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR, which corresponds with the high proportion of Medical sciences researchers who reported that their article(s) had been placed in an OAR by someone else.

Figure 7 Version of article deposited, by subject

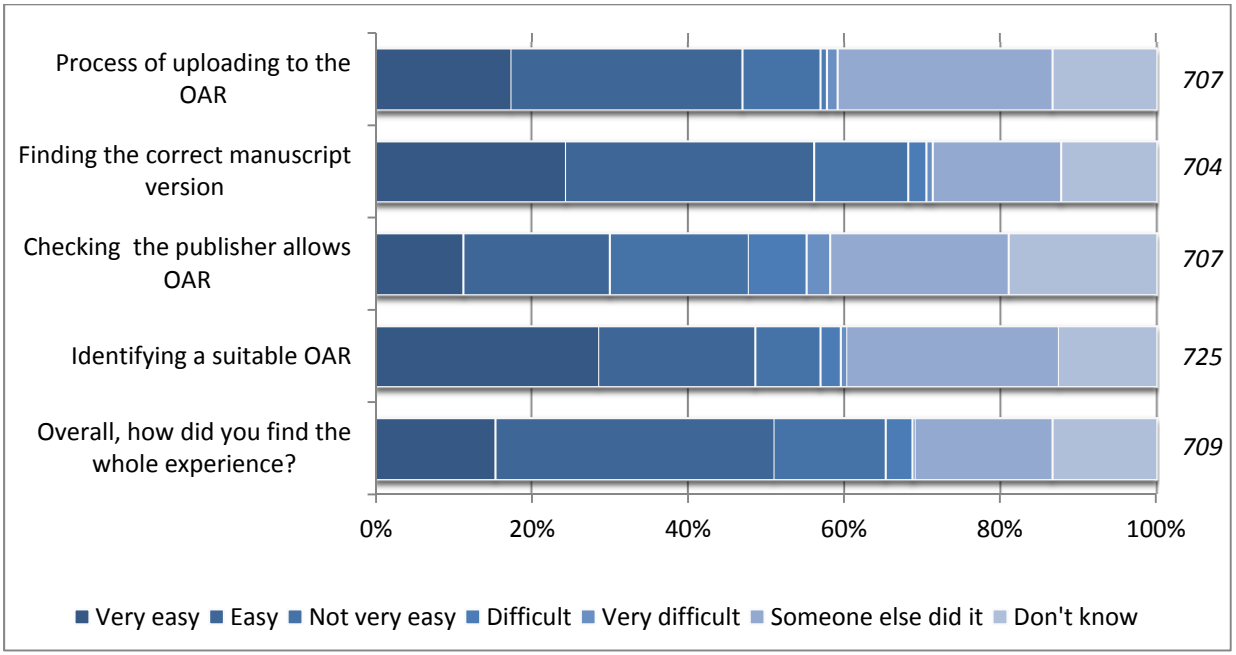


(Weighted total numbers of responses given in italics)

What are the perceived difficulties, if any, in the deposit process?

The survey results suggest that authors do not generally experience much difficulty when depositing a copy of their article(s) in an OAR themselves. Figure 8 illustrates the reported ease of the various steps involved in uploading a copy of their article to an OAR. The overall experience of self-archiving reported was rated as 'very easy' or 'easy' by 51% of authors. Breaking down the process of self-archiving into each component step, authors found that identifying the correct version was the easiest step, with 56% of respondents indicating that this process was 'very easy' or 'easy'. On the other hand, the process of checking that the publisher allows the article to be placed in an OAR proved to be the most difficult step of the deposit process, with less than one third of respondents finding it 'very easy' or 'easy'. Checking publishers' OA policies would appear to be the main barrier to authors managing the self-archiving process themselves. In this respect, the qualitative results also highlighted that authors are often unclear about publishers' OA policies and the nature of any copyright agreements that they have signed, with the perception that they are more constrained in terms of self-archiving than is the case in practice.

Figure 8 Ease of repository deposit procedures

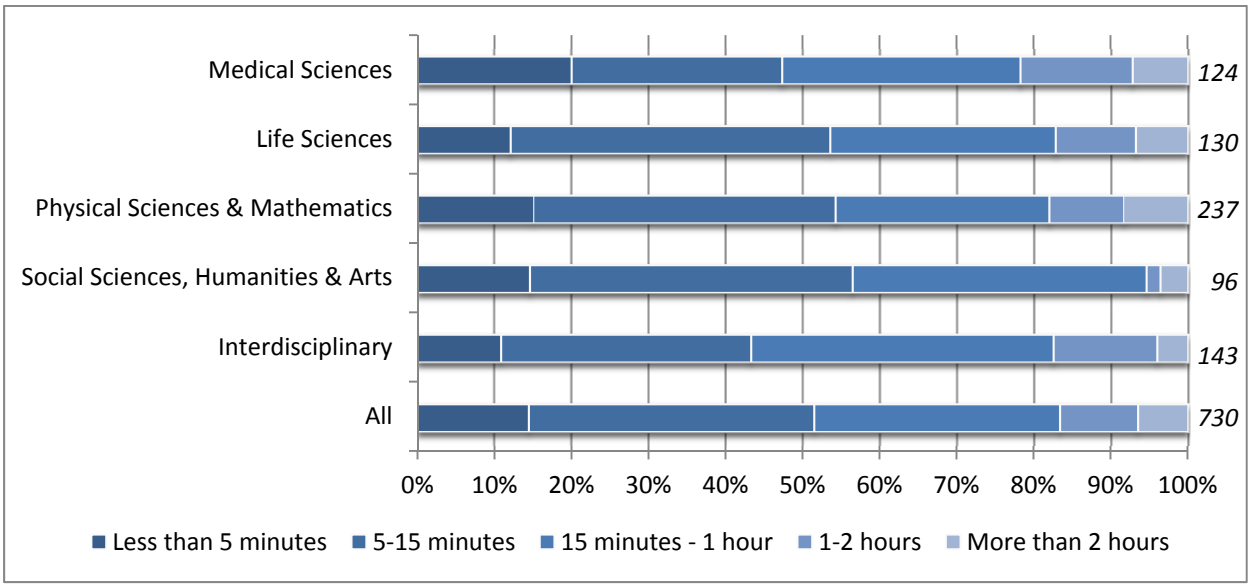


(Weighted total numbers of responses given in italics)

Amongst those authors who had undertaken the repository deposit process themselves, and could rate its difficulty, there were significant differences in experience between the broad discipline groups. Medical scientists were least likely to find it 'easy' or 'very easy' to identify a suitable repository for their articles, but most likely to find checking the publishers' policy on self-archiving to be 'easy' or 'very easy'. Life scientists were most likely to find the process of uploading to the repository to be 'easy' or 'very easy', whilst respondents from the Physical sciences & mathematics were most likely to rate finding the correct manuscript version to be 'easy' or 'very easy'. Respondents from the Social sciences, humanities & arts were most likely to rate finding a suitable repository as 'easy' or 'very easy', but least likely to find checking the publishers' copyright policy, and finding the correct manuscript version to be 'easy' or 'very easy'. Overall, Medical scientists were least likely to rate the whole experience as 'easy' or 'very easy', whilst respondents from the Social sciences humanities & arts were most likely to do so.

