

OPEN ACCESS – WHAT SHAPE THE FUTURE?

January 2009

Executive Summary

Open access is a business model which challenges the traditional way scholarly communication has operated – it relies on alternative funding mechanisms to underpin the costs of electronic publishing and moves away from the restrictions imposed by the subscription and site licensing processes. It is part of a wider movement within ICT whereby free access to all forms of information available through the Internet is becoming the norm. However, there are a number of forms which open access can take, some short term (the Green or Grey routes) and some more long term (the Gold route), but in these differences there lie tensions between the respective advocates of the movements.

Whilst it is felt by publishers in particular that the real challenge created by open access is the effect it will have on the quality of the scholarly record, and in the commercial viability of the industry per se, the real significance of open access is the impact it will have in assisting the migration away from a traditional 'article dominance' to a more multimedia, multifunctional approach to information dissemination. Open access, when seen in the context of the growth of (free access) to data, and (free participation) is social networking and social collaboration, and when considered alongside the growing power and digital content of free-to-access search engines such as Google, the real challenges of open access become apparent.

For ICSTI members the strategic significance of the open access movement is the change in the business paradigm which emerges with new players, new business models and new attitudes to the creation and use of scholarly information. Open access is as much a part of a social evolution of the scholarly communication process as it is a technical issue – whilst many of the technical infrastructural issues have or are being addressed the real question is how quickly the millions of authors and readers worldwide will adapt to the new open access opportunities. How quickly will the restrictive barriers, necessary to protect the quality of content, be dismantled to allow for a wider and more open access distribution? The answer to this will dictate how rapidly ICSTI members will need to assess the significance of their corporate

missions and embrace alternative strategic positioning in the new information world.

Definition of Open Access (OA)

'Open Access' within the context of scholarly communications is a set of business models that enables publicly funded research results to be disseminated to the widest possible audience at no cost for use of the published information.

It challenges the traditional journal subscription model and its online site licensing derivatives in the restrictions these place on access by any given individual or institution. The restriction is based on an institution's ability or willingness to pay for a subscription to a journal or series which in all likelihood includes more information than may be required. The premise of open access is – 'take away these commercially-based restrictions and potentially a much wider audience within society can be reached and a broader community therefore stands to benefit'.

As is claimed in a European Commission report on open access, 'Researchers, interested lay people, inventors, patients, teachers and journalists are all confronted with a tollbooth at the entrance to the gardens of knowledge when seeking the latest information for private or professional purposes'. These tollbooths are in many cases insurmountable for individuals. They become part of the vast current "disenfranchised" in the scholarly community, a community which open access claims to address.

Foundations of open Access

Although open access is seen to be a creature of the Internet, the genesis of the open access movement could be said to lie further back, within the specific culture created by such as the artificial intelligence lab at MIT in the 1970s, rather than the Internet in the 1990's. At that earlier time software programmes were still generally written as part of large-scale centralised projects, and distributed on floppy disks or tapes. Nowadays the Internet has emerged as the information distribution medium.

During the late 1990's a number of vocal advocates used international platforms to promulgate their own ideas of 'openness'. These proposals were based on the iniquity of the existing situation – that research was being funded by the public purse and undertaken by researchers whose career progression was heavily dependent on a research publication. Yet the written results of the research became 'owned' by publishers who not only set high

access prices in some cases but also applied these to the very same institutions that nurtured the research.

It also became evident that the price rises being set by publishers on scholarly and research works far exceeded other social price indices, notably the institutional/library collection budgets. Libraries, which were seeing only marginal increases in their collections budgets at best, led campaigns against such alleged 'price gouging'.

As implied earlier, there is also an inbuilt inefficiency in the subscription model as libraries and institutions pay for more information than they need – it comes in bigger bundles (particularly in the so-called 'Big Deals') than libraries or their customers are able to digest.

Finally, there was evidence reported by publishers such as Elsevier of the 30-40% gross profit margins earned on their scholarly publishing operations. This made libraries and policy makers believe that they were paying both for the research and again for the resultant publications, whilst the publishers were making a handsome, monopolistic profit. The frequent cry was that the existing system was 'dysfunctional'.

But the key claim was that there was a restriction to publicly funded research experiments, and it was felt that society's welfare would be enhanced, its productivity and efficiency increased, if the world's research findings were to be made available for all to access and use as and when they wanted. All this was based on supposition – until recently there has been little hard evidence to substantiate such claims.

