INTERDISCIPLINARY DIFFERENCES IN ATTITUDES TOWARDS DEPOSIT IN INSTITUTIONAL REPOSITORIES

James Allen

CONTENTS

FIGURES and TABLES

ABSTRACT

The attitudes and behaviours of academics from different disciplines towards depositing their work in institutional repositories are compared. This is achieved through the use of a survey strategy, and by examination of the contents of a twenty-five UK institutional repositories. The survey targets humanities academics, and the data is compared to that from previous surveys focusing on scientific, technical and medical (STM) disciplines.

The number of humanities documents in institutional repositories is currently far lower than that in STM disciplines. Awareness of Open Access amongst humanities academics is also low. However they perceive many advantages to depositing their work in institutional repositories, especially for the reader, not for themselves. Around two-thirds of respondents would deposit work in institutional repositories, despite having several concerns. Those who would not deposit work in this way perceive the same disadvantages: potential for plagiarism, the apprehension of interfering with publishing their work elsewhere, and the fragility of online means of dissemination.

Increased depositing in institutional repositories in the future depends on encouraging authors of the advantages of doing so, not only to others but also to themselves. At this early stage of development understanding the attitudes of academics in different disciplines is crucial

INTRODUCTION

Open Access had become an increasingly strong movement in recent years. The aim is to make research literature, especially peer-reviewed academic articles, free for anyone in the world to access. Dissatisfaction with the traditional publishing models, which have remained the same for hundreds of years, has encouraged experimentation with different models. Three important factors have come together to enable this:

- increasing subscription costs of journals, which have forced libraries to cut titles, especially where budgets are tight,

- objection to subscriptions restricting access to publicly funded research, forcing researchers to buy back their own work,

- and the technology of the Internet, which allows information to be disseminated widely and cheaply.

On their own these may not precipitate change. But strong recent support from major research funders, such as the Wellcome Trust in the UK and the National Institutes for Health in the USA, as well as from governments including the UK, and Universities such as MIT and Southampton, has forced the issue onto the agendas of academics, researchers, librarians and publishers alike.

Academic articles can be made 'open' via two main routes, but there are also hybrid publishing models in between. The first is via an Open Access journal. Rather than gaining income through selling copies of the finished journal, Open Access journals usually use an 'author pays' model, whereby a fee is paid when a paper is accepted to be published: by retrieving the costs in advance all subsequent access via the Internet is free.

The second method to achieve Open Access is by depositing work in specially designed repositories or archives. These are either developed to capture work from a certain discipline, or based within an institution or university. Once deposited, the work can then be accessed by anyone for free via the Internet, and located through special search interfaces which search across a number of separate repositories. This latter model is often seen to be complimentary to the traditional journal system: a paper can still be published in a traditional journal, bringing with it advantages including peer recognition within the field, but those who do not or cannot pay for a subscription to the journal itself can still access the article. Of course, in the traditional system, publishes own the copyright of articles, so their cooperation in allowing their information to be made available in this way is essential.

It is this second form of Open Access which the UK House of Commons Select Committee was particular enthusiastic about, and which universities and research funders in the UK, and elsewhere, are promoting most heavily. For such a network of archives, especially institutionally-based repositories, to become an accepted part of the dissemination process, it is crucial that a significant proportion of work is deposited. It is notable that most of the successful Open Access projects, and most of the literature and research on the subject, in recent years have been in the scientific, technical and medical (STM) fields. How will institutional repositories fare now that they are targeting the entire range of academic fields?

The aim of this study is to consider the first stage of this question: what are the attitudes of academics outside STM fields towards the concept of depositing their work in institutional repositories? The importance of this is that it will act as baseline against which development can take place. By identifying particular concerns with the model, and by establishing what level of support there is for such changes, then those developing institutional repositories will be in a better position to tackle them.

Specifically, this research will achieve four objectives. The first necessity is to discover how much academics outside STM fields know about the Open Access movement. What terms do they understand, what services have they used, and how did they hear about them? A questionnaire aimed at a sample of academics will establish this. The second objective is, via the questionnaire and follow-up interviews, to compare the attitudes towards institutional repositories of academics from the two broad disciplines. The survey will focus on academics from humanities fields, and the data will then be compared to results from previous surveys of STM academics. The follow-up interviews will also achieve the third stage, which is to establish why these attitudes are held: what elements of the institutional repository system appeal to authors, and what discourages them from depositing work? The interviews are the best way to uncover the complex reasons why authors publish in particular ways and places.

The fourth element of the research will focus on the actual behaviour of academics from different disciplines in depositing their work. This will be through a survey of the current contents of institutional repositories in the UK. In the past few years, many of the research-focused universities in the UK have developed institutional repositories, but how are they being used? What trends are visible, and are there any patterns along disciplinary lines?

By comparing the attitudes of the two disciplines, and by considering their actual behaviour in using the institutional repositories, this project will show how and where development can

take place. Of course these results to be very specific to a single point in time – in the near future the JISC-funded SHERPA project will produce its evaluation, and the UK research councils will begin mandating the deposit of work in repositories within a few months (RCUK, 2005). But that is why the research is so important. At this early stage in the development of what promises to be the most widespread Open Access project yet, the attitudes, behaviours and understanding of the issues involved are central to the success of failure of the project.

This report will be structured as follows. Firstly, the background to the study will be better explained, and a comprehensive literature review undertaken. This will be followed by a discussion of the methodology employed in the different stages of the study. Then the findings of the survey of the contents of UK institutional repositories will be described, followed by the results of the questionnaire. Finally all the quantitative and qualitative data will be brought together in a discussion. General conclusions, and suggestions for further research, will be given at the end.

BACKGROUND AND LITERATURE REVIEW

Much has been written about Open Access in all its forms in recent years. The majority of the debate has been the personal opinions of people involved at some stage in the academic publishing cycle, whether authors, librarians or publishers themselves. Some have been descriptive articles on particular projects or developments, but little has been actual research. Research has tended to focus on the attitudes and behaviors of both authors and readers of academic research material, considering to what extent they would find OA useful, and what problems and concerns they would have in making their work available in this manner. These have mainly taken the form of surveys, either of particular institutions or broadly aimed, but it is difficult for this kind of research to ever be fully reliable, valid or generalisable. They are, however, the most appropriate means for the research, and will also be employed in this study, but it is worth remembering that they are not perfect.

To describe and explain the context in which institutional repositories sit it is necessary to:
- explain the Open Access movement, including the reasons for its emergence,
- describe the methods by which Open Access can be approached, including their advantages and disadvantages, and support for and objections to them,
- consider the impact that Open Access and institutional repositories will have on scholarly publishing, on publishers, on libraries and on users.

Finally the actual use of Open Access means to deposit and find information will be tackled.

Open Access (OA) is not a straight-forward concept: it has many different definitions and approaches, numerous supporters and abundant detractors. The concept has existed since the mid-1990s, with champions such as Stevan Harnad (1994) expressing his 'subversive proposal.' In a brief statement, Harnad claimed that authors of 'esoteric' articles, for which no payment is received, only want to publish their work not sell it. He recommends a two step process to make access to such material 'open': first to place pre-prints, that is papers that have not yet been peer-reviewed or quality controlled, into globally accessible archives; then, once they have been peer-reviewed and made ready for publication, substitute this version into the archive. Alongside these archives, journals could still exist, but they would no longer be able to charge subscription fees, but would be forced to gather the costs before publication through 'authors' page charges, learned society dues, university publication budgets and/or governmental publication subsidies.' This original proposal sparked a heated debate, but the basic tenets of OA were set.

More recently, OA has found much wider support. The Budapest Open Access Initiative (Chan et al., 2002), signed by Harnad and several others, elaborated on the 'unprecedented public good' that OA could do. The BOAI stresses the importance of Internet technology in making it possible for scholarly literature to be distributed on a global scale, and made available free of charge. The stated advantages of this are to:

> 'accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge.'

The response to this call to action has been positive: over 4000 people have added their signature to the declaration, and innumerable universities, libraries, organisations and even governments have declared their support for the initiative. Similar calls for support of the OA model have also been successful, with the Public Library of Science (PLoS) raising over 30,000 signatures on an open letter threatening to boycott journals not granting open access within six months of initial publication (Ramachandran, 2004).

OA appears to have popular support, but why has it emerged in recent years? The technology may permit this model of disseminating information, but just having the technology does not mean change will take place. The technology not only makes worldwide information dissemination possible, but it also offers many advantages over the print system: distribution costs and time are reduced, there is a much larger potential audience, communication can become discursive more easily, and 'the danger that anything not on the web will be neglected' is precluded (Odlyzko, 2002, p.8). Even these advantages, however, do not address the basic objections to the traditional system felt by the 30,000 PLoS supporters: this widespread support suggests dissatisfaction with the current, traditional model of scholarly publishing. Why?

Academic librarians have, for a number of years, been vocal on the topic of a 'serials crisis,' which is the process by which the cost of journal subscriptions has taken up an increasing share of library budgets. This incongruity has developed from both increasing subscription prices – Lawal (2002) notes that libraries were paying three times more 'in real dollars' for 7% fewer journal titles in 2001 than in 1986 – and stagnating library budgets (Crawford and Gorman, 1995). The result of this has been the cancellation of journal subscriptions, and thus decreasing levels of access to research information for those who need it. This situation explains why many librarians have misgivings about the traditional journal system and publishers. There is also an issue of 'double payment' involved: much of the contents of journals to which universities subscribe is publicly funded, so should it not be available free to the public? Research funders, such as the Wellcome Trust, are coming to accept and

encourage the view that research they fund should be available for free. But publishers argue that anyone, including the public, can already access research results, via their local libraries or document supply services, although not necessarily for free. And as Reich (2003, p.137) points out, taxes are also spent on farm subsidies but 'I don't see anyone handing me free loaves of Wonder Bread.'

Others take a more fundamental view, criticising the system for imposing barriers to access, whatever the price. Subscriptions, licenses and pay-per-view make up the trio of barriers denounced by Harnad (2001), who points out that 'most refereed articles are inaccessible to most authors.' Likewise, the Wellcome Trust (2003) claim that the monopoly held by publishers in the current system does not act in the interests of either the academic community or the public. Focusing on access to journal articles also highlights the geographical disparities that exist between resource rich and resource poor countries, as well as between institutions within the same country. Free access to scholarly material for all users would widen the audience and recognition, and thus the impact and number of citations: Lawrence (2001) demonstrates that, in computer science, articles that are freely available online do receive more citations. A broader audience, across geographical and disciplinary boundaries, will also increase the pace and variety of research, as well providing equal opportunities for researchers to keep up to date with current information. In contrast, Esposito (2004, p.6) suspects that, due to the small readership that a technical research paper would attract, the idea of a wider audience is misleading, and that these articles virtually reach their full potential audience already:

> 'it appears that the wider readership that OA promises may be mostly illusory.'

Publishers are able to extract profits from scholarly journal articles due to a curiosity that does not apply to other forms of publishing, namely that authors receive no financial reward for publication. Authors are willing to give away the copyright to articles they have written in exchange for the services of the publisher – 'peer-review, quality labeling, marketing and disseminating' – which the author cannot or do not want to do themselves (Bjork, 2004, p.8). In return, the author's work is then published in quality, peer-reviewed journals, improving their recognition in the academic field and their career is advanced. The OA movement encourages authors who submit work to traditional journal to take more interest in the copyright declarations they are presented by publishers, and to retain copyright when possible to allow the possibility of self-archiving the work. OA advocates also remind authors that much of the work which Bjork (2004, above) attributes to the publisher is actually carried out by academics: of particular note is that peer-review is actually undertaken by researchers for

free (Harnad, 2003). Authors who are giving so much into the system should be demanding more out of it.

The BOAI describes two routes to achieve OA, namely OA journals and self-archiving. These two pillars of the OA movement tackle the issues of perceived unfair access from different angles, and thus have different advantages and disadvantages. OA journals have been growing in number and stature over the past five years, demonstrated by the growing numbers listed in the Directory of Open Access Journals (DOAJ). OA journals can retain all of the valued features of traditional journals, notably the peer-review process, and thus are in direct competition with them: for example, *PLoS Biology* was explicit in its rivalry with *Cell, Nature* and *Science* (Delamothe and Smith, 2004, p.1).

The journal structure, which places individual papers into a branded collection, is crucial to the current academic reward system, especially in the scientific, medical and technical (STM) fields. Under this system an academic's performance – and hence their pay, promotions, tenure decisions and the UK Research Assessment Exercise – can be measured by their published output. The quality of this output is judged by the impact factors of the journals in which they publish, which in turn is calculated by the number of citations that papers in a journal receives. By retaining this structure, OA journals can establish impact factor ratings and thus establish themselves in the mainstream of their discipline. There are difficulties though: it takes at least two years for an impact factor to be calculated as citation data is required, and this is from the time that ISI starts to track a journal, not when it first appears. New journals, whatever their business model, will always be at a disadvantage when competing with established, prestigious journals (Bjork, 2004). Thus, while ISI recognised two of BioMed Central's (BMC) OA journals 'as being among the top 10% of journals, by Impact Factor, in their respective fields' they were not at the time tracking BMC's most cited journal, the *Journal of Biology* (Cockerill, 2004). Advocates argue that OA journals have potentially higher readership, as their content is not trapped behind subscription or other fees, and therefore that they could attain higher impact factors.

The financial model employed by publishers such as BioMed Central and the Public Library of Science has been much discussed. They have adopted the 'author pays' model, whereby the production, peer-review and dissemination costs are paid for in advance by the author. As the distribution of information over the Internet is very low, even as the number of hits and downloads increases, subsequent access is free of charge. In reality authors rarely pay the fee themselves: several major research funders, such as the US National Institutes for Health and the Wellcome Trust 'have agreed to cover the cost of open access publishing,' and BMC's

institutional members include the NHS and 180 British universities, so virtually all of the UK's biomedical researchers can publish in BMC journals for free (Delamothe and Smith, 2004, p.2). The estimated cost of peer-review, and therefore the cost of author pays publication varies. The lowest estimate is $200, allowing Harnad (2001) to claim that 'quality control costs account for only 10% of the collective tolls actually being paid per article [through library subscriptions and licenses]. In fact BMC charges $500 to publish an article compared with PLoS's $1500. Both of these current charges are well below estimates of the revenue being made at present per article by publishers: 'well-over $4500' but varying substantially between individual journals, disciplines and publishers (Velterop, 2003, p.172). Even if the money to pay for 'author' charges does essentially still come from the same place – universities, governments, and research funders – the OA journal model seems to offer substantial savings.

Would such a model be sustainable? At the UK's Science and Technology Select Committee hearings in 2003, Nature Publishing Group estimated they would need to charge authors in *Nature* up to £30,000 per article to replace their current revenues, and predicted that most authors would be unwilling to pay this (Pickering, 2004). The high costs, they claimed, are caused by the high rejection rate of the journal – editorial costs are incurred but would not be recouped for papers not accepted. Elsevier (2004, p.2) similarly question the sustainability of the model, stating that:

> '… Open Access publishers are unlikely to cover production costs with revenues of just $1500 per article, assuming they provide similar levels of quality, peer review, functionality and accessibility as researchers receive today. They would almost certainly be unable to invest in technological innovation to any significant extent or in nurturing emerging areas of science.'

In contrast to the views of these publishers, October 2003 saw the financial institution BNP Paribas announcing that 'the commercial journal industry is not as sustainable as its Open Access competition' (Open access now, 2003). There are also other revenue streams that could be exploited in addition to 'author pays' fees: all of BMC's research articles are OA, but they charge for secondary services such as review articles; and the *British Medical Journal's* website offers content for free by gaining revenues from advertising not author charges. It should be noted, however, that falling subscription numbers have forced the *BMJ* reintroduce charges to view articles online (Grivell, 2004).

The impact that a widespread change to the 'author pays' model would have is also disputed. A wider, global audience seems certain, despite Esposito's (2004, p.8) claim that the market for research articles is not 'elastic.' Academics in institutions in developing countries would

benefit from the removal of access barriers, although many already do benefit from reduced rates. In addition the 'market' for research in different disciplines does vary, and while Esposito (2004) notes that a layperson would barely understand a technical paper, the same would not necessarily be the case of a paper in the arts or humanities.

There are potential problems though. Payment on acceptance to a journal would mean that under the OA system the countries and institutions that produce the most research could end up paying more. The UK, claim Elsevier (2004, p.9), would be hit hard by this change as it produces 5% of research papers but currently only pays 3.3% of total subscription payments. In contrast, commercial institutions would benefit from the change as they publish few papers, and so could see their costs fall by up to 90% (Elsevier, 2004, p.9). Another common criticism is that while access would be increased for resource-poor countries, institutions and individuals, this same group would be less likely to publish research due to the charges involved:

> 'It is bad enough that economic pressure prevents some scientists from having access to research results, but it is even worse to think of research that is simply not done, or done and cannot be published, because the scientists involved cannot afford to pay for its publication.' (Graczynski and Moses, 2004, p.3)

Others are concerned that the quality of articles in OA journals would be less than in traditional journals. There would be a bias towards accepting a paper for publication rather than rejection: the Internet places virtually no limit to the number or size of articles published, and rejecting a paper would mean expending effort on refereeing it but not receiving any revenue (Friend, 2004). While Suber (2002) proudly lists eight major advantages that OA journals can do as well as traditional journals (including peer review, quality and prestige) Esposito (2004) claims that this desire to retain the features of the traditional system limits the efficacy of OA journals. In particular, Esposito (2004) states that the peer review system is suitable for print publishing, where any alterations or mistakes would be expensive to fix, but it is not necessary for electronic publication. Likewise Odlyzko (1995, p.7) proposes a 'continuum of peer review' which would allow a preprint to be made available online immediately. Comments and responses could be added to the paper, and then an official peer review process would begin, taking into account any comments already made. Such a system would be more suitable to electronic publishing, it is argued, as it could increase the speed of dissemination, improve the discursive elements of academia, and still retain an official, organised and trusted system of refereeing articles. The only condition required is that a reader is made aware of what stage of the continuum each article is at.