The amount of time required to place an article in an OAR, as estimated by authors, varied a great deal (Figure 9), with the most frequently reported estimates being between 5-15 minutes and 15 minutes to one hour. The process of self-archiving was explored in more detail in the participatory workshop, being described by some participants as 'tedious' and 'time-consuming', with the need to check and resolve potential copyright issues adding a layer of complexity, and was perceived as a significant barrier.

Figure 9 Time taken to deposit, by subject



(Weighted total numbers of responses given in italics)

Some disciplinary differences emerged with regard to the time taken to deposit an article version into an OAR. Researchers from the Physical sciences & mathematics and the Social sciences, humanities & arts, were more likely to report that the process took 5-15 minutes, while respondents from the Social sciences, humanities & arts were least likely to indicate that the deposit process took more than one hour (Figure 9).

Readers' use of open access repositories

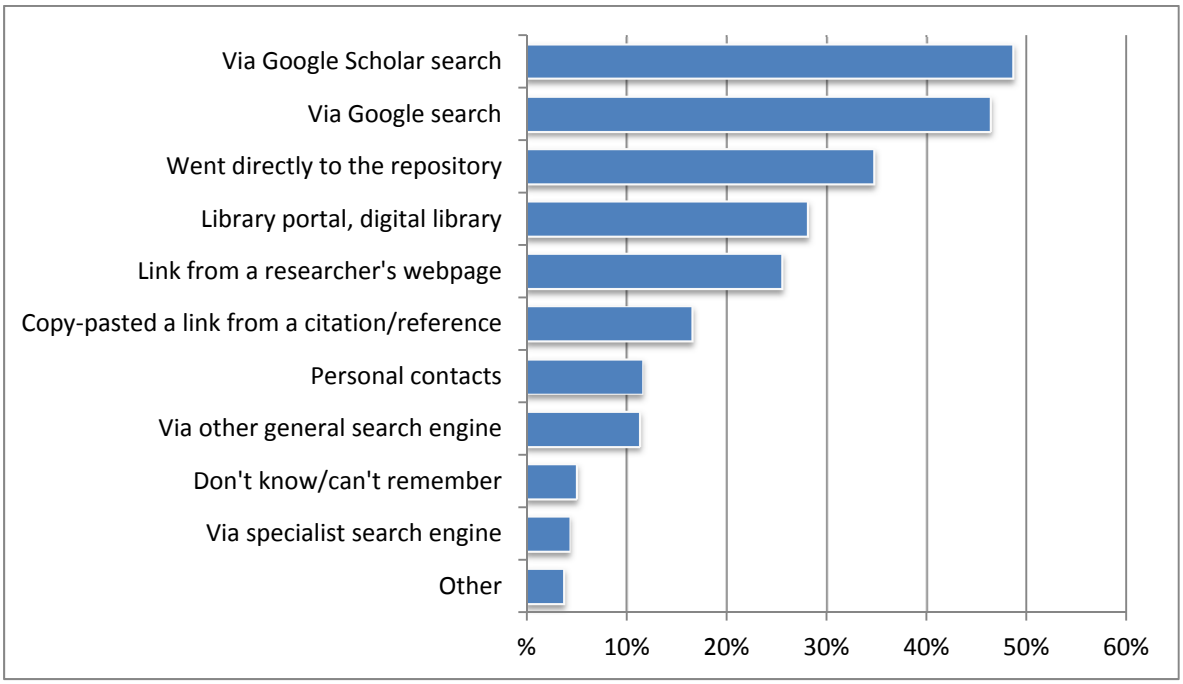
Just over two-thirds of researchers had accessed journal articles held in an OAR in the last year, with 16% stating that they had not accessed such articles, and a further 16% unsure whether they had or not. The research, via the survey and participatory workshop, focused on understanding the motivations and processes supporting each component step of researchers' information behaviours, i.e. how they arrive at the information source, the purpose for which they want the information, and whether the information source found fulfils their information needs.

How do readers arrive at open access repositories?

The research indicated that, regardless of discipline, researchers tend to use a narrow range of information resources, and their strategies are highly individualised – although there is evidence in the literature that for some disciplines, such as the Biomedical sciences, information search strategies follow a well-defined and structured pattern (Blake and Pratt, 2002).

Our findings indicate that researchers were generally unlikely to go directly to an OAR, other than those in disciplines or sub-disciplines served by well-established subject-based repositories such as PubMed Central, arXiv (an e-print service for physics, mathematics, computer science and related fields), RePEc (Research Papers in Economics) or SSRN (Social Science Research Network). These disciplines and sub-disciplines aside, searching specifically for OA resources did not appear to be a feature of researchers' information seeking behaviours. Various ways of accessing OARs were reported, with the most likely route being Google or Google Scholar. Almost half of researchers indicated that they had located an OAR in that way, whilst only 35% reported that they had gone directly to a repository (Figure 10). Less frequently reported ways of accessing OARs included specialist search engines, with just 4% of researchers using these.