Brief historical perspective

An early pioneer in addressing this issue in a novel business way was Vitek Tracz, a former publisher who had built up a formidable war chest based on creating and selling traditional publishing companies. In May 2000 he turned from being poacher to gamekeeper by establishing BioMed Central. BMC is an online journal publisher in the biomedical area which relies on payments received from authors to provide open access to quality (refereed) journal material. He went against conventional practice of depending on subscription sales, and in so doing raised a potential threat to traditional publishing practice. Initially each author submitting an article for publication within the BMC stable of online journals was charged \$450; in more recent years this rate has been increased. But the main point was that once this payment had been made by the supplier of the manuscript (to pay for all publishing

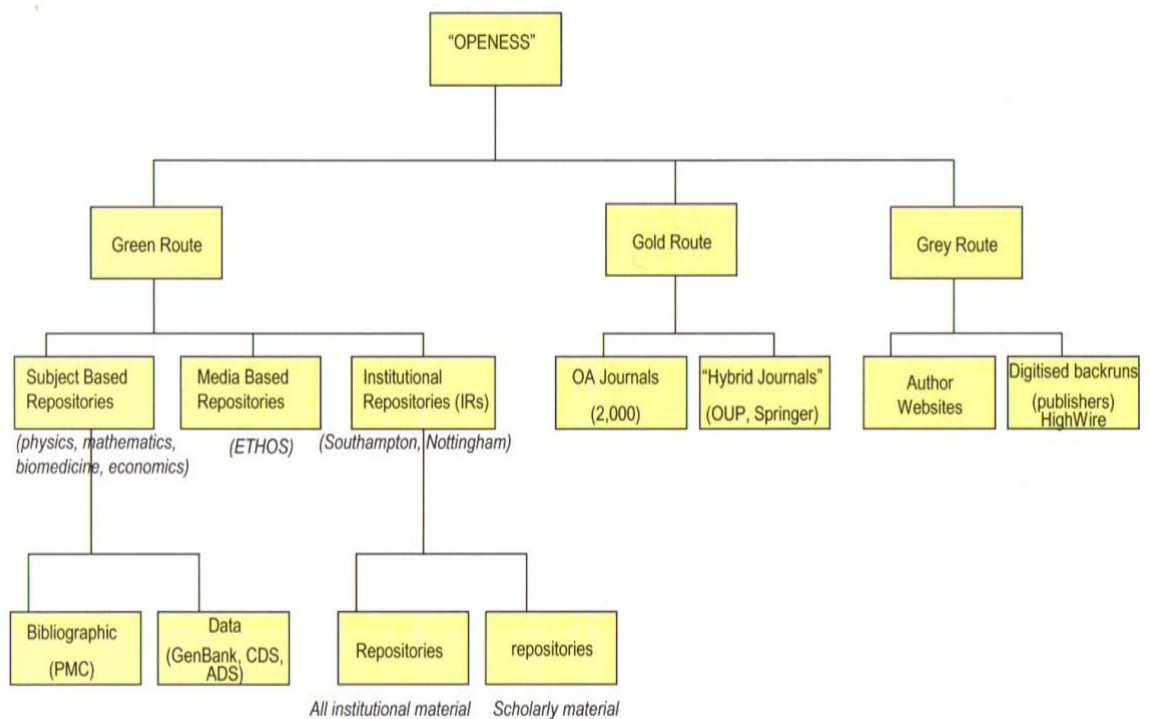
services) anyone, anywhere, could access that work online. There would be no restrictions to access. The tollbooths were removed.

On December 1-2, 2001, the Open Society Institute (OSI) supported by the Soros Foundation convened a meeting in Budapest, Hungary, attended by leading proponents of the then disparate open access movements. The goal was to see how far the many current initiatives could assist one another and how OSI could use its resources to help the cause. Once broad agreement was reached the Budapest Open Access Initiative was endorsed by over 300 institutions and 3,600 individuals. It was the first international platform to propose this new way of open scholarly communication. The Budapest Initiative was followed by similar declarations from meetings held in Bethesda, Maryland, and then in Berlin.