Odlyzko's (1995) discussion of a 'continuum of peer review' introduces us to the second pillar of OA, where quality-control is of even more concern: self-archiving. Bjork (2004, p.3) identifies three forms of self-archiving: 'research-area-specific archive (e-print) servers, … institutional repositories of individual universities, and self-posting on authors' home pages.' The former two of these represent more recent, organised forms of archiving, whereas the latter comprises a dispersed system in which it is difficult to locate information. The significant amounts of research available in this way, as demonstrated by Andrew's (2003) survey at the University of Edinburgh, suggests that there is a desire amongst academics to share their research for free.

The main difference in the processes of OA journals and self-archiving repositories is the lack of peer review or quality control at the point of self-archiving. Peer review, in some form would still be required to differentiate between general information and quality, academic material. Suber (2002) states that 'the primary body of literature for which we want open access consists of peer-reviewed research articles,' implying that documents would be quality-controlled before being archived. Therefore the peer review process must still be conducted, presumably by the publishers, leading many to claim that OA self-archiving is complementary to journal publishing rather than in competition with it (Bjork, 2004).

Others disagree, and Odlyzko's (1995) discussion of the peer review process above reveals how preprints could be archived first and peer reviewed later. Whether preprints or postprints are being self-archived, the consent of publishers through copyright agreements is essential. Although Harnad (2003) does suggest a strategy to 'get around restrictive copyright legally,' involving self-archiving the preprint, whilst the copyright indisputably remains with the author, then amending the copyright transfer agreement to explicitly allow the peer reviewed article to be placed in an OA repository. If this is impossible the author can still legally add corrections to the original preprint to reflect any changes that occurred during peer review.

But can author self-archiving make research material more easily accessible? Andrew (2003) notes that his survey involved searching about half a million individual pages on the university's website, and such a distributed collection of information would be impractical for the uniformed searcher. Repositories, arranged either at an institutional level or arranged by subject area, make articles much easier to locate. One of the most well-known subject-based archive is the arXiv system for high-energy physics, used for preprints by a high proportion of researchers in this, and related, fields to both deposit and search for material. The main advantage of this system is the speed of dissemination: authors can deposit a preprint as soon as it is complete, thus establishing priority of the research and sharing research results with

the community for free. However in these areas of physics the traditional journal has also remained important in the scheme of scholarly communication and publication, especially for 'archiving and for prestige and reward allocation' (Kling and McKim, 2000, p.1308).

Harnad (2000, p37) predicts that comprehensive self-archiving is 'inevitable in all disciplines within a very short time,' but there are few disciplines so far that have achieved this and none as successfully as arXiv. Johnson (2002) is optimistic about the changes that widespread self-archiving could bring about. He argues that the traditional peer review system will become obsolete, and that the journal will be replaced by the individual paper as the unit of communication. Each paper would be ranked by quality depending on citations and comments made about it through an on-going peer review system: where access to all research is universal, the best papers will be cited and poorer quality articles will not. Johnson (2002, p.6) is convinced that the whole academic reward system would be revolutionised by this, with quantitative measures like Impact Factors being replaced by 'a more robust qualitative analysis.'

However the critical mass of OA self-archiving is far from being reached. Kling and McKim (2000) highlight an error in Harnad's (2000) assumptions above, noting the immense differences between disciplines which lead to different levels of acceptance of the OA models. ArXiv has proved popular amongst physicists partly because there was a strong culture of sharing preprints within the community stretching back to the 1970s: using the Internet was merely an application of new technology to an old system (Kling and McKim, 2000). Biomedical sciences have been less keen to self-archive preprints, attributed to 'fear [of] the clinical or social consequences' (Ware, 2004). This might explain why OA journals which incorporate peer review, such as BMC and PLoS journals, have proved more popular in these fields.

More recently development the institutional repository has emerged as the favoured tool of self-archiving. As the name suggests, these aim to archive the research output of a single university, or a small group of universities. Johnson (2002, p.3) defines an institutional repository as being 'institutionally defined, scholarly, cumulative and perpetual, and open and interoperable.' By the latter he means that if a number of institutional repositories all follow the same protocol – the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) – then searches can be conducted across the repositories as a whole making the system as a whole far more useful. A researcher does not need to search each institutional repository separately, but the set acts as a single large archive.

To this end two elements of software have been developed. The first is software for the archive itself, to store and organise papers, and make the metadata available for harvesting. Two of the most frequently used are E-Prints and DSpace, which can be downloaded free from the University of Southampton and Massachusetts Institute of Technology, respectively. In the first three months of being available, there were over 1500 downloads of the DSpace system (Smith et al., 2003). The E-Prints software is currently used by over 160 archives, although not all at the institutional scale. The second type of software is that which gathers the metadata being exposed by repositories to create a single collection, and allow searching across them. There are several web-based services available, many still in experimental stages, including OAIster and the ePrints UK project. Internet search engines can similarly take advantage of the metadata exposed by OA repositories, and the Google Scholar service includes this feature.

Institutional repositories have the same advantages as other types of author self-archiving: global accessibility, increased speed of dissemination and potentially reduced subscription charges for institutions. In addition though they can be available to more authors to deposit their work, and could therefore speed up the course towards all articles being available via open access means. The interoperability of a series of institutional repositories should also make finding information easier for researchers: a single interface could theoretically search across disciplines and institutions, facilitating inter-disciplinary research, and giving simple access to information to researchers across the globe. Harnad (2005) comments on the practical problems of achieving OA via journals and institutional repositories: while either setting up an OA journal, or converting an established journal to an OA model, is extremely risky, depositing work in institutional repositories is virtually risk free. In addition, while there are very few options when choosing an OA journal (about 5% of all journals), over 90% will allow articles to be subsequently deposited in an institutional repository. For this reason, Harnad (2003) states:

> 'Universities need to mandate the self-archiving of all peer-reviewed research
> output in order to maximise its research impact for exactly the same reasons
> as they currently mandate publishing it.'

The benefits are therefore not just for the reader, who gets better and cheaper access, or the institutions, who save on journal subscriptions, but also for authors: better access leads to more citations, better impact and ultimately better career prospects.

There are also several potential disadvantages of archiving research papers in a series of interoperable institutional repositories. The long term preservation of research may be threatened by this system, some argue, as repositories themselves – either subject based or institutional – may not prove to be permanent. New technologies and new publishing models

should not be at the expense of preserving the research record (Kling and McKim, 2000). Many OA journals which are established do not survive more than a few years, and if repositories do not reach the critical mass, then there is the risk of 'losing scientific reports that are published in what will soon become disestablished and little visited electronic islands' (Kling and McKim, 2000, p.1315). This relates to the underlying values and goals of academic endeavour: work published somewhere that is not permanent has greatly reduced value if the goal is to disseminate research as widely as possible. Lynch (2003) insists that institutional repositories must reflect the needs of the community itself, and 'advance the needs of campus communities and of scholarship today,' rather than being used to address the problems faced by libraries, or to join in with a trend amongst other universities.

Anderson (2004) discusses some of the 'author disincentives' to making their work available via OA means. One such disincentive is the 'requirement that they give up copyright in their original work if they wish to publish that work in an open access environment' (p.290). However publishing work under the traditional journal system also requires that copyright is given up by the author, to the publisher. What Anderson (2004) means is that, rather than a publisher being restrictive about what others can do with the work, OA agreements tend to be less restrictive by allowing users to copy, distribute and otherwise use the work. From a legal standpoint the OA author does retain copyright, but Anderson's (2004) point is still sound: while an experienced, multinational publisher will protect the rights given to them by copyright, a single, inexperienced author may be less able to monitor and control use of his or her work.

Clearly support for OA is not unanimous, and there remains considerable debate about the future of scholarly publishing and communication. Most agree that the Internet will have, and indeed already has had, an enormous impact on the journal system. Odlyzko (1995, p.1), for example, predicted that traditional journals would 'disappear within 10 to 20 years' although he does not predict OA as the only means of publishing that will survive. Others note the importance of reaching a 'critical mass' before any new system of publication or communication gains acceptance (Day, 2003). In any case, the technology exists for dramatic change to continue in the near future: so far it has been the publishers who have taken advantage of the new technologies rather than those trying to subvert the traditional journal system.

The exact impact of OA journals and self-archiving could have on traditional publishers is unclear. Many predict a decreased role for publishers should OA become dominant in scholarly publishing. They could remain as the organisers of an independent peer review

system while pre- and post-prints were made available via OA means. This, according to Harnad (2003), is the only *essential* aspect that publishers contribute to the current process of scholarly publishing. Ramachandran (2004) describes the potential role of the publisher as that of a midwife: they would 'provide the services of peer review and editing and hand the papers over to be posted on a self-archive.' This would still have to be paid for in some way, but if this type of change was to happen then publishers could expect their profits from serials to be much decreased.

It is interesting to note the changing responses to OA from publishers over the past few years. In 2002, Derk Haank, then chairman of Elsevier Science, explained how his company had similar aims to PLoS (Kaser, 2002). Similar to the OA publisher, Elsevier has responded to falling subscriptions, higher prices and discontent from unsatisfied users and librarians by trying to provide wider access at cheaper prices. The 'big deal', as it is often called, takes advantage of the cheap distribution costs of the Internet:

> 'The situation is that a library starts with a current paper holding, then switches to an electronic holding. It pays a small surcharge for ScienceDirect functionality and as part of the deal, normally gets a lot more access. On average a library's holding will go up by almost 100 percent' (Poynder, 2002)

Haank explains that his idea of Open Access does not free access, just immediate access to anyone who wants it:

> 'Our end goal is that we want the whole world to have access to our material 24 hours a day with no additional cost to the end user. That is also open access, but it's paid for by the librarian' (Owens, 2003, p.742).

Elsevier's (2004) comments about OA are clearly disparaging about the chances of success of such projects, and highlight the potential damaging impact on the UK in terms of the cost of publishing research. Yet at the same time the company was experimenting with several OA projects: the chemistry preprint server, ChemWeb, is owned by Elsevier, the Scirus search engine covers Elsevier journals but also numerous 'free online scholarship' sources, science text books are distributed free via ebrary, and authors are free to self-archive preprints (Suber, 2002). Many other publishers are experimenting with OA models, notably Oxford University Press which is experimenting with author pays journals and subject- and institutional repositories (Richardson, 2005), and Springer, now under the control of Derk Haank.

Why are the publishers making these experiments in OA? The language of some publishers suggest that the experiments are aimed at proving the OA models will not work, thus giving justification to the traditional models. They are also protecting their interests in electronic journal publishing, having taken early advantage of the ever improving technology of the

Internet. Do the publishers view OA as a threat to their profits, and thus want to invest in the emerging publishing models? Odlyzko (2002) predicts that even with 50-100% growth per year, OA journals or self-archiving would take 7-14 years to be at the same level as traditional publishing. At some point a critical mass will be reached, so perhaps the publishers plan to pre-empt any negative impacts they may face when this occurs. There is also the question of why publishers allow articles to be self-archived: why should they allow papers that they publish to be made available for free? Why would their customers pay for material they could access for free? Johnson (2002) states that online access to articles also requires indexing and search mechanisms in place, and Elsevier and other publishers may not be threatened by a system which seems to lack the investment to provide these additional services.

In the UK, the University of Southampton is one of the most active institutions in this area, and has one of the best established repositories, while Cambridge University is a 'collaborating institution' of the DSpace Foundation Project. By June 2005, sixteen of the top twenty UK universities by research funding had institutional repositories, and others were being planned (SHERPA). Institutional repositories have recently won the approval of the House of Commons Science and Technology Committee (2004) which stated that the desire to disseminate information and research as widely as possible must be built into the core values of universities and the academic community, and that developing institutional repositories is central this goal. They also suggested that 'Research Councils and other Government funders should mandate their funded researchers to retain the copyright on their research articles' (p.64) and criticised the Government for not taking a more active interest in a development which would increase the impact of UK research, increase access to a global audience, and save considerable sums in the higher education sector relative to the traditional publishing system (House of Commons Science and Technology Committee, 2004). The academic community in the UK cannot be seen as a closed system, however, and the Select Committee recognises that institutional repositories must be implemented on a global scale to have the desired impact. Without this, spending on subscriptions would remain high to gain access to the majority of research produced outside the UK. Yet while there is this potential for a fall in the subscription costs paid out by universities, the publishing industry in the UK is concerned: widespread uptake of institutional repositories could also damage the £750m UK publishing industry (Morris and Olivieri, 2004, p. 115). There is similar concern for the long-term survival of learned societies, many of whom rely on subscription incomes to fund their other work.

The Select Committee, however, does want to see the UK taking the lead in implementing institutional repositories, and to this end the SHERPA project, funded by JISC and CURL, has been taking major steps to develop institutional repositories in several UK universities. This cooperative approach to development will help to alleviate the concerns of Lynch (2003), who worries that 'repositories will be offered hastily and without much real institutional commitment' to appease the academic community and to fit in with current trends.

Research into the uptake and use of eprint repositories, and of OA more generally, has centred on the STM field. This may be because it is within these fields that the 'serials crisis' has hit the hardest (Crawford and Gorman, 1995), where immediate access to new results is essential (Bjork, 2004), and hence where the majority of OA projects have taken place: for example, arXiv in physics, and PubMed Central BioMed Central, and the Public Library of Science in medicine and biology. A global survey in 2002 suggested that 32% of physicists deposited their work in preprint servers, with a lower percentage for eprints, and significantly fewer academics depositing work in every other discipline questioned (ALPSP, 2002). In the same year, Lawal's (2002) study of eprint repository use in the USA and Canada reported nearly 52% of physicists using eprint archives to find information, with eight other scientific disciplines at lower proportions down to 0% for chemistry: however this paper does not clearly report the percentage posting material to these archives. Lawal (2002) does consider the reasons for 'non-use' and finds 'not relevant,' 'against the policy of publishers' and 'technological restraints' as the main factors involved, although it is not clear here whether respondents were given any other options. By considering the ALPSP (2002) and Lawal's (2002) surveys together, a clear gap is visible between use of repositories and deposit in them. Why is this the case?

Swan and Brown's (2004) more stringent analysis is helpful, and investigates authors' concerns about publishing in OA journals. The top four factors they discover relate to the perceived low 'impact factors' of OA journals: the chance of winning grants, the impact of the work, the chance of promotion and the careers of co-workers are all thought to be jeopardized by publishing in an OA journal. Would these perceptions also be felt about publishing in OA repositories? By being complimentary to publishing in traditional subscription journals, these disadvantages need not exist, but authors may not realise this. The majority of the respondents in the survey were from STM disciplines, so would there have been significant differences if arts, humanities and social sciences academics had been better represented?

Open access journals have been most successful in the biomedical disciplines, and subject repositories are most well-known in the physical sciences, but institutional repositories target the whole range of academic fields. They extend the potential audience and users of OA to the whole academic community, with the aim of documenting the entire output of an entire institution. Therefore it is now essential to extend research about authors' attitudes to self-archiving beyond the STM field. Pelizzari (2003) does make a start in examining these attitudes in the social sciences, but rather than comparing this discipline with one from the STM field, he compares attitudes and behaviours of academics in an economics department to those in a law department. He discovers little statistically significant data, hindered further by the relatively small sample size of 62. A case-study of an single Italian university may disguise broader opinions under local conditions, and could also be difficult to compare directly to the UK situation: support from government and funding bodies will be different, and even the fact that many of the academics do not publish in English will affect the outcome.

However, Pelizzari's methods are usefully described in detail. He used a questionnaire and follow-up interviews as the main data collection methods, which are the most appropriate to the situation. His questionnaire, based on a survey by the RoMEO project, helped in the design of a new questionnaire for this study. His choice of interviewees was based on Marshall and Rossman's (1995) claim that the most prominent members of the community will provide the best information. As such, Pelizzari interviewed Department Directors, Deans and the Pro-Rector but interviewing academics with a broader range of positions in the hierarchy could have yielded more varied results, and perhaps a more reliable reflection of opinion.

As well as research into the opinions and attitudes of academics towards OA, there has been recent research into actual levels of self-archiving. Notably there are two case studies of UK universities. Andrew (2003) analysed the web pages of the University of Edinburgh, noting in each disciplines how many staff there were and how many of them placed their work on the website. In the School of Science and Engineering, between 7% (Engineering and Electronics) and 32% (Informatics) of staff self-archived their work. In all the departments in the College of Humanities and Social Science and the College of Medicine and Veterinary Medicine less than 3.5% of staff had self-archived any work, with the exception of Philosophy, Psychology and Language Sciences in which nearly 13% had done so. In this university, self-archiving was certainly most common amongst scientific and technical academics, but less so amongst both humanities and medical staff. Hey (2004) conducted a similar survey of Southampton University's website, and found broadly similar results.

Engineering, science and maths disciplines had the highest proportion of papers available full-text on the website, and the highest total number of articles listed. Are these results translated into self-archiving via institutional repositories? Ware (2004) briefly reports on a survey of 45 institutional repositories in the USA: excluding the archives with no documents at all, the average number of documents was around 1250. Only one-fifth were peer-reviewed, published articles. As on the Southampton and Edinburgh websites, the most popular subjects were scientific and technical, with fewer humanities and social science documents, but no medical articles at all.

At the present time, when institutional repositories are in the early stages of development and implementation, an understanding of academics' perceptions and needs is invaluable as it can be used to plan repositories more efficiently and encourage their adoption by the academic community more effectively. It is, afterall, the authors themselves who will make or break the institutional repository initiative (Nicolas et al., 2004), and if the UK is to take a global lead in this, attention must be focused on their attitudes across the whole range of disciplines.