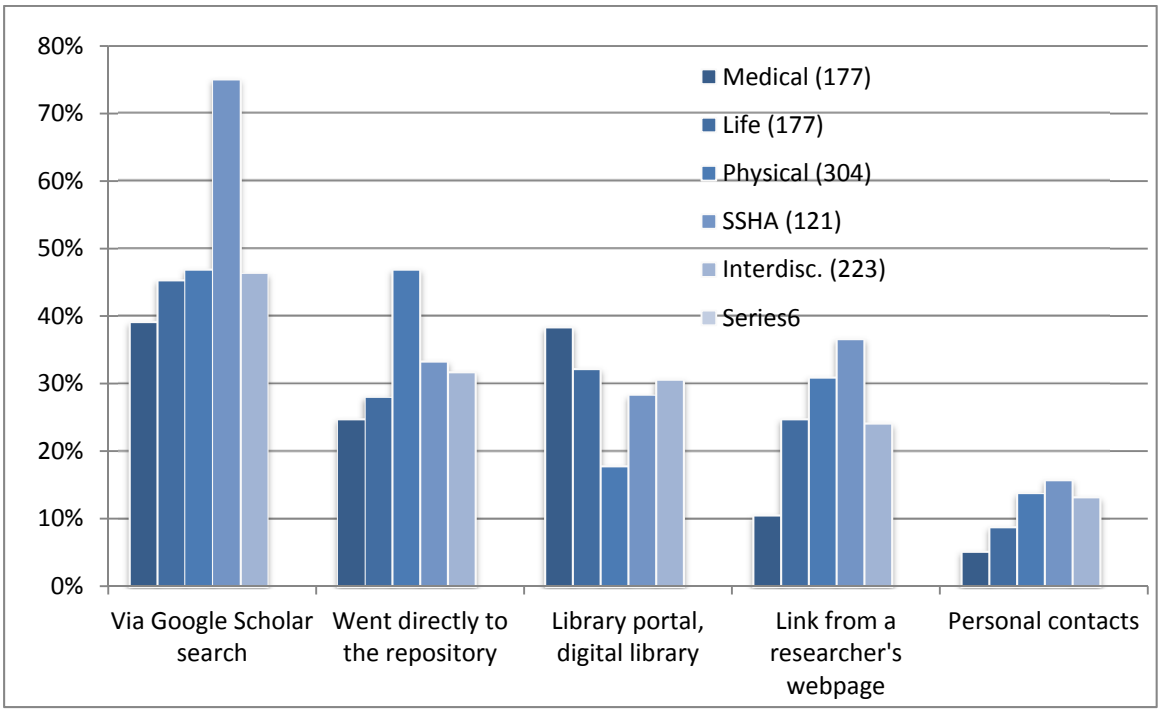
Figure 10 Methods of accessing repositories



Total respondents = 1,012

Researchers in the Social sciences, humanities & arts were more likely to have accessed a repository via a Google Scholar search than any other broad discipline group; whilst respondents in the Physical sciences & mathematics were more likely to have gone directly to a repository than respondents in any other broad discipline group (Figure 11). This latter finding is consistent with the self-reported behaviour of physicists in relation to arXiv (Creaser et al., 2010). Given the importance of PubMed Central to researchers in the Medical sciences (Fry et al., 2009), it might be expected that Medical scientists would report starting their search with this information resource, but our findings show a different pattern of search behaviour. Medical scientists indicated that they were more likely to use their home institution's library portal to access OARs than researchers in the other broad-discipline groups. Respondents from the Social sciences, humanities & arts were more likely to have accessed the repository through a link on a researcher's webpage or through personal contacts, than researchers in the other broad discipline groups.

Figure 11 Methods of accessing repositories, by subject



(Weighted total numbers of responses given in brackets)

Satisfaction with the versions of journal articles held in open access repositories

Three different versions of the same journal article may be made available in an OAR: the pre-print or submitted version (stage-one manuscript), the authors' final peer-reviewed and accepted version (stage-two manuscript) and the final published version (stage-three manuscript). The perceived value of the article version may vary according to the subject areas of researchers.; the Physics and Economics communities are well-known examples of disciplines placing a high value on the dissemination and reading of pre-prints, whereas Fry et al. (2009) found anecdotal evidence that researchers in the Medical sciences had trust issues with the use of pre-prints, and to some extent with any version other than the final published version, not the least because of the fundamental requirement within their own research culture that they build their research upon validated results. In this case validation is achieved through formal peer review and publication in a reputable journal.

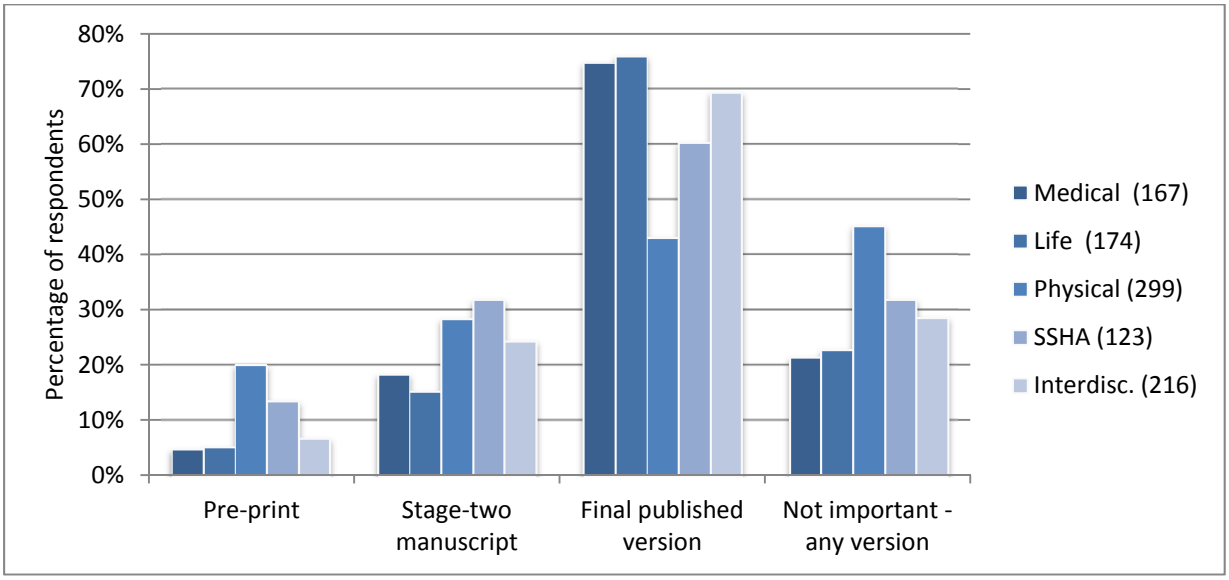
In this context, given the variety of versions of journal articles held in repositories, and given the place of the journal article in researchers' information landscapes, it was deemed crucial to understand researchers' expectations and attitudes towards what version is available to them in OARs. We therefore asked for details about the version researchers expected to find on their most recent visit to an OAR, the version they actually accessed and whether they felt the version they accessed was satisfactory for the task at hand.