Another significant event was the circulation of an 'open' letter originating in the USA in early 2000 and signed by 34,000 scientists in support of action by librarians and authors to boycott commercial publishers because of their pricing and access restriction policies. This particular suggestion disappeared in the mists of time. Instead it spawned Public Library of Science (PLoS), another new open access journal publisher, which like BMC eschewed conventional business practices.

Since then various national research agencies threw their support behind open access and the topic has been kept in the forefront of librarians, policy makers and funders through the activities of a few notable individuals such as professors Jean-Claude Guedon (Montreal), Stevan Harnad (Southampton) and Peter Suber, acting on behalf of the SPARC library consortium. These individuals have become idols or villains depending on one's view of the open access movement. But despite their common belief in 'open access', the way this was to be achieved varied according to the individual concerned – Harnad is a vocal advocate of one form of open access, and is often in conflict with J-C Guedon who believes open access can be achieved in a different way. So we are not faced with a united Church as far as open access is concerned – in fact the following diagram shows the disparity between the different forms which OA can take.

Types of Open Access Material



The two main formal types of open access are the so-called Gold Route and the Green Route. However, not to be overlooked is the less institutionalised Grey Route which includes author's depositing their reports and supplementary material on their own personal web sites.

Business Models for Open Access

The essence of the Golden open access is a business model which has been turned on its head as far as traditional operations in the scholarly information market are concerned. It moves the money away from an institutional purchase of information, mainly by academic and research libraries, to a different supply-sponsoring system. This can either be based on income received from authors, advertising related, subsidised by an institution,

sponsored by an interest group or charity, or maintained voluntarily by a group or individual. The central feature is that by making the content itself 'free' it reaches a much wider market – so it is claimed - than if it is controlled by a subscription-payment system.

The basis for the Green open access movement is that there is no business model in the conventional sense. Everything is done voluntarily – authors volunteer to deposit their articles in repositories which are maintained by institutions or the community without creating rechargeable costs.

The means to achieve the open access end depends on whether a short-term approach is adopted (primarily the Green Route) or a more strategic longer-term view is upheld (the Gold Route to open access). This distinction has become critical.

A. Open Access Journals – the Gold Route to open access

During the early years of this millennium the advocacy and public relations exercises in the way scholarly publishing should be conducted was being won by the Gold open access movement. Some articulate advocates highlighted how inequitable and restrictive was the traditional 'toll-based access' (TA) publishing system. However, nothing is free, particularly the creation and support of a quality-based refereeing system which sifts the wheat from the chaff in terms of knowledge generated. This refereeing system is something that provides benefits to authors and the community. Authors can bathe in the credibility conferred by having their article included within the 'brand' created by publishers for a particular journal title or book series, one that had become well known and respected in their area. Therefore it was suggested that, since authors stand to benefit most from the publication/refereeing system, they should pay for the costs of publication.

Underlying this essential emotional appeal was an attempt to quantify the benefits of an open access scholarly society. The most sophisticated of these came from an Australian, John Houghton.

Economic theory sustaining open access

Professor John Houghton, professional fellow at the Centre for Strategic Economic Studies at Victoria University, Melbourne, Australia, has produced reports that analyse macro-economic factors and the impact which open access could have on a nation's economy. In reports such as "The Economic Impact of Enhanced Access to Research Findings" (CSES Working Paper 23, July 2006) he freely admits that the data is not yet robust enough to be conclusive. Nevertheless his general impression was that open access will

produce a net improvement in society's wellbeing. He marshalled some figures that point in this direction.

His argument was based on a number of examples which suggested that existing government funded R&D produces both a private and a social rate of return of around 25-50%. By applying growth models which incorporate 'access' and 'efficiency' into the equation, substantial additional net benefits to national economies can be achieved. He speculated that the improved access and efficiency factors could be attributable to the substitution of a closed subscription by an open access business model. Using statistical equations, models and formulae, Houghton attempts to convince readers that the benefits that could be obtained by the change to open access could be substantial for all countries, and particularly:

- In Germany, its \$58.7 billion investment in government funded R&D could have been improved to the extent of \$3 billion in 2003 from a 5% increase in access and efficiency.
- In the United Kingdom, the government's \$33.7 billion investment in R&D could have been increased by \$1.7 billion from a 5% increase in access and efficiency.
- In the United States, the government's \$312.5 billion could have been improved by \$16 billion as a result of a 5% increase in access and efficiency.