METHODOLOGY

Any research project involves planning the strategy and methods to be used. These are usually determined by considering what information is required to answer the aim and objectives of the study. From here a plan can be formulated to obtain or extract this information. Often there will be more than one strategy or method which will produce relevant information: in that case, the validity, reliability and generalisability of each must be considered. Validity is the extent to which the method actually measures the necessary criteria. Reliability concerns the actual data collection: is the information collected in a consistent and appropriate way? And generalisability is the extent to which the findings of a particular method could be applied to different sample or broader population.

Here, two strategies could have been used. A case-study could have been conducted: by investigating the attitudes and behaviour academics within a single institution the questions posed by this study could have been answered. However, would this approach be this best way to answer the questions? The main problem is that all the answers would reflect the local conditions of the single university – whether or not there is an institutional repository, what subjects are taught, how much research is done and so on – and so any conclusions would not be generalisable to other universities. In contrast, a broader survey strategy removes some of these problems. By targeting a wider population, a greater variety of opinions will be gathered, although depending on the size of the sample the results may still not be reliable. The sampling strategy used is crucial to achieving useful conclusions.

Within this survey strategy, the specific methods employed in this study were:
- a questionnaire, aimed at a sample of humanities academics,
- interviews with a smaller sample of those surveyed,
- qualitative comparison of information gathered from the questionnaire and interviews conducted with that from other research,
- evaluation of the content of the institutional repositories of UK universities.

In addition to these research methods it was essential to determine the context of the investigation through a detailed literature search. This was achieved through both searching bibliographic databases such as Current Awareness Abstracts and Library and Information Science Abstracts, and by using a 'pearl growing' technique to follow citations between articles. The results of this exercise – attitudes, opinions, research, surveys and case studies on a broad range of issues around open access and institutional repositories – are presented mostly in the previous section, but will also be referred to in later sections.

The most reliable way to determine the attitudes and knowledge of a particular group is through direct interrogation, so here a survey and follow-up interviews were used.

The questionnaire has the advantage of being a quick and simple method of reaching a large sample of people, whereas interviewing is far more time consuming: within the time and resource restraints of this study, only a small sample could be reached by interviews alone. Other advantages of the questionnaire include gathering both quantitative and qualitative data from a broad range of respondents, and that reliability is improved by each respondent experiencing the questions in the same medium, and even with the same wording. There is a risk of bias being introduced, however, by the difference between those who respond and those who do not: it is probable that people who are either enthusiastic users or strong opponents of OA or institutional repositories are more likely to respond: those who are unaware of the issues may think it is not relevant to them, or that they cannot help. The extent to which this did or did not occur limits the ability to generalise from the results to the wider population.

The sampling technique used also influences the generalisability of results gathered. The population in this study is all humanities academics. It is impossible to define this group, nor to calculate the size of it, introducing problems to the sampling strategy. The method chosen to take a sample from this population was to send the questionnaire to academic mailing lists, via the JISCmail service. This method was chosen because it is an efficient and reliable, yet cheap and quick, way to reach large numbers relevant people. JISCmail lists are categorised by subject area, the humanities section including the sub-sections history, archaeology, philosophy and religious studies. This gave a possible total of 202 humanities lists to use, which was narrowed to 32 by choosing lists with over 30 members, of a general discursive nature, and not limited to specific geographical areas. Five list moderators did not to forward the message onto their list, so in total 27 mailing lists received the questionnaire: an email was sent to all the lists on 22nd June 1005, followed by a reminder ten days later. This strategy reached a total of approximately 5500 individuals, as calculated by the figures on the JISCmail website.

Using mailing lists to distribute a questionnaire does have difficulties: it is impossible to know who receives the message, and indeed whether all those on the lists read the message. This, along with defining the exact size and nature of the population, mean that the results obtained were not generalisable. The academics who subscribe to JISCmail lists might be unrepresentative of the population in terms of their attitudes towards OA than those who do not. The questionnaire was sent by email, the survey itself being reached via a hyperlink in

the message. The questionnaire was designed using an online survey tool, SurveySolutions Express, which produced a professional, easy to use questionnaire, and stored and manipulated the results. Even by sending reminders, a low response rate was expected.

The questionnaire was designed to gather both quantitative and qualitative data through using closed and open questions (Appendix 1). It was designed by examining other surveys which has been conducted on similar topics, notably, Pelizzari (2003), Markland (2005), Lawal (2003), Swan and Brown (2003 and 2004), ALPSP (2004) and the ROMEO studies (Gadd, 2003). These suggested interesting lines of enquiry and helped in designing the wording of questions. Simple and unambiguous questions ensure that each respondent understands the issues correctly.

The first two questions determine the awareness of the respondents of the OA movement and of institutional repositories. By asking which of a number of terms the respondent is aware of, the question both reveals which terms or services are most well known, and gives an indication of the awareness of each respondent of OA. The latter can then be correlated with responses from other questions. The next series of closed questions determine whether the respondent's university has an institutional repository, whether they have or would use it to deposit work, and what type of documents they would use it for. Those who do or would deposit work in this way chose from a list of reasons why this is the case: the list aimed to be comprehensive, determined by examining the questions and answers of previous questionnaires. Then they had an open question to state any disadvantages or difficulties they foresee. The opposite was conducted for those would not deposit work in an institutional repository: they chose from an exhaustive list of reasons why not, then stated any perceived advantages an institutional repository might bring. Thus respondents justified their own behaviour, but were also encouraged to consider the alternative possibilities. The next series of closed questions determined whether respondents use other electronic means to make their work available, and whether they would publish work in different types of online journal. A number of facts were gathered about the respondent before they were asked for comments and whether they would take part in the interview stage of the research.

The data from the questionnaire were analysed in Microsoft Excel: the data were downloaded directly from the SurveySolutions Express website in the correct format. The data could have been manipulated and correlated in innumerable ways, but only pertinent factors were considered. Quantitative data were displayed graphically, and analysed using simple descriptive statistics. More complex statistical analysis could give more accurate results, but were not appropriate for the small sample size. Resulting correlations do not demonstrate

causation, but the qualitative data from the questionnaire and interviews helps to determine the reasons for the findings.

A number of follow-up interviews were conducted after the questionnaire to gather further qualitative, clarify views and ask further questions. From those who supplied contact details, a non-random sample was taken to include a range of opinions about depositing work in institutional repositories. Pelizzari (2003) recommended interviewing people in senior positions, but here the sample was from all levels of the academic scale to give a broader view. Face to face interviews have the advantages of encouraging rapport between interviewer and interviewee, and allowing non-verbal clues to be observed, both of which can lead to a more reliable result. However, due to the dispersed geographical nature of respondents, this was not possible. Potential interviewees were given the choice of a telephone interview or a set of questions emailed to them. The former retains an element of rapport, allowing interesting answers to be followed up and is more likely to extract opinions through the use of multiple questions. The latter however means that a record of the entire 'conversation' is kept but loses the rapport, and the ability to pose multiple questions. Both of these methods were used.

Semi-structured interviews were conducted, giving interviews a clear focus but not consisting of a fixed series of questions. Thus the interviews could be different for each interviewee and the exact form altered during the course of the interview. By email this required a series of questions and answers rather than a single interrogation. Interviews aiming to determine attitudes of the respondents, such as these, require this discursive element as these beliefs are complex, multidimensional and difficult to extract (Robson, 2002). Even when questions are carefully worded, bias can be introduced in many ways, including interviewees giving exaggerated responses or replies aimed to improve their image to the interviewer. Questions were also used to discover facts about interviewees of their use or non-use of institutional repositories. Factual questions such as these are most reliable when referring to a particular event or critical incident, so interviewees were asked to remember the last time they had published an article. Even so bias could be introduced through lapses in memory or through trying to give 'correct' answers.

While the reliability of any data gathered in interviews cannot be guaranteed, and it is not generalisable, a wealth of useful information can still be collected. Analysis of interview data was purely qualitative: quantitative analysis can be used, but not with a very small sample. Recurrent themes can be sought, and ideas not introduced during the questionnaire stage

examined. Keeping records of the interviews will allow comments made to be compared with responses to the questionnaire, and to the existing body of research and literature.

Comparison of the data gathered with those from other research in the STM fields will involve both the questionnaire and interview results. Due to the differences in the data collection methods of this and other studies, direct quantitative comparison is impossible: the sample and population would be incomparable, the questions different, and choices in closed questions not the same, for example. This does not exclude qualitative analysis, and indeed some quantitative analysis as long as it is realised that no firm inferences can be drawn. Any correspondence between the proportion of respondents answering similarly in the different studies will offer purely indicative information. Qualitative comparison is much easier: dominant or common perceptions can be correlated, thus determining issues and factors relating to the deposit of work in institutional repositories specific to academics in the humanities. The interviews gave greater indication why certain opinions are held, but again it is not possible to generalise from this data, nor can they firmly establish differences between STM fields and humanities fields.

The analysis of the contents of institutional repositories in the UK was also conducted using a survey strategy: the relatively small number of repositories and their limited contents in many cases, as demonstrated by Day (2003), means it is not necessary to limit the analysis to case study institutions. Despite Day's (2003) survey, the rate of development of institutional repositories over the last two years makes this a useful analysis. The repositories to investigate were identified firstly via the SHERPA project website, which contains an almost comprehensive list of UK institutional repositories. Secondly the OAIster website was used to identify others.

A total of 24 institutional repositories were identified and analysed. Very fine detail information could have been collected, examining individual disciplines, authors and even documents, but this was impractical given the time constraints. Such a precise analysis could have been done with a case study of a single repository, but, as before, any data would not have been generalisable. A survey gives a better overview of the diversity of ways that repositories are being used by universities.

For each repository analysed specific data were gathered:
- the total number of items deposited,
- the number of items in each of two broad categories, namely science, technical or medical (STM), and arts, humanities or social sciences (AHSS),

- the number of documents published each year,
- and any other distinguishing features.

The contents of repositories were classified in different ways in the repositories making consistent data collection difficult: some are arranged by Library of Congress Subject Headings (LCSH) others by university departments. Also some departments are difficult to classify into the two broad categories: geography, even though it contains some scientific elements, was classed in AHSS. The date of a document in the repositories is usually the date the work was published (in the case of peer reviewed, published articles) or produced (in the case of other items) although in some cases it could be the date it was deposited. It was not always possible to distinguish between these dates. The type of documents was also considered – peer-reviewed papers, preprints, theses, or other material – which was easier in the smaller archives. The contents of smaller repositories were analysed more precisely as items could be counted by hand.

This analysis provided quantitative data, but again these cannot draw concrete results. The differences between repositories are sufficient to affect the reliability of even descriptive statistics, such as calculating averages. The fact that many institutional repositories are newly established and have little content also makes inferences impossible as the next two years could see development along a different course. The main purpose of this aspect of the research, then, is to provide a snapshot of current activity in depositing work, and a view of trends over recent years. This will indicate whether there has been any significant difference in the use of institutional repositories in the UK between STM and AHSS academics up to the present time.

The combination of these methods provides an extensive approach to the aim of this project. The focus is not just on the attitudes stated by humanities academics, but also considers the actual depositing behaviour that can be determined from the institutional repositories themselves. Comparing the collected data to that from studies in the STM fields gives additional context to otherwise isolated facts.
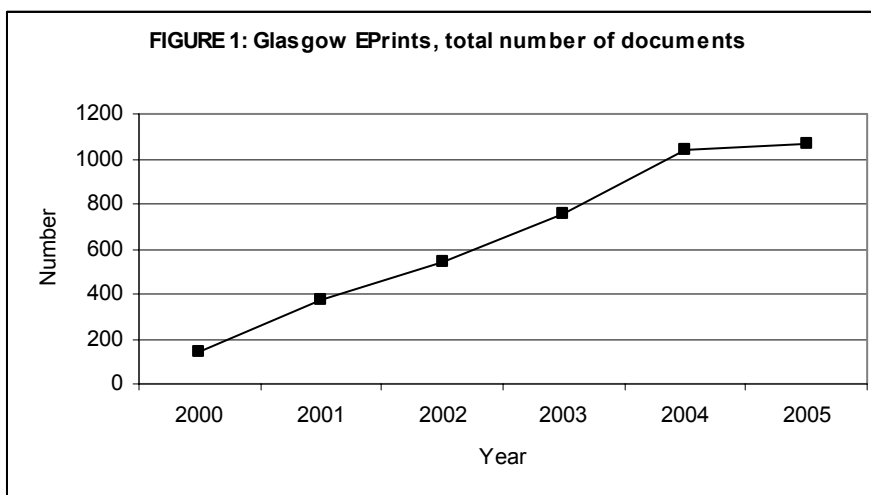
Firstly the results of the survey of the contents of institutional repositories will be presented, along with possible explanations of the findings, as this forms the context into which the questionnaire and interview results fit. The quantitative data will be described next, followed by an analysis of all the data together. This section will compare the collected data to that of previous research.

THE CONTENTS OF UK INSTITUTIONAL RESPOSITORIES, JUNE 2005

Analysis of the current contents of institutional repositories was designed to highlight the differences in the actual behaviour of academics from different disciplines in making their work available in this way. A total of twenty-five institutional repositories were identified and surveyed, as shown in Appendix 2. Glasgow University has two repositories, one for working papers, technical reports, theses and preprints, and the other for published and peer-reviewed papers. The Universities of Leeds, Sheffield and York all use the same repository under the name of the White Rose. Oxford University has a separate repository for the Institute of Mathematics, and Nottingham University likewise for Modern Languages: these two were incorporated into the totals for these universities. In the case of Southampton University, however, the four additional repositories were not included as their content was not provided exclusively by academics from Southampton University. Cambridge University's DSpace repository also contains much content which has been excluded from the analysis below: there were over 26,000 documents added by the chemistry department in 2004, mostly part of the Worldwide Molecular Matrix.

What is already clear from this description is that universities have approached the new avenue of dissemination in different ways. Cambridge University's massive archive of work, which even without the contribution of chemists is the second largest in the survey, does not exclusively contain copies of peer-reviewed papers. Analysis of this archive by department also shows curiosities: there are 302 anthropology documents, 297 of which are dated 2004. Likewise over 1,200 archaeology documents are not dated at all, and many are figures and maps from particular research projects. The content of this university's repository is made up of various documents from a small number of departments, mostly from 2003 and 2004, but not many peer-reviewed papers.

Glasgow University's use of both the EPrints and DSpace systems is unique amongst those surveyed here. The approach is useful to the researcher as it makes the distinction between unpublished and published work obvious. This is essential to the institutional repository system of disseminating work being successful: those finding the information must know exactly what it is. At Glasgow the archive of peer-reviewed, published papers (EPrints) includes considerably more documents that the archive of other material (DSpace). The EPrints archive has also shown a steady increase in the number of articles published each year since 2000 (Figure 1): perhaps academics are encouraged to deposit their work in a repository which is entirely populated with high-quality, peer-reviewed papers.

FIGURE 1: Glasgow EPrints, total number of documents

A line graph showing the total number of documents (vertical axis "Number", ranging from 0 to 1200) against "Year" (horizontal axis, 2000 to 2005). The values rise from approximately 140 in 2000, 380 in 2001, 550 in 2002, 760 in 2003, 1050 in 2004, to about 1070 in 2005.

By contrast Durham University's single repository includes a variety of document types. Over two-thirds are journal articles, although it is uncertain whether these are all peer-reviewed, but there are also preprints, conference papers, theses, book chapters and even three whole books. The level of submissions to the archive is far less than at Glasgow. This cannot be accounted for by the mixed nature of the contents, but is more likely to be due to differences in the way the repositories were implemented and publicised to the academics. The Durham University is in the early stages of utilising the repository, characterised by a single department accounting for a large proportion of the contents (Table 1).

TABLE 1: Contents of University of Durham repository by subject		
	Number	Percent
STM	15	17.0
Geography	31	35.2
Other AHSS	42	47.7
TOTAL	**88**	

Other universities have a policy of back-archiving documents. Newcastle University's oldest document dates from 1969. However, less than 5% of the total are from 1998 or earlier, and further examination reveals that almost 90% of the entire repository is from one department: computer science. The repository of St Andrew's University has a more even spread over time, with over two-thirds of the contents dated from 1998 or earlier, and 13 documents dating before 1950. Yet again further investigation shows that the majority of these older documents are from the same department: in this case most are the Computing Laboratory Newsletter. None of the twenty-five institutional repositories surveyed here had a back-archive of documents that was consistent across a number of departments.

The final repositories identified were small, virtually unused, and presumably newly established. Of the twenty-four repositories, three contained less than ten items and a further four less than fifty. Seven more repositories have less than 100 documents, and eight between 100 and 1000. Only three repositories have more than 1000 records.

The total number of documents in the twenty-five repositories was 15,036 (with the 26,000 chemistry documents from Cambridge discounted), which gives an average of just over 600 items in each. This is less than half the average of the institutional repositories in the USA surveyed by Ware (2004). In the twenty-four repositories for which data on disciplines were recorded, 19% of documents were from AHSS. This proportion can be adjusted to consider only better-established repositories by discounting those with small numbers of documents. When only considering the ten repositories with over 100 documents this average falls slightly to 17.7%. This figure does rise again to 18.2% for well-established repositories with over 500 documents.

However this latter result it misleading due to the unusual nature of the Cambridge repository, with its massive collections of archaeology and anthropology documents in particular. By only considering the other five repositories with more than 500 documents, the percentage of AHSS papers falls to just 8%. This implies that the larger, better established repositories are likely to contain mostly STM articles. A similar implication can be drawn from Figure 2 which plots the percentage of AHSS documents against the size of the repository. Without including the Southampton and Cambridge repositories on the scatter graph, because they are so much larger than the rest, there is a clear pattern: when a repository is small, up to about 200 documents in total, the range of percentage of AHSS documents is from 0 to 100%. This is due to the way that a single or small number of departments tend to dominate the contents of small repositories. And these departments can be from either STM or AHSS fields: at Warwick, 92% of documents are from physics, while at Nottingham 44% are from modern languages. With the sole exception of Cambridge, all the repositories with over 200 documents in total have less than 12% AHSS.