The research indicated that respondents were generally more likely to be seeking the final published version of the journal article. Sixty-two per cent of researchers reported looking for the final published version on their most recent visit to an OAR. Twenty-four per cent of researchers reported that they were expecting to find the stage-two manuscript. It is of interest to note that almost one-third of researchers reported that it was not important which version they were able to access[4]. Some disciplinary differences were observed in the responses, as illustrated in Figure 12. Unsurprisingly, researchers from the Physical sciences & mathematics were more likely to be looking for a pre-print than those from the Medical sciences or the Life sciences and Interdisciplinary researchers; this may reflect their long-standing pre-print culture. Researchers from the Physical sciences & mathematics, and the Social sciences, humanities & arts were more likely to be looking for the stage-two manuscript than respondents from the Medical

[4] Note that respondents could select one or more of the following categories: pre-print/authors' final version/final published version/ the version is not important.

sciences or the Life sciences. In contrast to this, researchers from the Medical sciences or the Life sciences were more likely to be looking for the final published version than respondents from the Physical sciences & mathematics, and the Social sciences, humanities & arts. Respondents from the Physical sciences & mathematics were also most likely to indicate that it was not important which version they found when visiting a repository.

Figure 12 Version of the article hoping to find, by subject



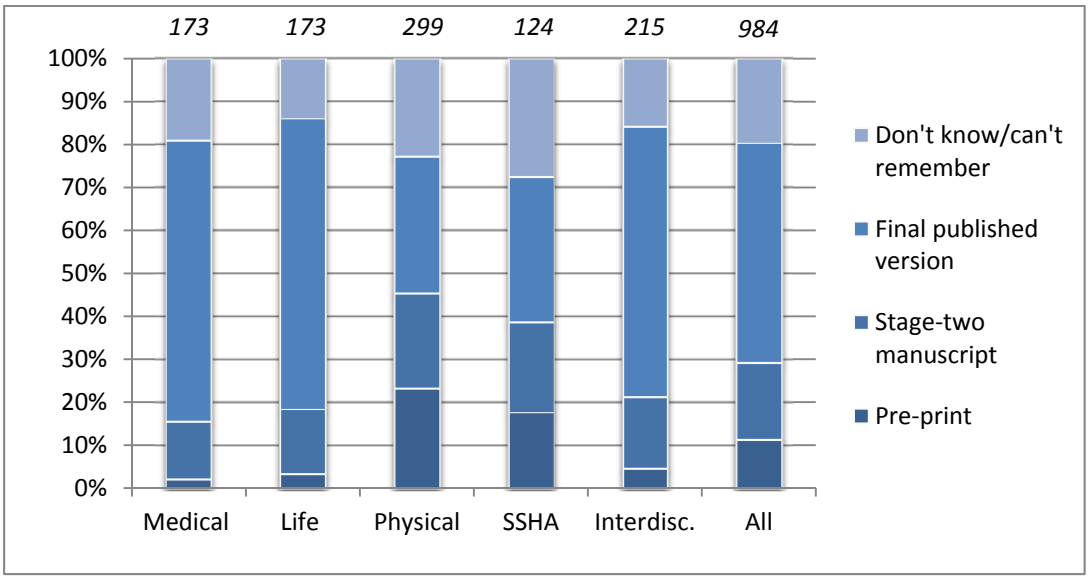
(Weighted total numbers of responses given in brackets)

Further disciplinary differences were noted concerning the article version sought by researchers during their most recent visit to a repository. Researchers from the Medical sciences and the Social sciences, humanities & arts were most likely to consider the article version important and least likely to 'trust' versions of articles held in a repository unless it was very clear to them that they had accessed the final published version. On the other hand, there was a greater acceptance of pre-prints by researchers from the Physical sciences & mathematics than by those from the other broad disciplinary groups. Uncertainty about article versions located in repositories seemed to be mitigated to some extent by author or journal reputation.

Half of all researchers reported finding the final published version on their most recent visit to a repository, although this varied by discipline (Figure 13). Researchers in the Medical sciences and Life sciences were more likely to report they had accessed the final published version; they were also more likely to have been looking for this version in the first place. It is unclear whether such a high rate of published versions reportedly found in repositories is representative of the version content held in OARs worldwide, although the disciplinary differences shown in Figure 13 are in line with known disciplinary OA practices and the use of PubMed central by researchers in the Medical and Life sciences. This may also, to some extent, raise the question of whether researchers conflate OARs with open access journals, or whether the clarity of the metadata for the version accessed in repositories is not sufficient to enable readers to distinguish between an author-produced PDF of a stage-two manuscript and the publisher-produced PDF of the final published version of the article.

As for the stage-two manuscript, 18% of survey respondents reported they had found this version; and a further 12% reported having found a pre-print version. Respondents from the Physical sciences & mathematics and the Social sciences, humanities & arts were the most likely to have found a pre-print; they were also most likely to have been looking for a pre-print.