Houghton's treatise focuses on the way a social return of 50% on R&D in general, and a further 5% increase in access and efficiency as a result of open access was arrived at.

However, just released is a follow-up study focused on the UK higher education sector alone which JISC has commissioned from John Houghton working together with Professor Charles Oppenheim from the University of Loughborough. It appears that this latest report takes a different approach and incorporates the experiences of publishing experts in assessing the economic value and returns from open access in the UK. However, the details of the procedures followed have only just been disclosed and still need analysing. In summary, Houghton et al suggest that core scholarly publishing system activities cost the UK higher education sector around £5 billion in 2007. Using the different models, the report shows, what the estimated cost would have been:

- £230 million to publish using the subscription model,
- £150 million to publish under the open access model and

- £110 million to publish with the self-archiving with peer review services plus some £20 million in operating costs if using the different models.

When considering costs per journal article, Houghton believes that the UK higher education sector could have saved around £80 million a year by shifting from toll access to open access publishing. He also claims that £115 million could be saved by moving from toll access to open access self-archiving.

In addition, the financial return to UK plc from greater accessibility to research might result in an additional £172 million per annum worth of benefits from government and higher education sector research alone.

The public launch of the Houghton report on the UK situation is scheduled for 26th March 2009. It is likely to stir up a hornet's nest as traditional scholarly publishers focus in on the models and data used to substantiate the economic case for open access in general.

Business models for Gold open access

Moving away from the macro to the micro, the cost estimates for Gold open access publications vary. The first commercial venture to offer OA journals – BioMed Central – came up with a \$450 per article processing fee. However, this figure was a bit vague as it depended on unknown submission rates (and rejection rates, which also incur costs but from which no revenues would flow). This is also true of traditional subscription based journal publishing – there is no consistency or agreement on a template for comparing journal publication costs, irrespective of the methodology adopted to generate revenues.

It is only early in 2008, in a report completed for the Research Information Network (RIN), that a better handle on all cost elements in electronic journal publishing was derived by Cambridge Economic Policy Associates (CEPA). Their findings, which will be commented on later, indicated that the cost of processing an article was \$2,300 (or £4,000 taking 'free' refereeing into account) – see "Activities, costs and funding flows in the scholarly communication systems", May 2008. Also, the Wellcome Trust, which strongly favours open access, supports a price level of between \$2,750 and \$3,500 per article as reflected in their own costing exercise done on open access publishing. Gradually 'gold' open access journal publishers are moving up to a level more in line with the CEPA and Wellcome Trust figures.

These costs set a benchmark for Gold open access journal publishers to achieve if they are to become sustainable and commercially viable in the long term. The main source of revenues to pay for the costs publishing services is seen to be payments from authors (or their institutions or funding councils). However, any price set on an author for the privilege of submitting a manuscript (which could be rejected) is contrary to what authors are used to. Why should they pay for something which has been – for them as authors – essentially a free service offered by established publishers? Besides the cost issue, without the ‘brand’ of an established journal or publisher, the author would not recognise a new open access journal as having the legitimacy to confer status and recognition among the relevant peer group. As such the build-up of interest and manuscripts from the author community has been slow.

A different approach has been adopted by the not-for-profit organisation based in the US, the Public Library of Science. As indicated earlier, it started life as a lobby group against commercial publishers, led by Nobel Laureate Dr Harold Varmus who had been director of National Library of Medicine (and is now scientific advisor to the incoming Obama administration) along with other notable experts such as Patrick Brown (Stanford) and Michael Eisen (Berkeley). Varmus and his colleagues created a prestigious online journal in biology (PLoS) with the help of a \$9 million grant from the Gordon and Betty Moore foundation. A title in medicine has followed, as have several more recent niche-type OA journals from PLoS, to reinforce the organisation’s commitment to excellence and the idea of open access journal publishing becoming, ultimately, a self-sustaining business model. The PLoS publication charge per article is \$1,500 to reflect the extra costs associated with rejecting a high proportion of submitted works (in comparison with BMC). But even these prices are considered insufficient by the professionals.