FIGURE 2: Percentage AHSS and repository size
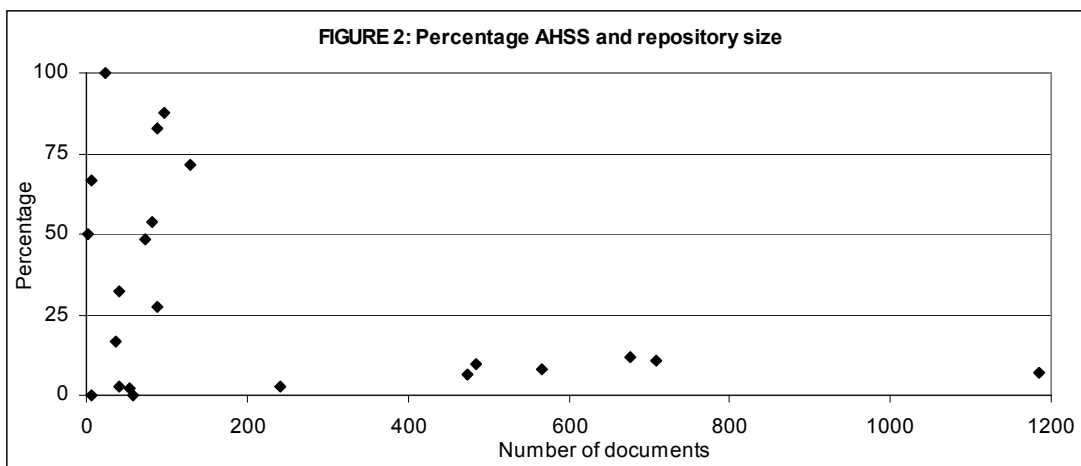
This could mean that as the smaller institutional repositories grow they become dominated by STM documents if they are not already, but this is not a certainty. The reason why the larger repositories are dominated by STM can only be guessed at: do these departments produce more papers? Probably. Do the academics in these departments perceive more benefits from institutional repositories than AHSS academics? Perhaps. Or have the universities themselves targeted particular departments to populate the repositories? More research is required here to explore the reasons for the clear patterns, although it must be noted that these figures do resemble those found by Andrew (2003) and Hey (2004) when examining self-archiving tendencies on university websites.

The chronology of the documents was recorded for twenty-four institutional repositories. Figure 3 shows the cumulative total of documents in these twenty-four repositories since 1969. Of course, this is not a reflection of depositing behaviour from 1969 onwards, but the date that the contents were published originally. There is a clear shift to a much steeper curve from around 1999. Some of the most recent documents will have been deposited simultaneously with their original publication rather than relying on archiving at a later date. The former is certainly easier, as the issue of copyright can be agreed with the publisher at the time of publication. Archiving articles after a number of years have passed introduces several extra complications. The concept of OA is relatively new so, even if an author does want to archive their work, copyright agreements from the 1990s or earlier will not provide allowances for such dissemination: are publishers prepared to give up some of their control, or do they still want to protect their financial rights? There is also the issue of whether the authors are still at the same institution: would anyone in an institution tackle all these problems to deposit the work of an academic who has long since left? Then there is the question of demand: the first few years after an article is published see the majority of its use,
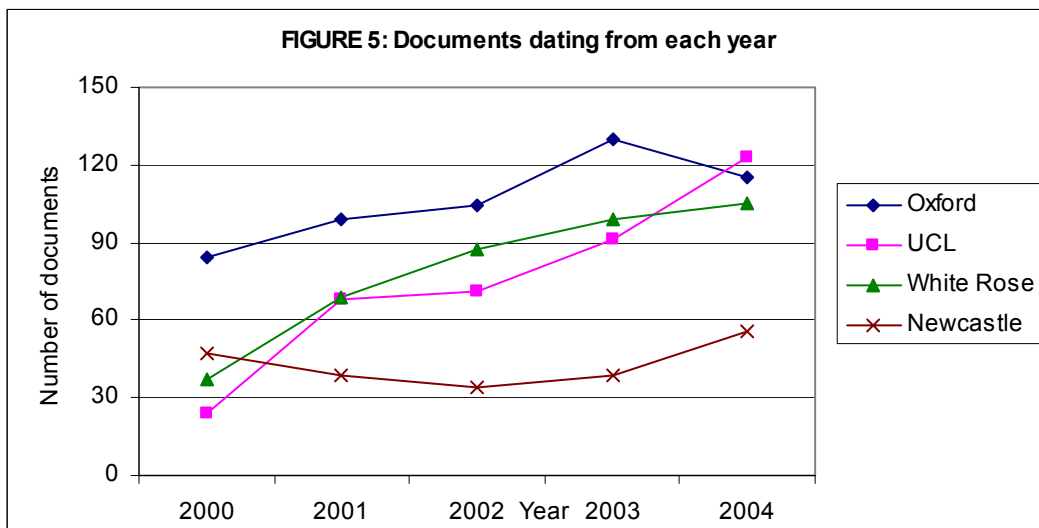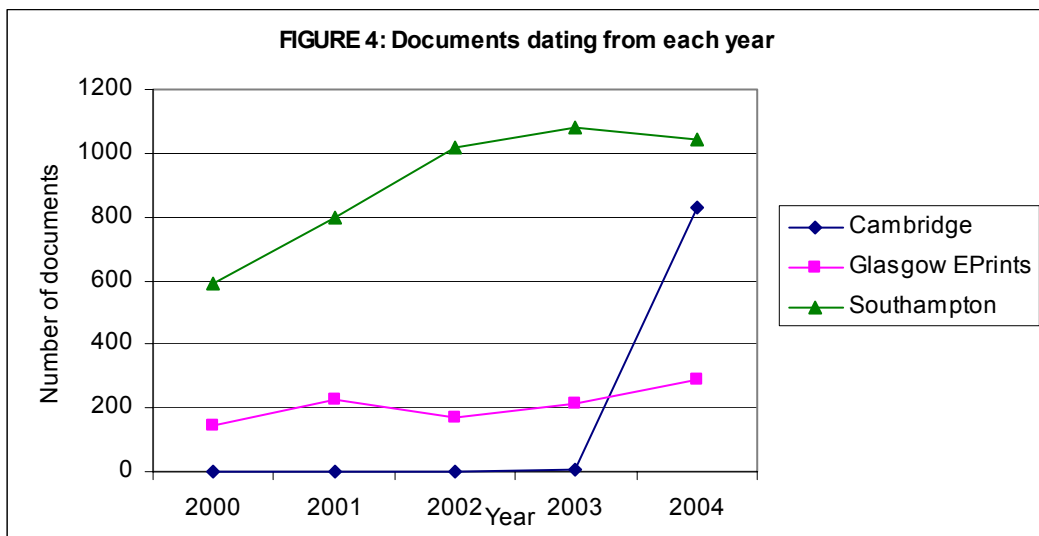
especially in STM fields, so are older articles worth the hassle when they might be infrequently used?

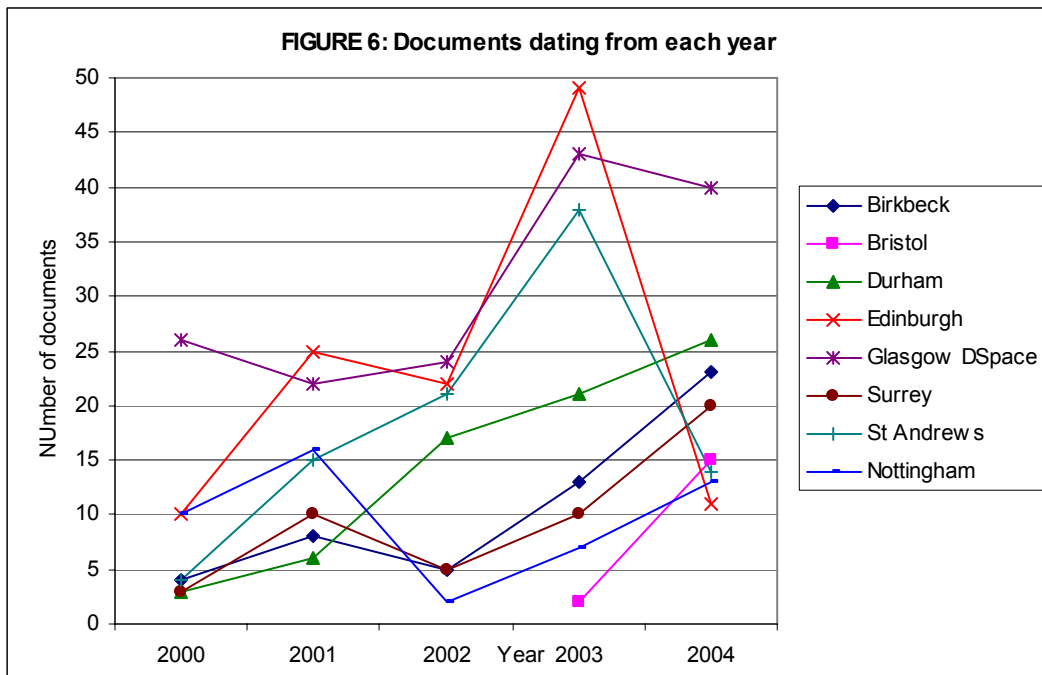**FIGURE 3: Cumlative total of documents in the 25 institutional repositories**
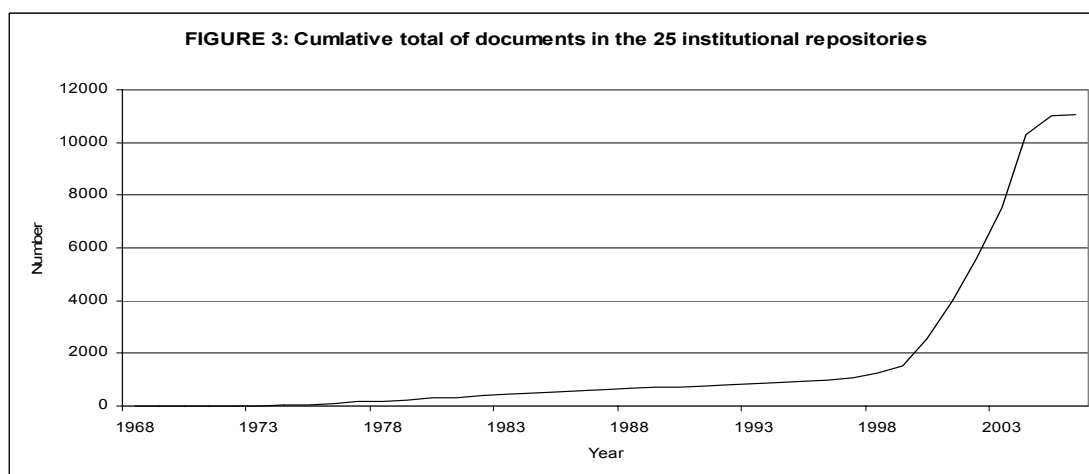
It is interesting that the total number of articles deposited in 2005 is only a quarter of the number in 2004, despite the survey being taken in July. This suggests either a catastrophic loss of interest in depositing work in institutional repositories or, more likely, a time delay between publication and deposit in the repository. This could be due to difficulties with publishers, similar to the six or twelve month embargoes before journals become OA to protect their subscription incomes, or difficulties in the process of depositing work. In contrast though, it is notable that some repositories already have documents due to be published in 2006. There is great variation in the depositing patterns. However, it is probable that the figure for 2005 will continue to rise sharply. Also likely is the continued growth in the number of documents from 2000 onwards, as these are still in heavy demand and so efforts will be made to back-archive these articles. With some many institutional repositories in the early stages of development it is also likely that the rate of increase in the total number of documents available in this way will grow sharply for many years. This initial survey could act as a benchmark against which to test these hypotheses in the future.

The final aspect of the repositories to be described is the number of documents from year within each institution. As the number of documents dated before 2000 is very small, this was only done from this year onwards. Nine repositories have been excluded from this analysis as they are so small that the data are insignificant. Figures 4-6 show the number of documents from each year in each repository. Of the three universities with large numbers of documents from each of these years, two have shown a steady increase followed by leveling off in the number of documents per year – perhaps they have reached an upper limit. Cambridge, however, had less then ten documents dated from 2003 or earlier, but over 800 dated 2004. The unusual nature of the Cambridge repository is perhaps due to the fact that

the university is one of the key implementers of the DSpace system: this may explain the enthusiasm that a number of departments showed for the repository at its initial inception. The four medium-sized repositories in Figure 5 all show healthy increases for each year, except for Newcastle which saw a decrease then an upturn. Eight of the institutional repositories (plus the other nine not included in this analysis) had fewer than fifty documents for each year from 2000 to 2005. At this lower level, there is far more variation between years within individual repositories: two of the archives notably saw a decrease in numbers of more than 60% between 2003 and 2004, although this is less significant than the overall low number of documents.

FIGURE 4: Documents dating from each year

FIGURE 5: Documents dating from each year

FIGURE 6: Documents dating from each year

This survey of the contents of twenty-five institutional repositories in the UK has revealed a picture of immense variation. Each repository has developed in a different way in terms of the publication dates of their contents. Those repositories which are better-established do have a healthy and rising rate of submissions each year, which bodes well for those which have been more recently implemented. Another aspect of the use of institutional repositories that requires further investigation is the proportion of papers that make it into the repository in different institutions: if a university is seeing 500 articles deposited in the institutional repository each year, this does not meet the goals of OA if 500 more are not available in this way. This proportion could also be investigated across disciplines and departments. This survey has only taken a glance at repositories' contents, but has discovered useful results. As well as the trends over time, it has demonstrated significant differences between the behaviours of academics from STM and AHSS disciplines, as reflected in the percentage of documents from AHSS departments.

RESULTS OF A SURVEY OF HUMANITIES ACADEMICS

In total 75 individuals responded to the questionnaire: it is not possible to calculate a response rate as the method of distribution makes it impossible to know how many individuals received the message. Over 60% of those who specified their institutions were from the UK, about another 15% each from North America and Europe, with the remaining 5 from India, China, Australia and New Zealand. Of the 75 respondents, fifty-nine indicted their status within their university. And of these 47% were PhD or post-doctoral students, with the remainder split at levels between lecturers and professors. Many of the remaining 21% who did not state their positions may not have fitted into the options given:

> 'I teach in the U.S. (as noted above) and wanted to observe that our faculty rankings are not reflected in this survey's Question 17. If it helps, I'm a Visiting Asst. Professor, which may be closest to your rank of Lecturer but I'm not entirely sure.'

Or they may not have been academics:

> 'I'm in a support service and none of the positions apply'

These two cases indicate the difficulties of using academic mailing lists to distribute a questionnaire when trying to reach a precise target audience. Thirty-seven of the respondents left contact details to take part in further research.

The answers to three pertinent questions can be used to divide the respondents into two groups: those who do or would deposit their work in institutional repositories and those who would not. Many of the questions in the survey can be analysed helpfully by dividing responses into those from these two groups. Those who knew that their university has an institutional repository were then asked if they have deposited work in it. Of these twenty-three respondents, just over half had deposited work in this way: a promising result for the OA movement, but still just 16% of the total number of respondents. Of those who knew their institution does not have a repository or who were not sure, almost two-thirds stated that they would deposit their work in a repository. Only three said they would not, and fifteen were unsure. The group of those who do or would deposit their work in an institutional repository (hereafter referred to as 'users') consists of 49 individuals (approximately two thirds of all respondents). The remained third (26 individuals) stated that they do not, would not or were uncertain about depositing work in an institutional repository: this group will hereafter be referred to as 'non-users,' as shown in Figure 7.