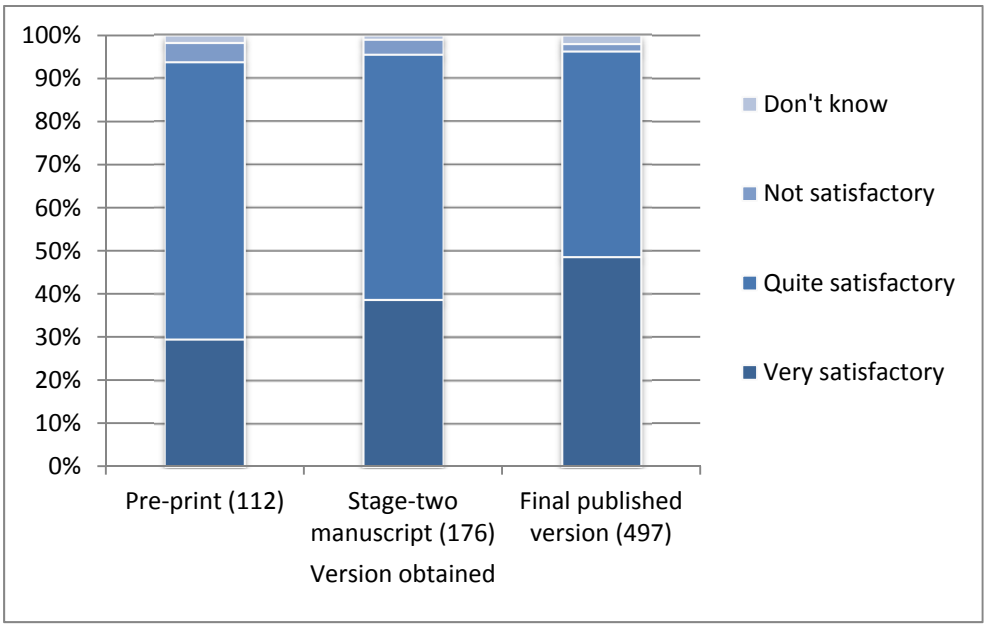
Figure 13 Version found, by subject



(Weighted total numbers of responses given in italics)

In terms of the extent to which the version that researchers found in the OAR met their requirements, just over half of the researchers surveyed indicated that the article they accessed was 'quite' satisfactory for the task they were undertaking, with a further 40% indicating that it was 'very' satisfactory. There is also a clear association between the version found and the level of satisfaction with it (Figure 14).

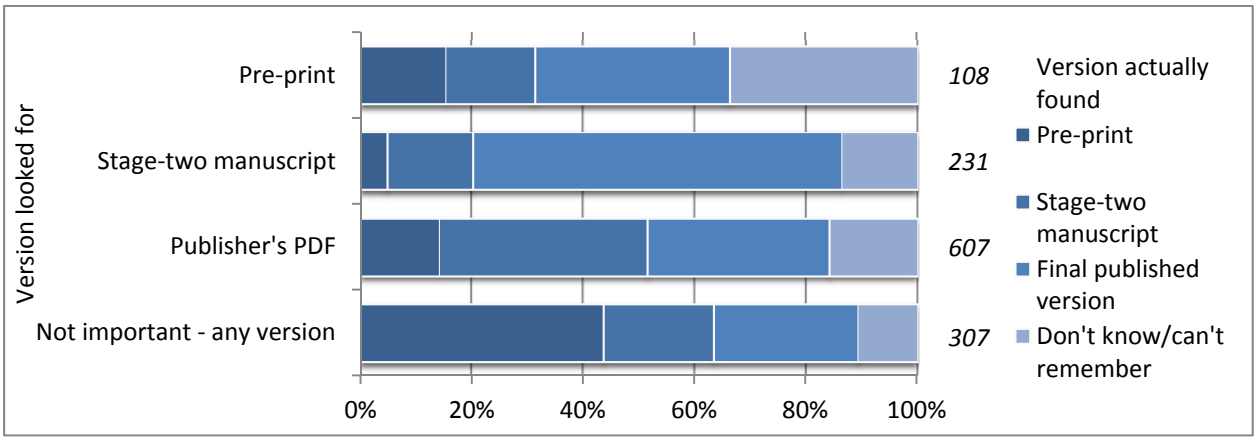
Figure 14 Satisfaction with the version obtained



(Weighted total numbers of responses given in brackets)

When comparing the version respondents were seeking with the version they actually found (Figure 15), statistically significant differences were found. Perhaps not surprisingly, in the case of pre-prints and the final published version, those respondents that had indicated they were looking for these versions, were more likely to have found them rather than other versions. Respondents that were looking for the author's final version were more likely to have either found this version or the final published version, whilst those respondents that had stated it was not important and were looking for any version were more likely to have found the final published version or be unsure which version they had accessed.

Figure 15 Version of article hoped for, by version found



(Weighted total numbers of responses given in italics)

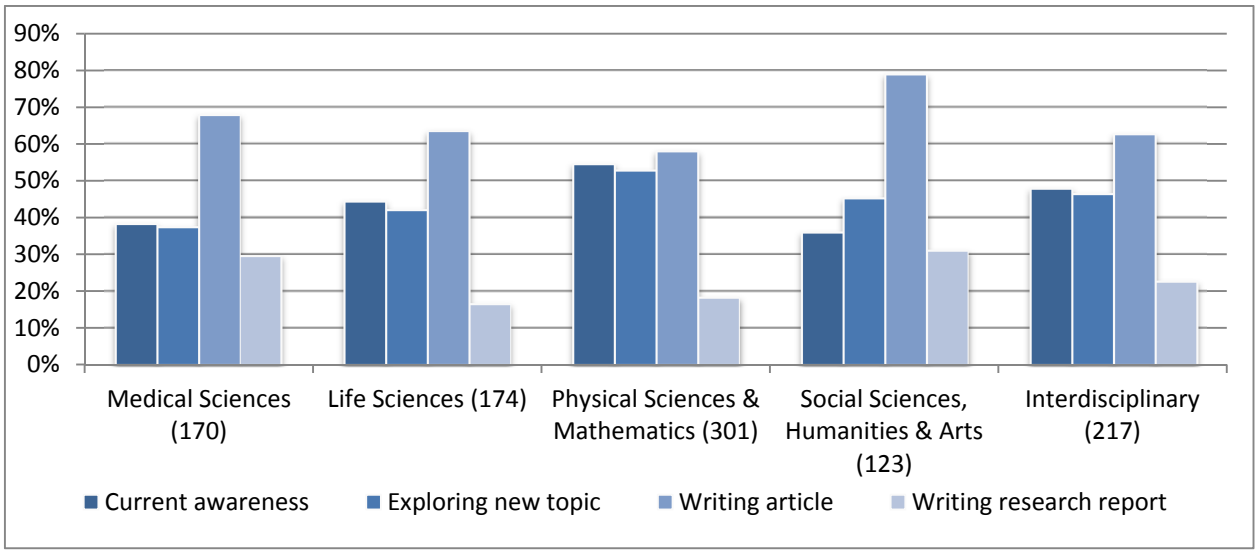
The survey findings suggested that evaluating the suitability of an article for the task they were undertaking did not seem to be problematic, with 64% of researchers finding the article they accessed either 'easy' or 'very easy' to evaluate. Workshop participants, however, indicated that they often found it difficult to ascertain which version of the paper they had accessed from OAR platforms. They indicated that this was not always clear either from the electronic file itself or from the accompanying metadata. This is in line with phase 1 survey results, indicating that 16% of respondents could 'rarely' or 'never' identify the article version from the repository information provided, while 31% could do so only 'sometimes', and 20% did not know (Fry et al., 2009).

Satisfaction with repository material in relation to the purpose of reading

The main mediating factor between purpose of reading and repository behaviour appears to be the article version being sought and clarity of the article version found. Our research found that, in general, researchers tend to perceive versions of articles found in OARs differently depending on the task at hand. In some cases, researchers may emphasise the need to work with validated results, i.e. results that have been through peer review and are published in a prestigious journal, and hence they may be looking for published versions of journal articles. Such behaviour was typically displayed in our study by researchers in the Medical sciences. In addition, the ease of reading that typically characterises final published versions (noted in the participatory workshop, especially in relation to tables and graphs) was often brought forward as an explanation for favouring reading the published version. Beyond these two elements of ease of reading and trust, the value researchers place on different versions of the same journal article will also depend on the purpose of their reading.

Overall, almost two-thirds of respondents indicated that on their most recent visit to a repository they were writing an article, whilst almost half of respondents stated that their purpose was current awareness, or exploring a new topic. Some disciplinary differences were noted, with respondents from the Social sciences, humanities & arts most likely to have arrived at the repository in the process of writing an article. Respondents from the Physical sciences & mathematics were more likely to have come to the repository for the purpose of current awareness than respondents from the Medical sciences or the Social sciences, humanities & arts, and were more likely to be exploring a new topic than respondents from the Medical sciences. In contrast, respondents from the Physical sciences & mathematics and the Life sciences are less likely to have arrived at a repository whilst writing a research report than respondents from the Medical sciences or the Social sciences, humanities & the arts (Figure).

Figure 16 Purpose of visiting repositories, by subject



(Weighted total numbers of responses given in brackets)

Analysis of the qualitative data (phase 1 focus groups and phase 2 participatory workshop) illustrates that researchers' different roles, reflected in the purpose of reading, may influence their acceptance of versions other than the final published version of a journal article. Indeed, our findings showed that, if researchers seek and read journal articles to inform their research with a view to writing a peer-reviewed journal article themselves, then they are most likely to try to locate the published version of the journal article that they have identified. Some workshop participants expressed their preference to have the final published version of their own journal article(s) cited, not only because it is the most authoritative version of their paper but also because it enables them to figure in citation counts. It was also suggested that citing a version other than the final published version would depend on the norms and research culture within the discipline and where this was not common practice within a discipline, there was a perception that it could have a negative effect on readers' perceptions of the value of the article, in terms of the quality and authority of the sources. Finally, citation practices also depend on a journal editor's specific requirements with regard to how a source should be cited, which may in some instances preclude the citation of versions other than the final published version of a journal article This may in turn influence readers' use of material in repositories if they perceive that they will not be able to cite such material in journal articles that they are writing.

The participatory workshop findings suggested that accessing the final published version of an article appears to be less critical for writing research outputs other than peer-reviewed journal articles, however. This was also the case for current awareness activities, particularly in novel areas of research where there may be only a limited number of journal articles that have been published on the topic of interest. Indeed, there is evidence that for readers who are not authors themselves, pre-prints or stage-two manuscripts could be satisfactory, in that their reading involves learning about a topic and using this knowledge without the need to formally cite material in scholarly works. However, it is important to reiterate the importance that workshop participants placed on accessing, reading and citing the final published version, which remains the version of choice for the purpose of writing a journal article. Participants across disciplines indicated on many occasions that they highly value final published versions as they were thought to be easier to read than pre-prints or stage-two manuscripts, especially with regard to graphs, tables and pictures.

Discussion

Self-archiving behaviours

According to ROAR (Registry of Open Access Repositories) the number of institutional or departmental research repositories worldwide has grown exponentially year on year and in recent years a number of subject-based repositories have come to the fore as *de facto* centralised resources in some disciplines. These developments have been coupled with initiatives and policies to mandate, or at least encourage, researchers to deposit copies of their published journal articles into OARs. These collective developments represent the so called 'Green Road' to OA and it might logically be anticipated that an increased level of technical developments and policy activity towards this route to OA would be accompanied by a corresponding increase in levels of awareness amongst researchers. Related research has reported that OARs are not currently on researchers' radar and are not highly visible within the majority of research communities (Harnad et al., 2009). Our findings show, however, that over half of all authors surveyed have placed, or given permission to have placed on their behalf, a version of their journal article(s) in an OAR in the past five years.

Subject-based and institutional repositories have emerged in different contexts and with a different set of goals. In some disciplines there are de facto centralised repositories, such as arXiv in physics, RePEc in economics and PubMed Central in the biomedical and life sciences, while in other disciplines the repository landscape is less well-established. In terms of authors' preference for type of repository in which to place a version of their article(s), there was a slight shift between phase 1 and phase 2 towards institutional repositories, which may reflect the current development of the European repository infrastructure, with a growing number of research institutions setting up their own open access digital repository, as reflected by ROAR data. When looked at from a disciplinary perspective some differences were notable, with researchers in the Physical sciences & mathematics expressing a marked preference for placing articles in subject-based repositories, whereas researchers in the other broad discipline groups expressed a preference for institutional repositories.

The preference for subject-based repositories by researchers within the Physical sciences & mathematics was, to a large extent, expected, as it reflects the high adoption rate of the subject-based repository arXiv by these disciplinary communities. Given the high profile of PubMed Central within the Medical science research community the preference amongst Medical science respondents for placing their articles in an institutional repository is somewhat surprising. This finding might, in part, be explained by the fact that Medical scientists reported that in general somebody else placed a copy of their article(s) in an OAR on their behalf and there may be some misunderstanding about the type of OAR into which material is being deposited.

Whilst researchers in the Medical and Life sciences reported that their journal articles were generally deposited by someone else on their behalf, researchers in the Physical sciences & mathematics reported depositing their journal articles into OARs themselves, which is consistent with arXiv's deposit procedures where there is no mediated deposit service. Within the Social sciences, humanities & arts no clear pattern of deposit behaviour emerged, which may be due to the extremely broad scope of this disciplinary grouping, for which internal and external initiators seem to influence repository type preferences: when the article is deposited into an institutional repository then somebody else is more likely to have deposited on a researcher's behalf, whereas when the deposit is made into a subject-based repository, then researchers are more likely to do it themselves.