FIGURE 7: Formation of the groups, Users and Non-users
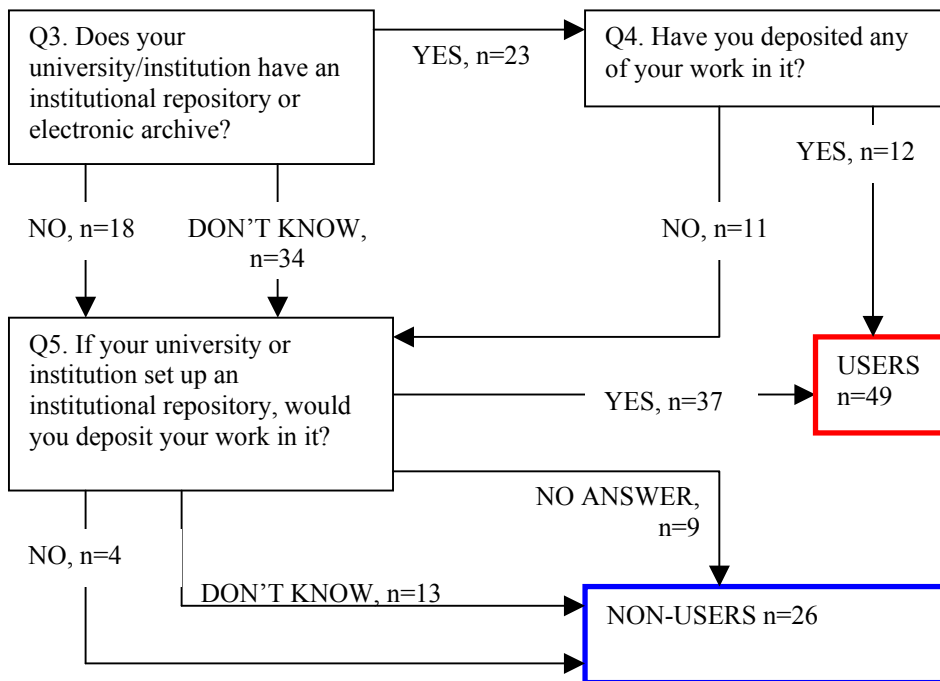
In the first question respondents were asked to select which names, services or terms they were aware of from a list of twenty-four displayed in a random order. The most well known terms (37%) were the general terms, 'the open access movement' and 'institutional repositories.' Also known by more than twenty of the respondents were the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH), the 'author pays' model and the Directory of Open Access Journals (DOAJ). The least known were Springer Open Choice (n=3), whereby authors publishing in Springer journals can choose to pay an author-fee to make the article Open Access, and the PhilSci subject-repository in philosophy (n=4). The frequencies are shown in Figure 8.
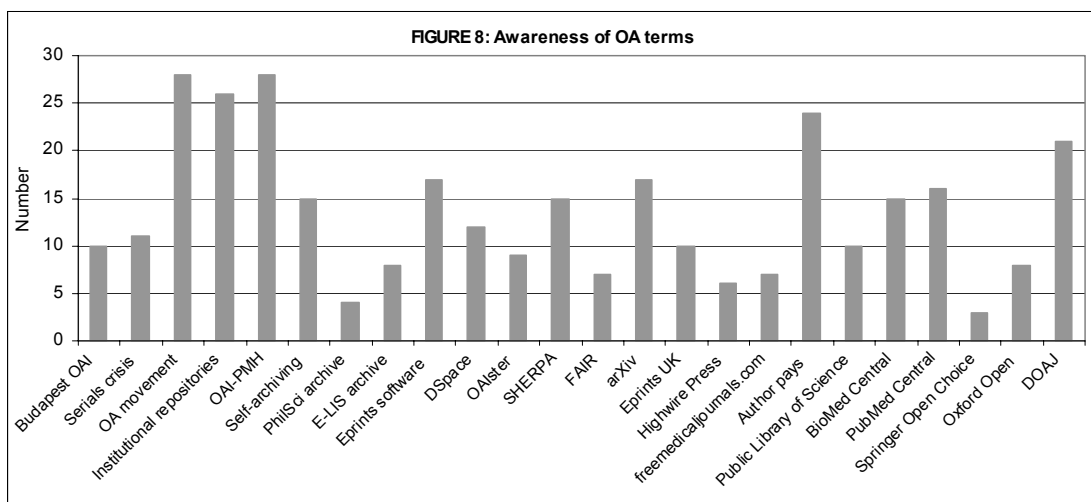
These twenty-four options can be split into three broad categories – those relating to OA generally, OA journals and institutional repositories. Figure 9 shows the percentage of respondents who are aware of none, up to one third, up to two thirds or more than two thirds in each category, and for the total of all twenty-four choices. Clearly awareness of OA services is low.

**FIGURE 9: Awareness of OA terms by category**
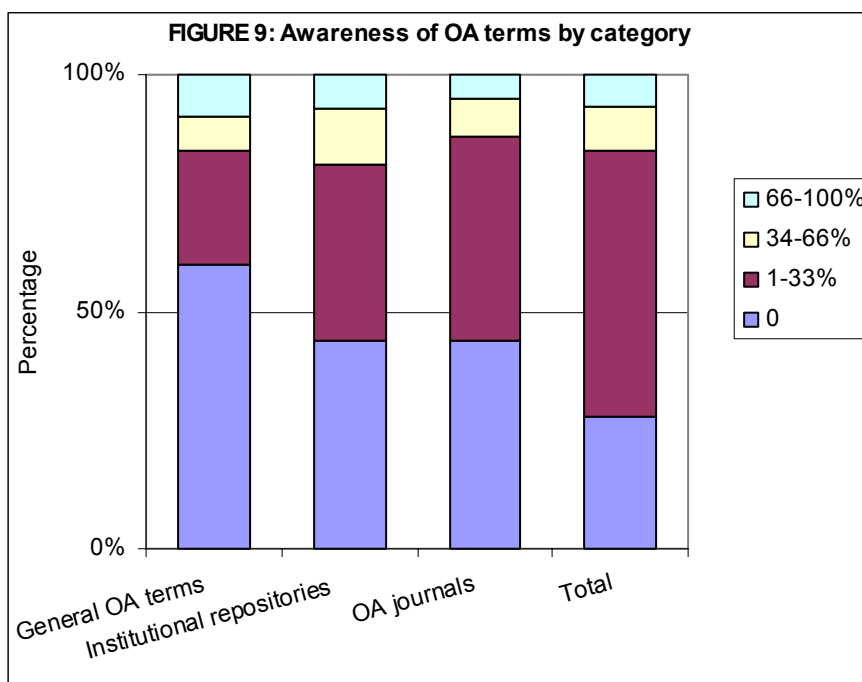
The most common way that these terms had been discovered was through searching the Internet. A similar number found out about OA terms and services through debates or via colleagues in their discipline. Relatively few respondents had first heard of these terms via another discipline or through a debate within their institutions. By comparing these answers to those from the first question, it becomes apparent that those who first heard about OA from their universities or another discipline are better aware. In particular those who heard about the terms from a debate in another discipline are the most aware of OA journals (Table 2): many of the terms in this category relate to the biomedical sciences, so perhaps this is where the colleagues in question came from. Not surprisingly, those who were aware of whether or not their institution had an institutional repository were also aware of more OA terms: an average of 7 out of 24 for those who knew that their institution did have a repository, 6 for those who knew it did not, and less than 2 for those who were not sure. Using these mean statistics hides huge variation, of course: one respondent who first knew of the OA terms through the Internet knew 19 of them, but it is unlikely that all of these would have been discovered in the same way.

TABLE 2: How respondents heard about OA, and their awareness of OA terms	Percent who found out about OA by each means	Average awareness of OA terms (out of 24)	Average awareness of OA journals (out of 9)
I discovered it myself whilst on the Internet	37	4.4	1.7
I heard a debate about it in my own discipline	19	7.8	2.5
I heard a debate about it in another discipline	6	10.0	4.0
I heard a debate about it in my University	9	10.2	1.6
I heard about it from a colleague in my own discipline	20	5.3	1.7
I heard about it from a colleague from another discipline	9	8.4	2.4

Comparison of users and non-users with their position in their universities shows that PhD and post-doctoral students are most committed to depositing their work in institutional repositories (Figure 10). Only one group has a higher number of non-users than users: professors (60% non-users, n=6), but only ten respondents were professors, so this figure does not have great reliability.

FIGURE 10: Users and non-users by academic level
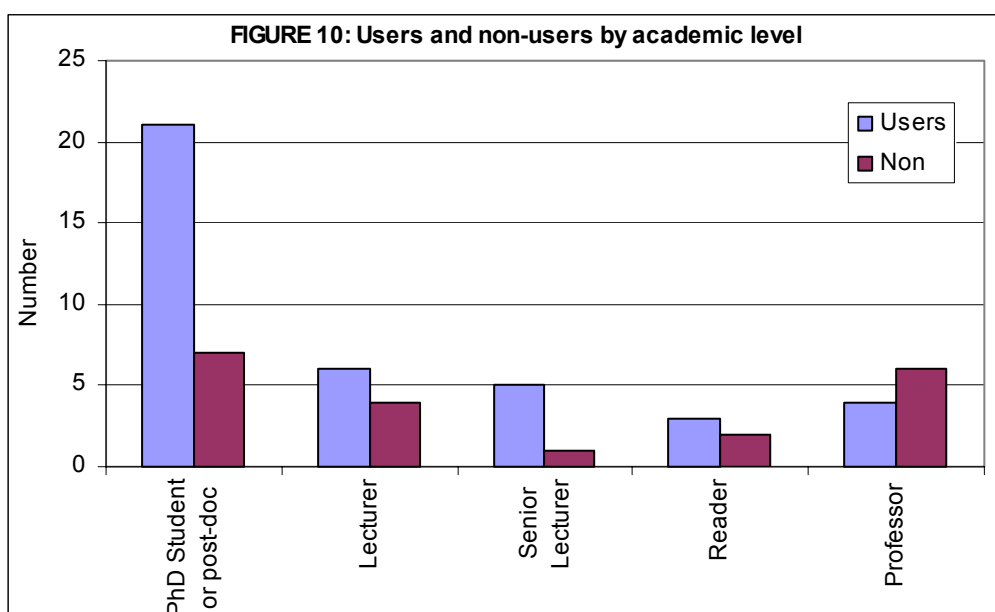
When asked who the target of research publications was, the research funder themselves was the least common target. Only 10% of respondents targeted their research funder, nearly 30% just a small group of researchers in their discipline, and over 40% chose each of the other three options: their whole discipline, all academics and the general public plus academia. However this desire to reach a wide audience does not seem to correspond to use and non-use of institutional repositories, despite the advantages that these could offer in this respect (Figure 11). While 91% of those targeting a small group of researchers would deposit work in an institutional repository, only 63% of those aiming work at the public would (Table 3). For the later group, one would logically assume that the advantage of widely disseminating work via an institutional repository would be extremely attractive.


FIGURE 11: Target audience for users and non-users
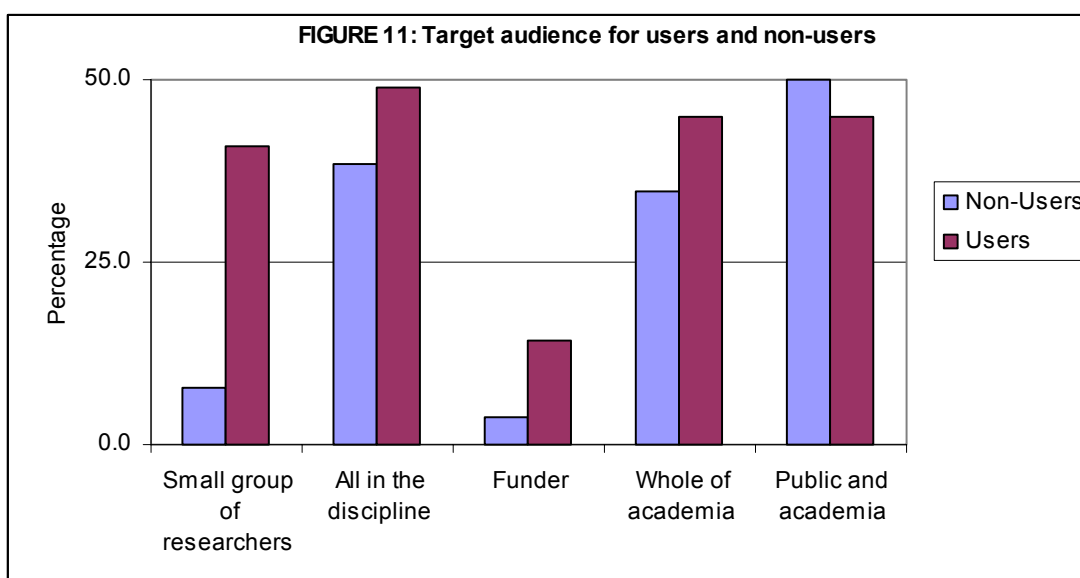
TABLE 3: Respondents targeting different audiences	
Target audience	Percent who would deposit in institutional repositories
Small group of researchers	91
All in the discipline	71
Research Funder	88
Whole of academia	71
Public and academia	63

Users of repositories – including the 41% targeting a small group of academics – are on average aware of a greater number of the terms in Question 1 than non-users (5.1 compared to

2.9 of the terms). Is it this increased awareness of OA that leads to increased likelihood of depositing work in an institutional repository? It is not possible to determine causality from the results of this questionnaire, but a consideration of the perceived advantages and disadvantages of institutional repositories may reveal a correlation. Users of institutional repositories chose which advantages they perceived from a list of seventeen. Figure 12 shows the number of respondents who ticked each statement: increased accessibility, increased impact and the permanence of the archive were the most common. The least commonly chosen were increased chances of promotion, redressing the serials crisis and reducing the expenditure of the university. The question did not have either an 'other' or 'none of the above' option. Thus, while the options given were supposed to be comprehensive, respondents may have wanted to add their own ideas. Similarly it is impossible to determine whether a respondent actively disagreed with each statement, or if they just did not attempt to answer the question: these two points reduce the validity and reliability of these answers.
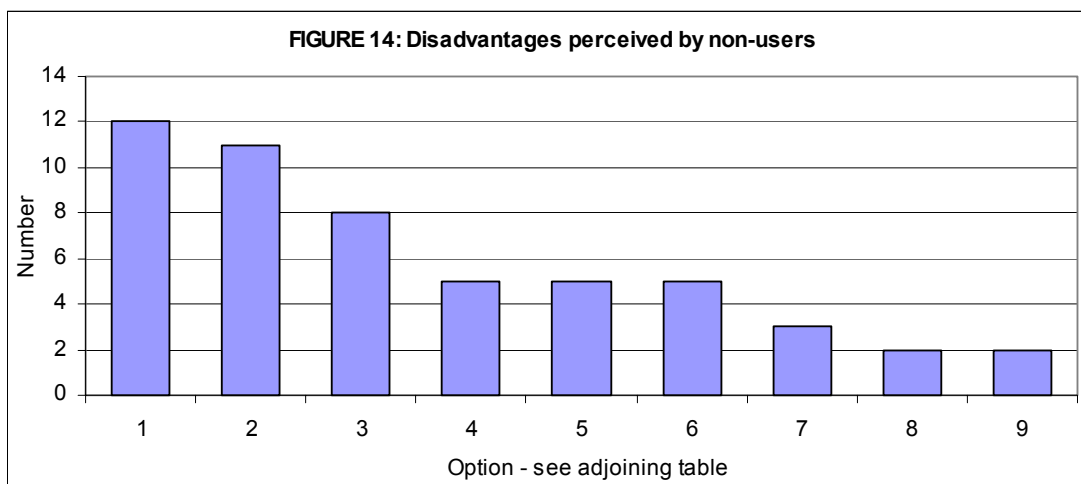

FIGURE 12: Advantages perceived by users
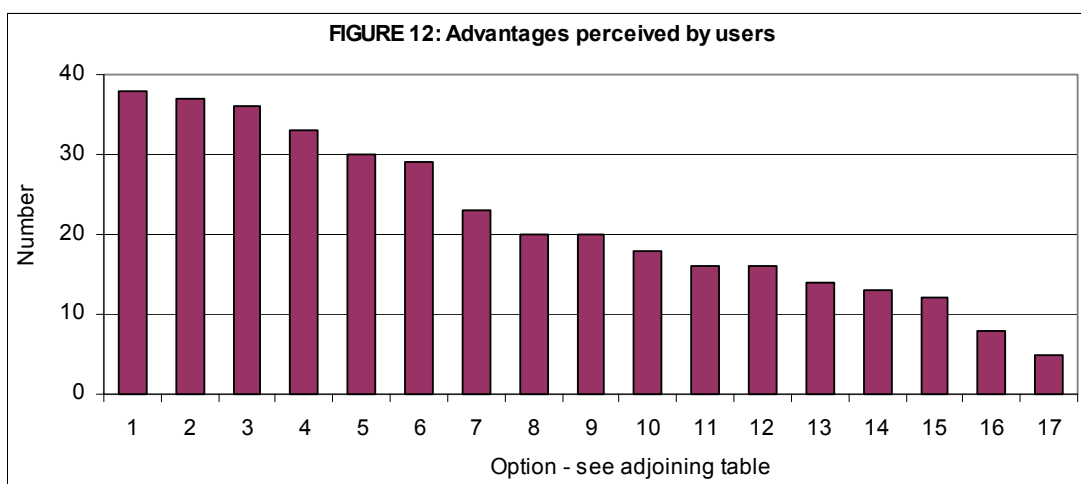
1	The accessibility of my work is increased	7	I retain the copyright of my work	13	The repository is easy to use
2	The impact of my research is increased	8	The number of citations my work gets increases	14	The repository is well indexed and archived
3	My work will be permanently archived and available	9	The public opinion of the University is increased	15	The University saves money
4	Access to the work is cheaper for others	10	I can add extra data to the work, such as photos, video, audio or datasets	16	The serials crisis is redressed
5	My published material is easy to find	11	Depositing my work in the repository protects it from plagiarism	17	My chances of promotion are increased
6	My work is disseminate more quickly	12	My work is published alongside other high quality research		

Comparing the total number of options chosen by users of institutional repositories in Questions 1 (OA terms) and 7 (institutional repository advantages) can show if there is any relationship between the two factors. Figure 13 does not show any firm relationship, but rather significant variation. Interestingly, the concentration of points along the lower portion of the chart implies that the perception of advantages of the system of institutional repositories is not dependent upon knowledge of specific OA services and terms. Some of those with low awareness are optimistic about the OA model, and are apparently quick to recognise the potential advantages of these methods of disseminating work. While this may be promising for OA advocates, the same variation still exists as knowledge of OA terms increases.