Furthermore, aside from voluntary deposits, Medical science and Life science researchers were more likely to deposit journal articles into OARs following an invitation from a publisher or a repository than the other disciplines. Researchers in the Physical sciences & mathematics and in the Social sciences, humanities & arts, on the other hand, were largely more likely to report having deposited their journal articles into OARs voluntarily; and, to a lesser extent, they were also likely to deposit their journal articles in response to a co-author request (particularly physicists & mathematicians) or a repository prompt (particularly researchers in the Social sciences, humanities & arts).

In terms of the article version deposited, Medical and Life scientists reported depositing the final published version, whereas researchers in the Physical sciences & mathematics and Social sciences, humanities & arts reported depositing either pre-prints or the stage-two manuscript. The deposit of final published versions in the Medical and Life sciences is likely to be linked to the primary deposit process for PubMed Central, whereby those journals that have an agreement with PubMed Central deposit the published version of articles on behalf of authors, after the embargo period (if any) has expired. The deposit of pre-prints or state-two manuscripts by researchers in Physical sciences & mathematics and Social sciences, humanities & arts is in line with the recognised pattern of self-archiving in subject-based or institutional repositories.

Readers' open access behaviours

The research found interesting disciplinary patterns in the way researchers access and use material held in OARs. Researchers in all disciplines, except the Physical sciences & mathematics, generally arrive indirectly at an OAR, often through Google or Google Scholar. Researchers in the Life sciences and in the Social sciences, humanities & arts indicated that they would often arrive at an OAR through a Google Scholar search, whilst medical scientists seemed to favour Google over Google Scholar when they search for information. Researchers in the Social sciences, humanities & arts were also likely to arrive at an OAR following a link from a researcher's webpage. Researchers in the Physical sciences & mathematics, on the other hand, were the only ones reporting going directly and purposively to OARs to access journal articles. This distinct behaviour of Physical scientists & mathematicians is in line with their use of arXiv - checking new papers in arXiv is part of their daily research routine (Creaser et al, 2010).

Researchers in all disciplines indicated that writing an article was generally the main purpose of their reading when they had accessed versions of journal articles held in OARs, resonating with recent studies showing the importance of journal literature in researchers' workflows (Tenopir et al., 2011). Our research found a marked difference in the purpose of reading between Physical scientists & mathematicians and the other broad disciplinary groups, however. Physical scientists & mathematicians reported that the purpose of reading varies equally between article writing, current awareness, and exploring a new topic when they access versions of journal articles held in OARs, whereas researchers in the other three disciplines generally reported being in the process of writing an article.

As for the version sought compared to that accessed, all disciplines showed coherent behaviours in terms of what they look for and what they actually access. Medical and life scientists usually look for, and find, the final published version, whilst researchers in the Social sciences, humanities & arts tend to look for the stage-two manuscript, and report accessing either a pre-print or the stage-two manuscript. Physical scientists & mathematicians tend to look for pre-prints or stage-two manuscripts and access one of those two pre-publication versions.

Conclusions

Approximately half of the authors reported self-archiving a version of their journal article(s) in an OAR, although a majority of these indicated that someone else had made their work available in an OAR for them. There were differences across the broad discipline groups in terms of the self-archiving process, with authors from the Physical sciences & mathematics and the Life sciences more likely to place a version of their article(s) in an OAR themselves, whilst authors in the Medical sciences or the Social sciences, humanities & arts more likely to have a version of their article(s) placed in an OAR by somebody else.

The type of OARs in which authors are likely to self-archive a version of their article(s) seem to be influenced by the nature of their motivation. Indeed, if the motivation is voluntary, requested by a co-author, or an invitation by a publisher then authors are more likely to choose a subject-based repository whereas if authors are invited by repository/library staff, or were mandated by their employer, they are more likely to choose an institutional repository. It is of interest to note that little importance was, however, given to funder and institutional mandates as motivators for repository deposit.

The final published version is the version most likely to be placed in an OAR, overall. Authors from the Physical sciences & mathematics and the Social sciences, humanities & arts were more likely to self-archive a pre-print or stage-two manuscript. Authors from the Medical sciences were most likely to be uncertain about which version of their article had been placed in an OAR, which correlates with the high proportion of Medical science researchers who reported that their article(s) had been deposited by someone else.

When accessing journal articles in OARs, the article version was most likely to be considered important by researchers from the Medical sciences and the Social sciences, humanities & arts. Researchers in those broad disciplinary groups were also least likely to 'trust' versions of articles held in a repository, unless it was very clear to them that they had accessed the final published version. There was a greater acceptance of pre-prints by researchers from the Physical sciences & mathematics than researchers from the other broad-disciplinary groups.

The clarity of repository metadata and readers' ability to distinguish between the different versions of an article seem critical to how the 'quality' (e.g. whether it is authoritative) of repository content is perceived. Researchers reported that whilst the difference between a pre-print and a final published version was reasonably clear, it was much more difficult to distinguish between a submitted pre-print and an accepted stage-two article.

Disciplinary behaviours tended to vary depending on whether researchers were reporting behaviours from the perspective of an author or a reader. Researchers in the Medical and Life sciences often signalled an inconsistency between their repository behaviours as an author or as a reader, whereas researchers in the Physical sciences & mathematics demonstrated a stronger alignment between their attitudes and behaviours as authors or readers.

This research has taken a broad-brush approach to investigating researchers' behaviours with respect to the use of OARs, both as authors for dissemination and as readers seeking information. It has highlighted differences and similarities, many related to the homogeneity of the broad discipline groups used in the analysis. Further analysis at a more granular disciplinary breakdown of the survey data is expected to add to our understanding of repository behaviours, particularly in sub-disciplines where there is already general acceptance of OA. Such understanding may help to inform and/or support policy makers and research funders in their decision to actively promote, or not, OA to research information. It would also provide further valuable information about the fine-grained characteristics at sub-discipline level underlining the (non) adoption of OA by specific research communities.

References

Björk, B., Welling, P., Laakso, M., Majlender, P., Hedlund, T. and Guõnason, G. (2010), "Open access to the scientific journal literature: situation 2009", PLOS One, 5(6).