FIGURE 13: Awareness and perceived advantages
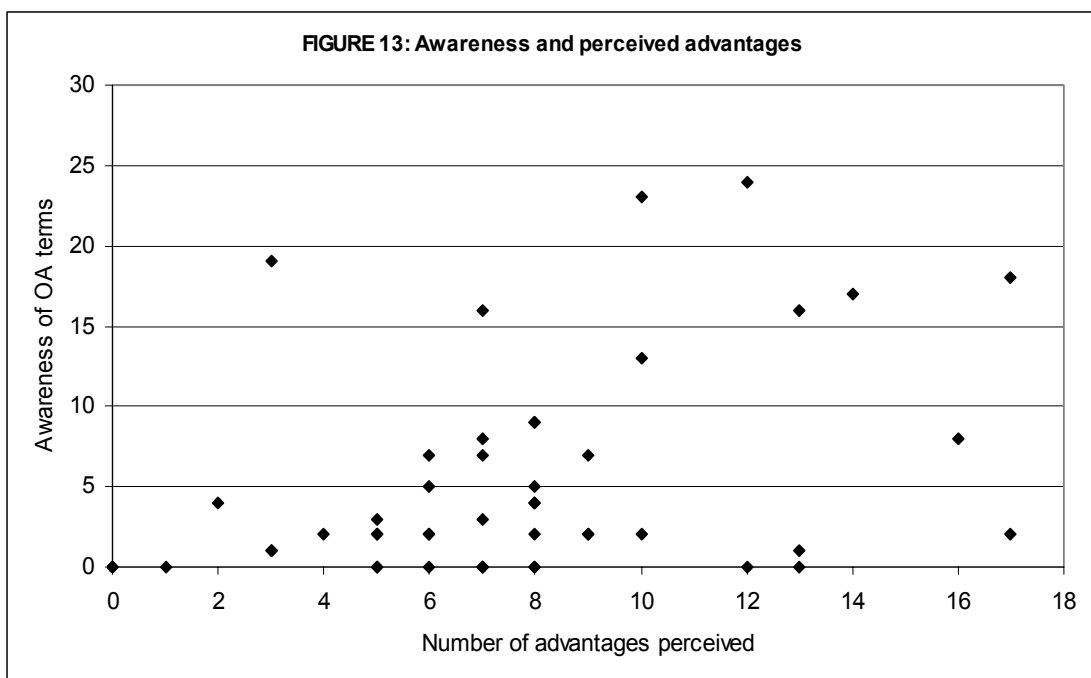
Of the 49 users, twenty-one stated disadvantages or difficulties with the institutional repository publishing system. Of these, five identified an increased risk of plagiarism from OA online documents, and another four copyright issues. The latter included worries about publishing the same article in commercial publishers' journals at a later date, and not knowing what publishers' policies were on the subject. Six respondents were concerned about technical issues, such as compatability, changes in file formats and poor indexing, and another six thought that depositing their work in an institutional repository would prove complicated or costly in other ways.

The twenty-six non-users were asked to indicate their perceived disadvantages of depositing their work in institutional repositories:

FIGURE 14: Disadvantages perceived by non-users

1	My work might be misused or plagiarised	**4**	Publishers would not let me put my work in a repository	**7**	It would be difficult and time-consuming to deposit my work
2	With no peer review or quality control process, the quality of content of the repository would be questionable	**5**	The impact of my work would be less if I deposited it in a repository, and so my promotion prospects would suffer	**8**	It would not be easy to find my work
3	I would break copyright agreements by making my work available in a repository	**6**	The Research Assessment Exercise would not take into account work in a repository, so the department would suffer	**9**	If I deposited work in a repository then I could not later publish it in a peer-reviewed journal, and so my promotion prospects would suffer

Non-users and users of institutional repositories alike seem to have the same concerns about depositing their work: plagiarism and breaking copyright agreements. There must be other factors which *users* consider of more importance than these two issues in deciding to deposit their work in repositories: increased impact, easier access for others or archiving in perpetuity, perhaps. Conversely the *non-users* of institutional repositories do not perceive any potential advantages to be strong enough to overcome their concerns: but still nearly 50% of them stated that there were advantages to be had, especially if depositing work in institutional repositories occurred on a large scale. Three-quarters of those who stated potential advantages mentioned increased accessibility for readers, as opposed to any definite benefits for the authors.

All respondents of the questionnaire were asked whether they currently placed any of these types of work elsewhere on their university's website. Amongst both users and non-users

teaching materials are most likely to be published on the university website. For users this was followed by conference papers and other work (such as photos, audio or video) (see Figure 15). Those who do or would deposit their work in institutional repositories were asked what type of work they would deposit: Figure 16 compares users' level of depositing work on the website to their potential use of the institutional repository. For almost all document types far more users would deposit work in an institutional repository than elsewhere on the university website. The only exception to this is teaching materials, which some academics may not think suitable for a freely accessible archive. Figure 16 also shows that peer-reviewed articles, conference papers and theses are the most likely documents to be placed in an institutional repository, compared to teaching and audio-visual material elsewhere on the website.

FIGURE 15: Work deposited on university websites

FIGURE 16: Use of university websites and institutional repositories to deposit
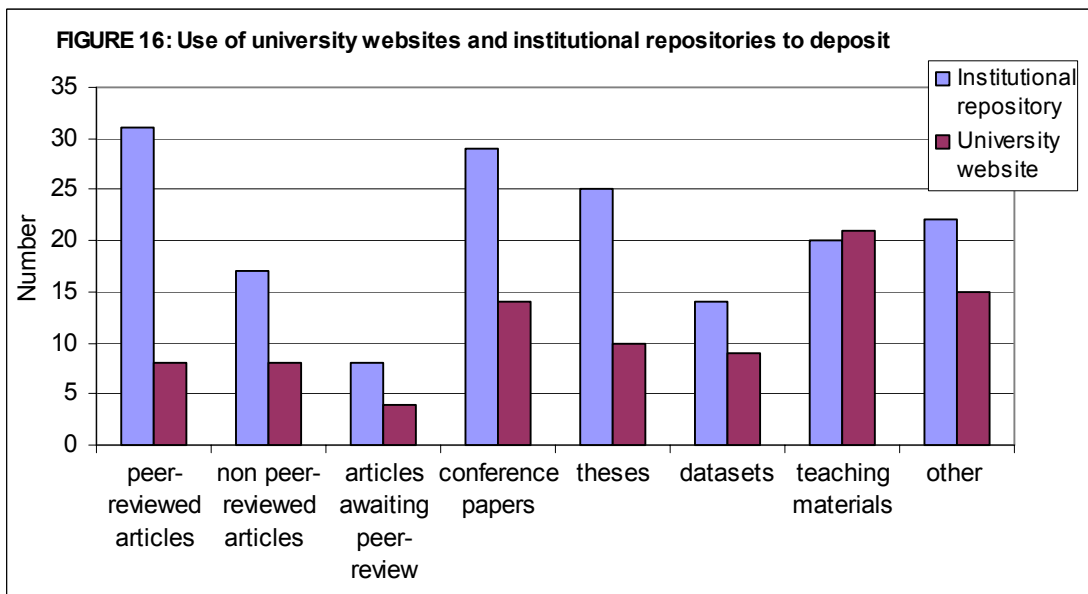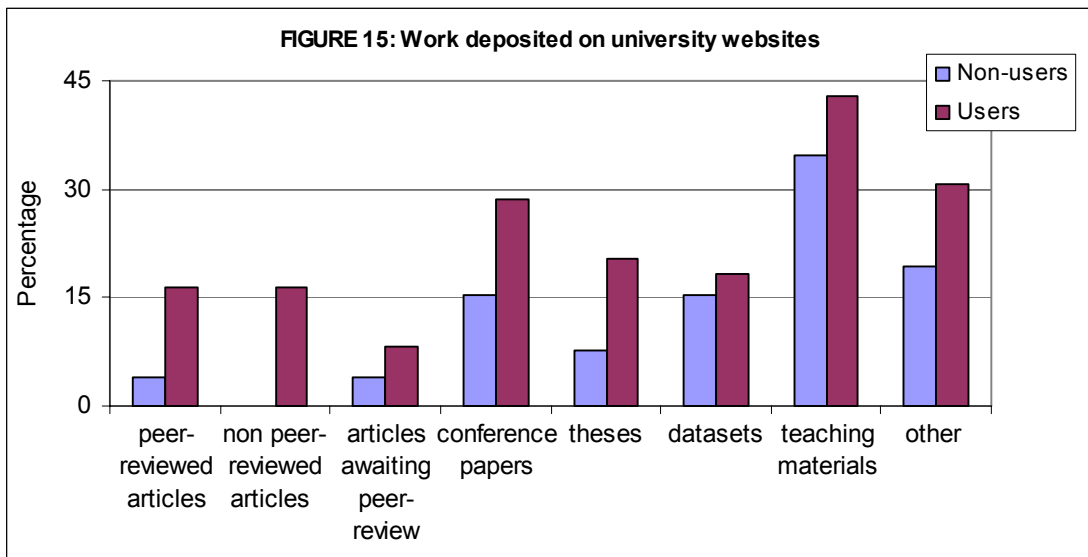
The factors influencing where humanities academics choose to publish their work were also explored through a series of three simple yes/no questions:

Question 12: Do/would you ever publish work in an electronic journal which produced *no printed copy*?
Question 13: Do/would you publish work in an electronic journal which produced no printed copy, and was *free to access*?
Question 14: Do/would you publish work in an electronic journal which produced no printed copy, was free to access, and was *not peer-reviewed*?

By adding an extra concept in each question the analysis should reveal which, if any, of the concepts is significant in the decision to publish in an OA journal or not. The expected result would be to see the number of 'yes' answers falling for each successive question, but the actual result was different, as seen in Figure 17.

FIGURE 17: Willingness to publish in online journals
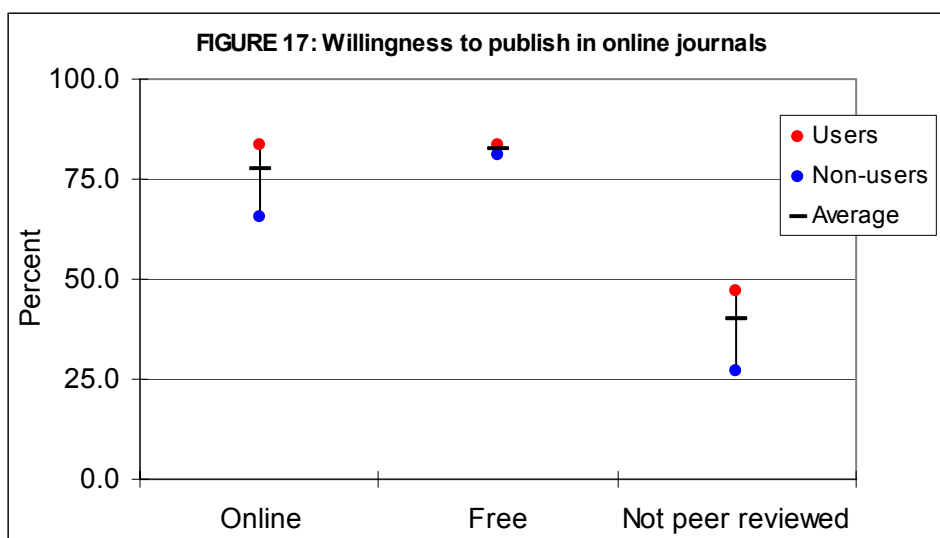
Compared to the 65% who would deposit work in an institutional repository, 77% of respondents would publish work in an online-only journal, and 83% in an online-only journal that was free to access. When asked about a non-peer-reviewed journal with the same two attributes, the proportion willing to publish fell to 40%. This latter figure is not a surprise: peer-review is clearly an important aspect of publishing amongst AHSS academics to gain recognition in the field. Interestingly 20% fewer non-users of institutional repositories would publish in a non-peer-reviewed, online-only, free-to-access journal than users. The latter group – with a higher proportion of PhD and post-doctoral students – are more likely to disseminate their work via less prestigious, non-traditional means. When deciding where to publish their work this group still considers the same issues:

'Within the discipline you pick up informally knowledge about where people you admire are being published, where the work you think is credible is being published, and then you think, Oh, I'd like to be published there.'

However, these ambitions might be compromised for other issues, for example:

'after the submission process I want it published as soon as possible. I don't want to be submitted to [name] which is very prestigious and has an awful lot of applications, wait six months then they reject it and I have to go through

the whole process again, when I could submit it to a less prestigious journal and get it published more quickly.'

Another unexpected result was that over 80% of non-users of institutional repositories *would* publish work in an online OA journal.

DISCUSSION

The questionnaire and interviews undertaken for this research have produced an enormous amount of useful data which will be analysed here. However the data can express far more in comparison with data from other surveys. In particular the data collected here about academics from AHSS fields can be compared to data collected about STM academics: are there significant differences in the way the academics from these two broad disciplines perceive OA and institutional repositories?

In this survey was that 65% of respondents stated that either they do or they would deposit their work in institutional repositories. Of course this figure masks a variety of different responses to the series of three questions shown in Figure 7. In fact only four (5%) of the respondents actually gave a firm 'no' to depositing their work in institutional repositories. This is promising for the future f OA and institutional repositories: 95% *might* deposit work. But what type and how much is not determined. How do these figures compare with the results of previous surveys?

Swan and Brown (2003) found that, across the six STM and two AHSS disciplines they questioned, an average of 11% did deposit their work in OA repositories. They did not limit this to institutional repositories, so it is likely, for example, that the high proportion of physicists using preprint archives is due to the popularity of arXiv. This 11% is close to the number in this survey who did actually deposit work in institutional repositories, and indeed the 62% who think 'reprint archives are important' is similar to the 65% in the 'users' group here. Gadd et al. (2003) show that only 9% of their respondents, of whom two-thirds are from STM disciplines, had deposited work in an institutional archive, but another 21% had done so in a subject-based archive, and almost twice as many had placed work on their own web pages. They put the low use of institutional repositories down to the small number available at the time, so this figure could be higher if repeated now. Rowlands et al. (2004), in a survey of almost 3,800 researchers of whom less than 10% were from AHSS disciplines, reported that 21% had deposited work in an institutional repository, and that 55% might do so in the future. The only survey which produced significantly different figures is that of Pelizzari (2003) which found that just 2% have deposited work in an OA archive. This variation could be due to the availability of such archives to authors of this nationality, and the particular circumstances of the single university the respondents were from.

As with the current survey, all these surveys show that current levels of deposit in institutional repositories is low across all disciplines, but that willingness to do so in the future is considerably higher. Will this potential be realised? All of these surveys are likely

to be affected by false positive reporting, as respondents try to come across in a positive light by supporting a worthy cause. In particular, when asking for probable behaviour in the future these surveys may have been affected by this phenomenon, but it remains the only valid method to determine this.

The knowledge and attitudes of academics are likely to have been reported more accurately than their probable future behaviour: they are reporting something real rather than potential. In general the respondents of this survey were aware of very few OA terms. The general terms, such as the 'OA movement' and 'institutional repositories,' proved the best known, although this might be because those who were aware of OA were more likely to respond. Others may have been put off responding to the questionnaire at all if they were unaware of all the terms in the first question:

> 'I did not recognize any services or terms in # 1. How I am to know about
> this information? Is it relevant to my research? Perhaps not...'

Rowlands et al. (2004) had similar results: they found that 34% of their respondents knew 'nothing at all' about OA, and a further 48% only 'a little.' The first of these figures is easy to compare, but the latter is a subjective term. It is surprising that, based on these two surveys, more STM researchers knew nothing about OA than the humanities academics, but that could be because Rowlands et al.'s (2004) survey was focused on a wider issue. They may have attracted more respondents who were not directly interested in OA. Also OA and institutional repositories have been increasingly newsworthy over the last two years so even in that time general awareness amongst academics must have risen. The same may account for Swan and Brown's (2003) finding that, even amongst those who had published in OA journals, over 70% knew of no eprint repositories.

Rowlands et al.'s (2004) study of mostly STM academics found that reaching the general public was a low priority for researchers: 'fellow researchers' and 'other researchers' were the most important. But this study's questionnaire of humanities academics produced very different results: reaching the general public and academia was important to these respondents. There could be many reasons for this. Firstly many STM papers are of little interest to those outside the field of research: there is not wider potential market as the public have no interest (Esposito, 2004). Comparatively, they are far more likely to be interested in articles about history, literature or other AHSS subjects. There may be more demand from the public, but does this translate into different methods of disseminating work? Not necessarily:

> 'I know that a journal will be read by a certain group of people who I need to
> get my work out to, whereas an institutional repository in theory has a wider

number of people looking at it, may not have the specific people who I want to be reading it.'

The respondents from the Rowlands et al. (2004) study were 'senior researchers,' whereas the respondents from this research comprised 37% PhD or post-doctoral students and only 13% professors. The younger, less established academics, at the beginning of their careers, may have more to gain from distributing their work more widely, and might have a more utilitarian view of publishing their work:

'For me, there is no point conducting academic work if you don't desire its dissemination far and wide.'

This suggestion is supported by Pelizzari (2003) who documents the variation between the willingness to self-archive of academics of different levels. Nearly 20% more of those with 'non steady employment' were willing to self-archive than those in steady employment. This lends support to the findings of this research that academics in lower positions are more likely to deposit work in institutional repositories.

Rowlands et al. (2004, p.37) do find that those respondents with experience of OA publishing do 'have a broader communications agenda,' but the direction of causality cannot be established. Is reaching a wider readership just a welcome side-effect from the new publishing model? In any case, the same has not been replicated in this survey about institutional repositories. Approximately half of both users and non-users of institutional repositories are targeting the general public, but paradoxically a far greater proportion of users of institutional repositories are aiming to be read by a small group of researchers in their field than amongst non-users.

This unusual result could link to the other perceived advantages of depositing work in institutional repositories. In physics, for example, the arXiv repository has proved very popular, but it is not the fact that it can be accessed by the general public that makes it so. Rather it facilitates the sharing of preprints in a quicker and more efficient way. By being open access the documents are available to *all* the interested researchers, and any potential additional readership is of little consequence. The data from the humanities survey suggest that they too are more interesting in targeting other academics. 75% of users see increased impact of their work as an advantage of depositing in institutional repositories. Qualitative data from the interviews supports this:

'One [purpose of publishing work] is to gain peer recognition in my field, the benefits of that is that it allows me to be in a stronger position if I want to have further research funding or if I wanted a job. So, peer recognition is

achieved mainly, I think, through publication in a peer reviewed journal. That I guess is the main benefit.'

And:

'*What influences your choice of journal to publish in?* ...the prestige of the journal (Research Assessment Exercise brownie points)...'

However, despite recognising the likelihood of increased accessibility and impact, the humanities academics do not link using the institutional repository directly to benefits to themselves:

'The advantages would be to readers, not authors.'

The lack of these end goals as perceived advantages may limit their enthusiasm for depositing work in repositories. It also reveals a need for these benefits to be explicitly stated to authors if institutional repositories are to succeed.

On top of not recognising all the potential benefits, it is clear that, even the users of institutional repositories, still have significant concerns about them. One major limitation perceived is that they may not be permanent:

'Funding bodies might be doing this [mandating the deposit of work in institutional repositories] under the misguided illusion that institutions are eternal 'beings.' Many departments close down; entire Schools disappear – what will happen to defunct institutions' repositories? And WHO will look after them/maintain them? And who will pay for them?'

And in addition to these practical issues, there are also technical ones:

'Digital = fragile'

The analysis of factors which influence publishing work in online journals relates to this last comment. Amongst non-users of institutional repositories, more would publish work in an OA journal than a subscription-based online journal (although the figure is above 50% in both cases). This may be a reflection of the fact that online journals, whether author-pays or subscription-based, are not traditional. Why publish in an online-only subscription journal when the same would be achieved in a traditional print journal, and without the risk of the journal failing and access to the work being lost. In contrast the OA journal only exists online, making this a more popular option: it is more radical, but with more to gain through the online-only route. Afterall evidence does suggest that publishing in OA journals is more common than depositing work in eprint archives.

Gadd et al.'s (2003) survey showed another issue that was of concern for both archivers and non-archivers. The most common concern that their respondents raised was that self-archiving work might interfere with publishing the same work in traditional journals. The same concerns were also apparent amongst the humanities academics, but not as strongly as amongst Gadd et al.'s (2003) mostly STM respondents. The latter survey did not report on

whether respondents were concerned about the permanence of archives – presumably they were not asked their opinion on that issue. While peer-review is central to maintaining the quality of published research in both STM and AHSS fields, the two broad disciplines do have different approaches to disseminating research. In STM fields dissemination of research before it has been peer-reviewed has been common for many years, and this has contributed to the success of the arXiv preprint archive (Esposito, 2004). By contrast in AHSS fields there is less interest in disseminating preprints: research does not move as quickly and is less competitive, so registering research at the earliest possible moment is not essential. Thus the relationship between publishers, copyright and self-archiving will be very different in STM and AHSS fields. In the latter, copyright agreements are often already in place before the process of self-archiving begins, adding an extra complication:

'Depositing takes time to check publisher policies on copyright.'

Swan and Brown (2004) quantify the concerns that respondents have about publishing in OA journals. For those who have published in OA journals, of whom only 6% were from AHSS disciplines, the main concerns involved the impact that publishing in this way, as opposed to traditional journals, would negatively affect their career progression in some way. The same issues were of concern to non-OA authors, 16% of whom were from AHSS fields, but to a greater extent.

From all these surveys it is certain that those difficulties perceived by non-users of OA methods are also common concerns for those who do deposit work in repositories. Swan and Brown (2003, p.32) also show this: institutional 'reprint' and 'preprint' archives are considered to be less important than discipline-based archives, and all far less than traditional journals. The future of academic research publishing remains highly uncertain in the perceptions of researchers, and many concerns contribute to author inertia.

Rowlands et al. (2004) asked their respondents what type of documents they publish in institutional repositories. The options given by Rowlands et al. and those in this survey differ, but comparison can still be made. The surveys both reveal a high use of institutional repositories to disseminate theses and conference papers, which are difficult to distribute through traditional publication methods. Both also show a significant level of use for peer-reviewed papers which have been published in or accepted by traditional journals: this is crucial as it is these documents that are the most important in scholarly communication, and the Harnad (2001) is targeting for deposit in this way. The main discrepancy between the two surveys is in the level of use of institutional repositories for pre-prints: the mainly STM survey produced a figure of nearly 30%, but the figure for AHSS academics was about half of

this: some reasons why AHSS academics do not disseminate pre-prints have already been discussed.

Rowlands et al. (2004) also considered what work their respondents self-published on home pages or Departmental websites: accepted papers, conference papers and multimedia were the most common, followed by pre-prints. Interestingly theses were not commonly disseminated in this way, despite their common deposit in institutional repositories. A similar pattern was seen amongst the humanities academics, but making peer-reviewed articles which had already been published available in this way was not common. The reason for this difference is difficult to explain, especially as both groups would disseminate this type of document via an institutional repository. It may reflect a historical difference in the way that these groups have exchanged their work: both Andrew (2003) and Hey (2004) show that academics in certain science and technology fields have long used informal deposit of work in home pages to distribute their work, but this practice was not as common in AHSS fields. These two surveys by Andrew (2003) and Hey (2004) reveal a baseline of self-archiving via websites which is reflected in the actual current behaviour visible in the twenty-five institutional repositories surveyed here. AHSS academics certainly lag behind their STM counterparts in self-archiving.

CONCLUSIONS

This research has approached several objectives through a number of methods. It has been established that academics from humanities disciplines are generally positive about the role that institutional repositories will play, and that most would deposit at least some of their work in them. Notably, while these respondents did perceive several advantages of the system, they considered the benefits to be mostly for people accessing the information, not for the authors themselves. They had poor awareness of terms and services associated with open access, with only the very general terms being well known: many of them first found out about the open access movement whilst searching on the Internet. Whilst making their work available via the Internet was not a problem for most respondents, peer-review was essential to most of the respondents: sharing of pre-prints has not been strong in the humanities, and does not look at though it will be common in the near future either.

The current contents of UK institutional repositories is mostly from the STM fields. There is also a concentration of documents from the last five years, with steadily rising numbers of documents dated in each year. The use of repositories varies considerably between universities, with many of those with few documents being dominated by a small number of departments, sometimes from the arts, humanities or social sciences. However, the larger repositories – in fact all but one with more than 200 documents – are dominated by documents from STM fields.

There are definite differences in the attitudes and behaviour of academics from different disciplines in depositing work in institutional repositories. Current behaviour indicates far lower levels of deposit by arts, humanities and social science academics. Many of the attitudes held by authors from different disciplines are common: the advantages revolve around increased access and impact of work, whilst the disadvantages are the difficulties of establishing publishers' policies and the quality control mechanisms in place. However, humanities authors – both those who would and would not deposit their work – are more concerned about the potential for plagiarism, and the permanence of the archives themselves. A significant proportion of them are undecided over whether they would deposit work in an institutional repository. These could be caused by less exposure to electronic means of disseminating information in the past. Whilst this is the case, it is clear that there is the potential for growth in the use of institutional repositories within the humanities to the same level as STM fields.

Much evidence has been collected and compared during this research project. Most of it suggests that the level of use of institutional repositories by humanities academics is very low. The contents of the twenty-five UK institutional repositories surveyed here show just 19% of documents archived in this way were from arts, humanities or social science department. This reflects the level of self-archiving on university websites, shown by Andrew (2003) and Hey (2004). The significance of this is that a pre-existing culture of self-archiving or preprint sharing within a discipline translates into a higher level of institutional repository use (Ware, 2004). In addition to this, the level of awareness of humanities academics of OA is low – although this is consistent with academics more broadly. The majority of literature about Open Access of all types, and the majority of OA journals or subject-based repositories are in STM fields. This experience and knowledge, however, is not being transferred to academics in the humanities, and nor are they being informed about OA from their universities.

All these factors, plus the concerns that humanities academics have about depositing work in institutional repositories, seem to contradict Harnad's (2000, p.37) enthusiasm that comprehensive self-archiving is 'inevitable in all disciplines within a very short time.' However, open access and especially institutional repositories are in the early stages of development, so these results are a product of this particular stage. The questionnaire results point to a more promising future for the advocates of OA. Around two-thirds of the respondents either already do, or would when the chance arose, deposit work in an institutional repository. And a further 17% are undecided on whether they would deposit their work in this way. Despite their low level of awareness of OA, and their current low usage, humanities academics are positive about the advantages that they could bring, especially for readers. They do still have concerns about the institutional repository model, but these do not seem important enough to stop them depositing work.

On top of this, more and more universities in the UK are establishing institutional repositories. The majority of the top-ranked research universities have already established them, and more are planned. The JISC-funded Focus on Access to Institutional Repositories (FAIR) programme and SHERPA projects, as well as the Resource Discovery Network's (RDN) E-Prints UK project, will also support and encourage further use of institutional repositories, and add further advantages to both readers and authors.

Research funders are also taking a keener interest in Open Access and institutional repositories. The Research Councils UK (2005) position statement is especially important for the UK. They state:

> 'Where research is funded by the Research Councils and undertaken by researchers with access to an open access e-print repository (institutional or subject-based), Councils will make it a condition for all grants awarded from 1 October 2005 that a copy of all resultant published journal articles or conference proceedings (but not necessarily the underlying data) should be deposited in and/or accessible through that repository, subject to copyright or licensing arrangements.'

The final clause of this quotation does mean that some articles will be exempt from this condition, but the majority of research will have to be deposited. But this remains a powerful statement, and especially so in for academics in the arts and humanities:

> 'Given the state of funding in the arts, then yes. ... Funding is so scarce that I couldn't afford to turn my nose up and say 'no I'm not going to do that [deposit work in an institutional repository].'

So the combination of factors outlined above points to a massive increase in the amount of material available via open methods over the next few years. In particular, given the support from research funders and universities, the deposit of articles in e-print repositories seems assured. This is the method which will least effect the traditional journal publication process, but in the longer term the impact on the publishing industry is unclear. Harnad (2003) expects that publishers will give up their role in disseminating research, and just retain the role of coordinating the peer-review process.

What ever the future impact of institutional repositories, this research has been conducted at a time on the cusp of change. As such it, along with other research which has been conducted in recent years, will act as a baseline against which future developments can be measured.

There are also many other issues which need to be addressed at the present moment, along with research into the likely impacts of widespread institutional repository use. Firstly a more detailed examination of the current contents should be conducted: it is not just the number of documents which are deposited which is important, but also the proportion of post-prints which are deposited within each discipline. Further studies into the attitudes of academics from all disciplines could be conducted to discover how willing they are to deposit post-prints. Both positive and negative opinions of open access and institutional repositories are required to determine how best these systems of dissemination should develop: such discipline-specific knowledge is important at this stage while these developments are still in their infancy.

BIBLIOGRAPHY

ALPSP, 2004. *Authors and electronic publishing: the ALPSP research study on authors' and readers' views of electronic research communication*. Worthing: ALPSP.

ANDERSON, R., 2004. Author disincentives and open access. *Serials review* **30** (4) pp.288-291.

ANDREW, T., 2003. Trends in self-posting of research material online by academic staff. *Ariadne* [online] **37** [cited 7 July 2005] <http://www.ariadne.ac.uk/issue37/andrew/intro.html>

BIOMED CENTRAL [online] [cited 26 January 2005] <http://www.biomedcentral.com>

BJORK, B., 2004. Open access to scientific publications: an analysis of the barriers to change. *Information research* [online] **9** (2) [cited 17 February 2005] <http://informationr.net/ir/9-2/paper170.html>

BRAND, A., 2001. CrossRef turns one. *D-Lib magazine* [online] **7** (5) [cited 20 January 2005] <http://www.dlib.org/dlib/may01/brand/05brand.html>

CHAN, L. et al., 2002. *Budapest open access initiative* [online] [cited 14 July 2005] <http://www.soros.org/openaccess/read.shtml>

CLAYTON, C., 2004. STM research literature and the general public. *Library and information update*, **3** (10) p.15.

COCKERILL, 2004. Assessing the impact of Open Access. *Open access now* [online] **23** [cited 26 January 2005] <http://www.biomedcentral.com/openaccess/miscell/?issue=23>

CRAWFORD, W. and GORMAN, M., 1995. *Future libraries: dreams, madness and reality*. Chicago: American Libraries Association.

CROW, R., 2002. *The case for institutional repositories: a SPARC position paper* [online] [cited 12 May 2005] <http://www.arl.org/sparc/IR/ir.html>

DAY, M., 2003. *Prospects for institutional e-print repositories in the United Kingdom: ePrints UK supporting study, no. 1, version 1.0* [online] [cited 7 July 2005] <http://www.rdn.ac.uk/projects/eprints-uk/docs/studies/impact>

DELAMOTHE, T. and SMITH, R., 2004. Open access publishing takes off: the dream is now achievable. *British medical journal*, **328** pp.1-3.

DIRECTORY OF OPEN ACCESS JOURNALS [online] [cited 23 May 2005] <http://www.doaj.org>

ELSEVIER, 2004. *Elsevier's comments on evolutions in scientific, technical and medical publishing and reflections on possible implications of Open Access journals for the UK* [online] [cited 17 June 2005] <http://www.elsevier.com/authored_news/corporate/images/UKST1Elsevier_position_paper_on_stm_in_UK.pdf>

EPRINTS, Archives using GNU EPrints [online] [cited 17 February 2005] <http://software.eprints.org/archives.php>

E-PRINTS UK PROJECT [online] [cited 14 July 2005] <http://www.rdn.ac.uk/projects/eprints-uk/>

ESPOSITO, J., 2004. The devil you don't know: the unexpected future of Open Access publishing. *First Monday* [online] **9** (8) [cited 21 January 2005] <http://firstmonday.org/issues/issue9_8/esposito/index.html>

FRANK, M., REICH, M., and RA'ANAN, A., 2004. A not-for-profit publisher's perspective on open access. *Serials review* **30** (4) pp.281-287.

FRIEND, F., 2004. How can there be open access to journal articles? *Serials* [online] **17** (1) pp.37-40 [cited 14 May 2005] <http://www.metapress.com>

GADD, E., OPPENHEIM, C., and PROBETS, S., 2003. RoMEO studies 2: how academics want to protect their open access research papers. *Journal of information science* **29** (5) pp.333-356.

GRACZYNSKI, M. and MOSES, L., 2004. Open access publishing: panacea or Trojan horse? *Medical science monitor* **10** (1) pp.ED1-ED3.

GRIVELL, L., 2004. Access for all? *EMBO reports*, **5** (3) pp.222-225.

GUEDON, J., 2004. The 'green' and 'gold' roads to open access: the case for mixing and matching. *Serials review* **30** (4) pp.315-328.

HARNAD, S., 1994. *Scholarly journals at the crossroads: a subversive proposal for electronic publishing* [online] [cited 18 July 2005] <http://www.arl.org/scomm/subversive/sub01.html>

HARNAD, S., 2000. E-knowledge: freeing the refereed journal corpus online. *Computer law and security report* [online] **16** (2) pp.78-87 [cited 17 February 2005] <http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad00.scinejm.htm>

HARNAD, S., 2001. The self-archiving initiative. *Nature webdebates* [online] [cited 17 February 2005] <http://www.nature.com/nature/debates/e-access/Articles/harnad.html>

HARNAD, S., 2003. Open access to peer-reviewed research through author/institution self-archiving: maximising research impact by maximising online access. *Journal of postgraduate medicine* [online] **49** (4) [cited 14 April 2005] <http://www.jpgmonline.com>

HARNAD, S., 2004. Fast-forward in the green road to open access: the case against mixing up green and gold. *Ariadne* [online] **42** [cited 13 August 2005] <http://www.ariadne.ac.uk/issue42/harnad/>

HEERY, R. and ANDERSON, S., 2005. *Digital repositories review* [online] [cited 14 July 2005] <http://www.jisc.ac.uk/uploaded_documents/rep-review-final-20050220.pdf>

HEY, J., 2004. Targeting academic research with Southampton's institutional repository. *Ariadne* [online] **40** [cited 13 August 2005] <http://www.ariadne.ac.uk/issue40/hey/intro.html>

HOUSE OF COMMONS SCIENCE AND TECHNOLOGY COMMITTEE, 2004. *Scientific publications: free for all?* London: The Stationery Office Limited

HUBBARD, B., 2004. *Evidence on scientific publications for the Science and Technology Committee of the UK Parliament from the SHERPA project* [online] [cited 4 April 2005] <http://www.sherpa.ac.uk/documents/SHERPA_Evidence.pdf>

JISCMAIL, Humanities [online] [cited 23 February 2005] <http://www.jiscmail.ac.uk/mailinglists/category/Humanities.htm>

JOHNSON, R., 2002. Institutional repositories: partnering with faculty to enhance scholarly communication. *D-Lib magazine* [online] **8** (11) [cited 17 February 2005] <http://www.dlib.org/dlib/november02/johnson/11johnson.html>

KASER, D., 2002. Ghost in a bottle: Elsevier Science chairman Derk Haank responds to the Public Library of Science initiative. *Information today* [online] **19** (2) [cited 14 July 2005] <http://www.infotoday.com/it/feb02/kaser.htm>

KLING, R., and McKIM, G., 2000. Not just a matter of time: field difference and the shaping of electronic media in supporting scientific communication. *Journal of the American society for information science*, **51** (14) pp.1306-1320.

LAWAL, I., 2002. Scholarly communication: the use and non-use of e-print archives for the dissemination of scientific information. *Issues in science and technology librarianship* [online] **36** [cited 17 February 2005] <http://www.istl.org/02-fall/article3.html>

LAWRENCE, S., 2001. *Online or invisible?* [online] [cited 21 July 2005] <http://ivyspring.com/steveLawrence/SteveLawrence.htm>

LYNCH, C. A., 2003. Institutional repositories: essential infrastructure for scholarship in the digital age. *ARL bimonthly report* [online] **226** [cited 20 January 2005] <http://www.arl.org/newsltr/226/ir.html>

MacCOLL, J. and PINFIELD, S., 2002. Climbing the scholarly publishing mountain with SHERPA. *Ariadne* [online] **33** [cited 10 May 2005] <http://www.ariadne.ac.uk/issue33/sherpa/intro.html>

McVEIGH, M., 2004. *Open access journals in the ISI citation databases: analysis of Impact Factors and citation patterns* [online] [cited 26 January 2005] <http://www.isinet.com/media/presentrep/essayspdf/openaccesscitations2.pdf>

MARKLAND, M. *Institutional repository questionnaire* [online] [cited 17 May 2005] <http://ico-trg.mmu.ac.uk/sherpa>

MARSHALL, C. and ROSSMAN, G., 1995. *Designing qualitative research, 2nd edition.* London: Sage.