Blake, C. and Pratt, W. (2002), "Collaborative information synthesis", in Toms, E.G. (Ed.), ASIST 2002: Proceedings of the 65th ASIST Annual Meeting., Medford, NJ: Information Today, pp. 44-56.

Case, Donald O. (2012), *Looking for information: a survey of research on information seeking, needs, and behaviour* (3rd ed.), Emerald, Bingley.

Creaser, C., Fry, J., Greenwood, H., Oppenheim, C., Probets, C., Spezi, V. and White, S. (2010), "Authors' awareness and attitudes toward open access repositories", New Review of Academic Librarianship, 16(S1), pp. 145-161.

Dallmeier-Tiessen, S., Darby, R., Goerner, B., Hyppoelae, J., Igo-Kemenes, P., Kahn D., Lambert, S., Lengenfelder, A., Leonard, C., Mele, S., Polydoratou, P., Ross, D., Ruiz-Perez, S., Schimmer, R., Swaisland, M. and van der Stelt, W. (2010), "First results of the SOAP project. Open access publishing in 2010", available at: http://arxiv.org/ftp/arxiv/papers/1010/1010.0506.pdf (accessed 12 January 2012).

Department for Business, Innovation and skills (2011), "Innovation and research strategy for growth" [Report presented to Parliament by the Secretary of State for Business, Innovation and skills], available at: http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.pdf (accessed 12 January 2012).

European Commission. (no date), "Research & Innovation - Policy initiatives" [website], available at: http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1294&lang=1 (accessed 12 January 2012).

Working Group on Expanding Access to Published Research Findings (2012) Accessibility, sustainability, excellence: how to expand access to research publications, available at http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf (accessed 13 July 2012)

Fry, J., Oppenheim, C., Probets, S., Creaser, C., Greenwood, H., Spezi, V and White, S. (2009), "PEER Behavioural Research Baseline report", available at: http://www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf (accessed 12 January 2012).

Fry, J., Probets, S., Creaser, C., Greenwood, H., Spezi, V and White, S. (2011), "PEER Behavioural Research: Authors and Users vis-à-vis Journals and Repositories (Final report)", available at: http://www.peerproject.eu/fileadmin/media/reports/PEER_D4_final_report_29SEPT11.pdf (accessed 12 January 2012).

Goodman, D., Dowson, S. and Yaremchuk, J. (2007), "Open access and accuracy: author-archived manuscripts vs. published articles", Learned publishing, 20(3), pp. 203-215.

Harley, D., Accord, S. K., Earl-Novell, S., Shannon, L. and King, C. J. (2010), "Assessing the future landscape of scholarly communication: an exploration of faculty values and needs in seven disciplines", UC Berkeley: Center for Studies in Higher Education, available at http://cshe.berkeley.edu/research/scholarlycommunication/ (accessed 12 January 2012).

Harvard University Library. (2010), "Office for scholarly communication - Open Access policies" [website], available at: http://osc.hul.harvard.edu/policies (accessed 12 January 2012).

Harnard, S., Carr, L., Swan, A., Sale, A. and Bosc, H. (2009), "Maximising and measuring research impact through university and research-funder open-access self-archiving mandates", Wissenschaftsmanagement, 15(4), pp. 36-41.

Hjørland, B. and Albrechtsen, H. (1995), "Toward a new horizon in information science: Domain-analysis", Journal of the American Society for Information Science, 46, pp. 400-425.

Hjørland, B. (2002), "Epistemology and the socio-cognitive perspective in information science", Journal of the American Society for Information Science and Technology, 53, pp. 257-270.

King, D., Tenopir, C., Choemprayong, S. and Wu, L. (2009), "Scholarly journal information-seeking and reading patterns of faculty at five US universities", Learned Publishing, 22(2), pp. 126-144.

Organ, M. (2006), "Download statistics – What do they tell us?", D-Lib Magazine, 12(11), available at: www.dlib.org/dlib/november06/organ/11organ.html (accessed 12 January 2012).

Research Information Network (2009) Communicating knowledge: How and why UK researchers publish and disseminate their findings, available at http://www.rin.ac.uk/system/files/attachments/Communicating-knowledge-report.pdf (accessed 13 July 2012)

Rowlands, I., Nicholas, D. and Huntingdon, P. (2004), "Scholarly communication in the digital environment: what do authors want?", London: CIBER, available at: http://www.homepages.ucl.ac.uk/~uczciro/ciber-pa-report.pdf (accessed 12 January 2012).

Shepherd, P.T. and Wallace, J.M. (2009), "PEER: a European project to monitor the effects of widespread open access archiving of journal articles. Based on a presentation given at the UKSG seminar 'Mandating and the scholarly journal article: attracting interest on deposits'?", London, 29 October 2008. Serials, 22(1), pp. 19-23.

Swan, A. and Brown, S. (2004), "Authors and open access publishing", Learned Publishing, 17(3), pp. 219-224.

Swan, A. and Brown, S. (2005), "Open access self-archiving: an author study", Truro, UK: Key Perspectives Limited, available at: www.jisc.ac.uk/media/documents/themes/infoenvironment/oaselfarchivingauthorstudy.pdf (accessed 25 January 2012).

Talja, S., Vakkari, P., Fry, J., and Wouters, P. (2007) "Impact of Research Cultures on the Use of Digital Library Resources". Journal of the American Society for Information Science and Technology, 58(11), pp.1674-1685.

Tenopir, C., King, D., Edwards, S. and Wu, L. (2009), "Electronic journals and changes in scholarly article seeking and reading patterns", Aslib Proceedings New Information, 61(1), pp. 5-32.

Tenopir, C., Mays, R. and Wu, L. (2011), "Journal article growth and reading patterns", New Review of Information Networking, 16(1), pp. 4-22.

Tenopir, C., King, D.W. and Bush, A. (2004), "Medical Faculty's Use of Print and Electronic Journals: Changes Over Time and Comparison with Other Scientists", Journal of the Medical Library Association, 92(2), pp. 233–241.

Thatcher, S. (2011), "Copyediting's role in an open-access world", Against the grain, 23(2), available at: http://www.against-the-grain.com/ (accessed 25 January 2012).

Wates, E. and Campbell, R. (2007), "Author's version vs. publisher's version: an analysis of the copy-editing function", Learned publishing, 20(2), pp. 121-129.

Whitley (2000), The intellectual and social organization of the sciences (2nd ed.), Clarendon Press, Oxford.