MEYERS, B., 2004. Open access: a matter for definition. *Society of scholarly publishing, issue status report* [online] June 2004 [cited 21 July 2005] <http://www.sspnet.org/i4a/pages/index.cfm?pageid=3578>

MORRIS, S., 2004. Open access: how are publishers reacting? *Serials review* **30** (4) pp.304-307.

MORRIS, S. and OLIVIERI, R., 2004. The secret life of STM publishing. *Serials*, **17** (2) pp.111-117.

NICHOLAS, D., ROWLANDS, I., and HUNTINGTON, P., 2004. Open access publishing: what authors say. *Library and information update*, **3** (11) pp.34-35.

OAISTER [online] [cited 17 February 2005] <http://www.oaister.org>

ODLYZKO, A., 1995. Tragic loss or good riddance? The impending demise of traditional scholarly journals. *Notices of the American mathematical society* [online] **42** (1) [cited 14 July 2005] <http://www.ams.org/notices/199501/forum.pdf>

ODLYZKO, A., 1997. The economics of electronic journals. *First Monday* [online] **2** (8) [cited 14 July 2005] <http://firstmonday.org/issues/issue2_8/odlyzko/index.html>

ODLYZKO, A., 2002. The rapid evolution of scholarly communication. *Learned publishing*, **15** (1) pp.7-19.

OPEN ACCESS NOW, 2003. Financial analysts warn about impact of changes in the scientific publishing industry. *Open access now* [online] **11** [cited 17 February 2005] <http://www.biomedcentral.com/openaccess/news/?issue=11>

OPPENHEIM, A., 1992. *Questionnaire design, interviewing and attitude measurement, new edition*. London: Continuum.

OPEN ACCESS TEAM FOR SCOTLAND [online] [cited 26 July 2005] <http://scurl.ac.uk/WG/OATS/declaration.htm.

OWENS, S., 2003. Revolution or evolution? A shift to an open-access model of publishing would clearly benefit science, but who should pay? *EMBO reports* **4** (8) pp.741-743.

PELIZZARI, E., 2003. Academic staff use, perception and expectations about open access archives: a survey of social science sector at Brescia University. [online] [cited 17 February 2005] <http://eprints.rclis.org/archive/00000737/>

PHELPS, C., 1998.Achieving maximal value from digital technologies in scholarly communication. *ARL proceedings* [online] **133** [cited 28 June 2005] <http://www.arl.org/arl/proceedings/133/phelps.html>

PICKERING, B., 2004a. BioMed Central hits out at open access 'myths.' *Information world review,* **201**, p.1

PICKERING, B., 2004a. Big guns fire at open access. *Information world review,* **201**, pp.14-15.

PINFIELD, S., 2003. Open archives and UK institutions. *D-Lib magazine* [online] **9** (3) [cited 10 May 2005] <http://www.dlib.org/dlib/march03/pinfield/03pinfield.html>

POYNDER, R., 2002. Not pleading poverty: Elsevier Science chairman Derk Haank addresses industry and end-user issues. *Information today* [online] **19** (4) [cited 14 July 2005] <http://www.infotoday.com/it/apr02/poynder.htm>

POYNDER, R., 2004. Interview: put up or shut up. *Information today* [online] **21** (8) [cited 14 July 2005] <http://www.infotoday.com/IT/sep04/poynder.shtml>

PUBLIC LIBRARY OF SCIENCE [online] [cited 26 January 2005] <http://www.plos.org>

RAMACHANDRAN, R., 2004. The 'free access' debate. *Frontline* [online] **21** (2) [cited 7 July 2005] <http://www.flonnet.com/fl2102/stories/20040130000807900.htm>

REGAZZI, J., 2004. The shifting sands of open access publishing: a publisher's view. *Serials review* **30** (4) pp.275-280.

REICH, M., 2003. Peace, love and PLoS. *Physiologist* [online] **46** (4) p.137-141 [cited 7 July 2005] <http://www.the-aps.org/publications/tphys/images/tphys8x03.pdf>

RESEARCH COUNCILS UK, 2005. *RCUK position statement on access to research outputs* [online] [cited 3 August 2005] <http://www.rcuk.ac.uk/access/statement.pdf>

RICHARDSON, M., 2005. Open access and institutional repositories: an evidence-based approach. *Serials* [online] **18** (2) pp.98-103 [cited 10 August 2005] <http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=18&issue=2&page=98>

ROBSON, C., 2002. *Real world research:* a resource for social scientists and practitioner-researchers*, 2nd edition*. Oxford: Blackwell.

ROMEO [online] [cited 17 February 2005] <http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html>

ROWLANDS, I., NICHOLAS, D., and HUNTINGDON, P., 2004. *Scholarly communication in the digital environment: what do authors want?* [online] [cited 24 January 2005] <http://ciber.soi.city.ac.uk/ciber-pa-report.pdf>

SHERPA [online] [cited 17 February 2005] <http://www.sherpa.ac.uk>

SMITH, M. et al., 2003. DSpace: an open source dynamic digital repository. *D-Lib magazine* [online] **9** (1) [cited 20 January 2005] <http://www.dlib.org/dlib/january03/smith/01smith.html>

SUBER, P., 2002. Where does the free online scholarship movement stand today? *Cortex*, **38** (2) pp.261-264.

SUBER, P., 2003. *Removing the barriers to research: an introduction to open access for librarians* [online] [cited 11 May 2005] <http://www.earlham.edu/~peters/writing/acrl.htm>

SWAN, A. and BROWN, S., 2003. Authors and electronic publishing: what authors want from the new technology. *Learned publishing,* **16** (1) pp.28-33.

SWAN, A. and BROWN, S., 2004. Authors and open access publishing. *Learned publishing,* **17** (3) pp.219-224.

TALJA, S. and MAULA, H., 2003. Reasons for the use and non-use of electronic journals and databases: a domain analytic study in four disciplines. *Journal of documentation*, **59** (6) pp.673-691.

TWYMAN, N., 2004. Launching PLoS Biology: six months in the open. *Serials*, **17** (2) pp.127-131.

VAN DE SOMPEL, H. et al., 2004. Rethinking scholarly communication: building the system that scholars deserve. *D-Lib magazine* [online] **10** (9) [cited 24 January 2005] <http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html

VELTEROP, J., 2003a. Public funding, public knowledge, publication. *Serials*, **16** (2) pp.169-174.

VELTEROP, J., 2003b. Should scholarly societies embrace open access (or is it the kiss of death)? *Learned publishing*, **16** (3) pp.167-169.

WARE, M., 2004. Universities' own electronic repositories yet to impact on Open Access. *Nature webdebates* [online] [cited 7 July 2005] <http://www.nature.com/nature/focus/accessdebate/4.html>

WATSON, L., LOGIN, S., and BURNS, J., 2003. Exploring new ways of publishing: a library-faculty partnership. *Journal of the medical library association*, **91** (2) pp.245-247.

WELLCOME TRUST, 2003. *An economic analysis of scientific research publishing, revised edition* [online] [cited 24 January 2005] <http://www.wellcome.ac.uk/assets/wtd003182.pdf>

APPENDIX 1

Copy of questionnaire sent by email to humanities academics.

An online copy of the survey can be found here:
<http://survey.perseus.com/287fb63a.htm>

This survey forms part of my research for an MA in Library and Information Management at Manchester Metropolitan University.

The aim of the project is to gather the perceptions of academics about the open access movement, and in particular their attitudes towards depositing their work in institutional repositories. I would greatly appreciate your views on these subjects.

This survey is aimed at academics from humanities disciplines. All information provided via this survey will be kept confidential and any data used in the final report will be anonymous.

If you have any questions or comments about the study or this survey, please contact me at james.d.allen@student.mmu.ac.uk.

Please could you complete this questionnaire by July 1st 2005. Thank you for your help.

ONE.
Which of the following names, services or terms are you aware of ?
Please tick as many as apply.

☐ Budapest Open Access Initiative (BOAI)

☐ Self-archiving

☐ Springer Open Choice

☐ Oxford Open

☐ BioMed Central

☐ www.freemedicaljournals.com

☐ SHERPA project

☐ The serials crisis

☐ The arXiv / Los Alamos archive

☐ Highwire Press

☐ E-Prints software

☐ Public Library of Science

☐ Directory of Open Access Journals (DOAJ)

- FAIR (Focus on Access to Institutional Repositories)
- E-Prints UK project
- E-LIS archive
- PubMed Central
- Open Archives Initiative (OAI-PMH)
- Author-pays publishing
- www.OAIster.org
- PhilSci Archive
- The open access movement
- Institutional repositories
- D-Space software
- The open access movement
- Institutional repositories
- Budapest Open Access Initiative (BOAI)
- Open Archives Initiative (OAI-PMH)
- Self-archiving
- PhilSci Archive
- Author-pays publishing
- E-LIS archive
- The serials crisis
- E-Prints software
- Springer Open Choice
- D-Space software
- Oxford Open
- www.OAIster.org
- Directory of Open Access Journals (DOAJ)
- SHERPA project
- PubMed Central
- FAIR (Focus on Access to Institutional Repositories)
- BioMed Central
- The arXiv / Los Alamos archive
- Public Library of Science
- E-Prints UK project
- Highwire Press

☐ www.freemedicaljournals.com

Which of these best explains how you *first* knew of the terms above.

Please only choose one answer.

○ I discovered it myself whilst on the Internet

○ I heard a debate about it in my own discipline

○ I heard a debate about it in another discipline

○ I heard a debate about it in my University

○ I heard about it from a colleague in my own discipline

○ I heard about it from a colleague from another discipline

Does your university / institution have an Institutional Repository or electronic archive?

> *An Institutional Repository can be defined as a digital archive, managed and organised by a University, or small group of Universities, in which its researchers can deposit electronic copies of their work, subject to copyright agreement.*

○ Yes

○ No [Goto question FIVE]

○ Don't know [Goto question FIVE]

Have you deposited any of you work in it?

○ Yes [Goto question SIX]

○ No [Goto question NINE]

If your university / institution set up an Institutional Repository, would you deposit your work in it?

○ Yes

○ No [Goto question NINE]

○ Don't know [Goto question NINE]

What type of material have you / would you deposit in an Institutional Repository?

- [ ] peer-reviewed articles published in a journal
- [ ] non peer-reviewed articles published in a journal
- [ ] articles awaiting peer-review
- [ ] conference papers
- [ ] theses
- [ ] datasets
- [ ] teaching materials
- [ ] other (eg. photos, video, audio etc.)

SEVEN.

Which of these statements about depositing this work in the Institutional Repository do you agree with?

- [ ] My work is published alongside other high quality research
- [ ] My chances of promotion are increased
- [ ] The accessibility of my work is increased
- [ ] I retain the copyright of my work
- [ ] I can add extra data to the work, such as photos, video, audio or datasets
- [ ] Access to the work is cheaper for others
- [ ] The University saves money
- [ ] The number of citations my work gets increases
- [ ] The impact of my research is increased
- [ ] My work is disseminated more quickly
- [ ] The repository is easy to use
- [ ] My work will be permanently archived and available
- [ ] The public opinion of the University is increased
- [ ] My published material is easy to find
- [ ] The serials crisis is redressed
- [ ] The repository is well indexed and archived
- [ ] Depositing my work in the repository protects it from plagiarism
- [ ] The accessibility of my work is increased
- [ ] The impact of my research is increased
- [ ] The number of citations my work gets increases
- [ ] My work is disseminated more quickly
- [ ] The repository is easy to use

- ☐ My work will be permanently archived and available
- ☐ My published material is easy to find
- ☐ The repository is well indexed and archived
- ☐ Access to the work is cheaper for others
- ☐ I retain the copyright of my work
- ☐ Depositing my work in the repository protects it from plagiarism
- ☐ I can add extra data to the work, such as photos, video, audio or datasets
- ☐ My work is published alongside other high quality research
- ☐ The serials crisis is redressed
- ☐ The University saves money
- ☐ The public opinion of the University is increased
- ☐ My chances of promotion are increased

EIGHT.
What disadvantages or difficulties do / might you encounter from depositing your work in this way?

```
┌─────────────────────────────────┐▲
│                                 │▓
│                                 │
│                                 │
│                                 │
│                                 │▼
└─────────────────────────────────┘
```

Please now go to question ELEVEN.

NINE.
Which of these statements about depositing your work in an Institutional Repository do you agree with?

- ☐ It would be difficult and time-consuming to deposit my work
- ☐ Publishers would not let me put my work in a repository
- ☐ If I deposited work in a repository then I could not later publish it in a peer-reviewed journal, and so my promotion prospects would suffer
- ☐ My work might be misused or plagiarised
- ☐ I would break copyright agreements by making my work available in a repository
- ☐ The Research Assessment Exercise would not take into account work in a repository, so the department would suffer

☐ The impact of my work would be less if I deposited work in a repository, and so my promotion prospects would suffer

☐ With no peer review or quality control process, the quality of content of the repository would be questionable

☐ It would not be easy to find my work

☐ Publishers would not let me put my work in a repository

☐ It would not be easy to find my work

☐ My work might be misused or plagiarised

☐ With no peer review or quality control process, the quality of content of the repository would be questionable

☐ It would be difficult and time-consuming to deposit my work

☐ If I deposited work in a repository then I could not later publish it in a peer-reviewed journal, and so my promotion prospects would suffer

☐ The impact of my work would be less if I deposited work in a repository, and so my promotion prospects would suffer

☐ The Research Assessment Exercise would not take into account work in a repository, so the department would suffer

☐ I would break copyright agreements by making my work available in a repository

TEN.
What advantages might you encounter from depositing your work in this way?

ELEVEN.
Do you deposit any of these types of work on any part of your university's website (eg. department, research group or personal page)?

	Yes	No
peer-reviewed articles published in a journal	☐	☐
non peer-reviewed articles published in a journal	☐	☐
articles awaiting peer-review	☐	☐
conference papers	☐	☐
theses	☐	☐
datasets	☐	☐

	Yes	No
teaching materials	☐	☐
other (eg. photos, video, audio etc.)	☐	☐

TWELVE.
Do / would you ever publish work in an electronic journal which produced *no printed copy*?

☐ Yes

☐ No

THIRTEEN.
Do / would you ever publish work in an electronic journal which produced no printed copy, and was *free to access*?

☐ Yes

☐ No

FOURTEEN.
Do / would you ever publish work in an electronic journal which produced no printed copy, was free to access and was *not peer-reviewed*?

☐ Yes

☐ No

FIFTEEN.
Who is the intended audience for the research papers you publish?

☐ a small group of researchers in your particular field

☐ all researchers in your discipline

☐ your research funder

☐ the whole academic community

☐ the general public and the academic community

SIXTEEN.
What University / institution are you from?

- [   ]

SEVENTEEN.
What is your position in the University?

☐ PhD student / Post-doctoral researcher

☐ Lecturer

☐ Senior Lecturer

☐ Reader

☐ Professor

EIGHTEEN.

Do you have any other comments about publishing, open access or institutional repositories that you would like to make? Please enter them below.

```
[                              ]
```

NINETEEN.

If you are willing to let me contact you in the next few weeks to discuss your responses, please enter your email address or telephone number below.

```
[                              ]
```

Thank you for completing this survey.

[Submit Survey]

APPENDIX 2

List of twenty-five UK institutional repositories surveyed, and their web addresses.
The survey took place on 24-25th June 2005.

Birkbeck, University of London   \<http://eprints.bbk.ac.uk\>

Birmingham University   \<http://eprints.bham.ac.uk\>

Bristol University   \<http://rose.bris.ac.uk/dspace\>

Cambridge University   \<http://www.dspace.cam.ac.uk\>

Cranfield University   \<http://dspace.lib.cranfield.ac.uk/index.jsp\>

Durham University   \<http://eprints.dur.ac.uk\>

Edinburgh Univerity   \<http://www.era.lib.ed.ac.uk/index.jsp\>

Glasgow University - DSpace   \<https://dspace.gla.ac.uk/index.jsp\>

Glasgow University - Eprints   \<http://eprints.lib.gla.ac.uk\>

Imperial College London   \<http://eprints.imperial.ac.uk\>

King's College London   \<http://eprints.kcl.ac.uk\>

London School of Economics   \<http://eprints.lse.ac.uk\>

Newcastle University   \<http://rogue.ncl.ac.uk\>

Nottingham University   \<http://eprints.nottingham.ac.uk\>

(Nottingham University - Modern Languages archive

       \<http://mlpa.nottingham.ac.uk\>)

Open University   \<http://libeprints.open.ac.uk\>

Oxford University   \<http://eprints.ouls.ox.ac.uk\>

(Oxford University - Maths archive   \<http://eprints.maths.ox.ac.uk\>)

Royal Holloway, University of London   \<http://eprints.rhul.ac.uk\>

St Andrew's University   \<http://eprints.st-andrews.ac.uk

School of Oriental and African Studies   \<http://eprints.soas.ac.uk\>

Southampton University   \<http://eprints.soton.ac.uk\>

Stirling University   \<http://dspace.stir.ac.uk/dspace/index.jsp\>

Surrey University   \<http://epubs.surrey.ac.uk\>

University College London   \<http://eprints.ucl.ac.uk\>

Warwick University   \<http://www2.warwick.ac.uk/fac/sci/csc/eprints\>

White Rose repository   \<http://eprints.whiterose.ac.uk\>