Traditional publishers took time to respond to these initiatives. They focused on the impracticability of the central theme of the OA movement – would authors want to pay for publication services? Would authors give up a slice of their research budget to enable their article to be produced through an expensive refereeing system? What about authors in the third world who do not have discretionary funding in this area? Or impoverished authors generally? Was it creating elitism in the publication process?

OA publishers tried to get round this by offering waivers for those unable to pay the page charges. They also sought ‘subscriptions’ from institutions generally which would enable any author from that institution to have a discounted page rate, or even no rate charge at all.

A survey undertaken in 2006 by ALPSP looked at the proportion of OA journals which actually charged author fees. It emerged that only 48% of the journals charged an author-side fee. The rest were funded from a variety of sources including charitable subventions, grants, etc. Most of the latter could hardly be considered sustainable in the long run. (See: http://www.alpsp.org/ngen_public/article.asp?id=200&did=47&aid=270&st=&aid=-1)

Current impact of Gold open access

Eight years since the launch of the first commercial open access business model we are not seeing the sustainability which such ventures need. BioMed Central's stable of some 150 Gold open access journals has been recently sold by its owner (Vitek Tracz) to one of the largest TA or subscription based journal publishing companies – Springer Science and Business Media. Springer itself is owned by Candovar and Cinven, two London-based venture capital companies whose rules on business are distinctly non-charitable.

PLoS has been for some time struggling to make a surplus and their publishing operations in the USA and UK are only sustainable as they draw on the \$9 million funds from the Moore Foundation (which has of necessity been increased). A third major open access publisher is Hindawi. Their stable of some 60 journals relies on the pool of inexpensive but highly-qualified scientists in Hindawi's home country of Egypt.

More worrying have been revelations that some of the newer names to the open access publishing scene have been using tactics to gain visibility and manuscripts from authors which have caused great concern. Richard Poynder, an independent researcher, has investigated two such companies during the past year. The tactics employed by Bentham and SJI and have tended to debase the Gold open access coinage. It raises issues about quality control in an open access environment.

As far as authors are concerned, the costs of publishing their paper in a branded journal of repute were borne by another part of the information chain – an intermediary, the library budget. As such the costs of quality control and administering the refereeing system were invisible – all they were asked to do was to give up their copyright over the work (in certain instances) in return for having their work professionally disseminated throughout the world. On the other hand open access brings them face-to-face with the reality of costs. They or their institutions will be expected to make a financial contribution for the costs of bringing the article to market. This is something new for them. In some instances it will involve financial pain.

Hybrid Journals

Nevertheless, several commercial and learned societies have felt compelled, in the face of the growing public campaign in favour of open access, to experiment with open access publications. Hybrid journals have emerged, which includes a mixture of author paid (OA) and toll-based article access (TA). Oxford University Press and Springer S+BM have been leading lights in this latter process, though many others have also joined – perhaps half-heartedly – the bandwagon.

In these cases the subscription price to the hybrid journal in the year following publication will reflect the proportion of income received by the publisher from author-payments in the previous subscription year and this would reduce the following subscription price accordingly. Though it remains unclear who audits this.

Gold OA Journals – What shape the Future?

Open Access (OA) journals have not been launched with the same energy and commitment with which its proponents had hoped. The Directory of Open Access Journals (DOAJ) lists some 3,800 titles, growing at a rate of almost one title per calendar day, and includes some 250,000 articles. See <http://www.doaj.org>

Today open access in its most popular forms (gold and green) accounts for about 2% of all newly published scholarly articles (of which there are 1.2 to 1.4 million per annum (Elsevier/Mabe)). The author-pays model accounts for less than 1% of this overall total.

It is expected that the Golden route to open access will continue to grow but not to the extent that it will result in a revolution in the scholarly communication process. It will sit alongside the traditional TA system, edging an increasing market share as overall society and technology led features come into play in support of author-paid open access publishing. It has to be recognised that the culture of the Internet, with its assumption of free access to material, will have a considerable effect on the move towards open access. Users being able to download articles which they want without subscription restrictions is becoming a dominant feature of the web. Golden open access would be a long term beneficiary.

However, for the next five years at least one can expect only small increments in Gold OA publications as the author community struggles to adapt to a more

transparent publication system in terms of costs. As far as ICSTI members are concerned the Gold open access movement requires at best a slight adaptation to prevailing business models, at worst just a monitoring of developments as they slowly emerge.

B. Author self-depositing articles – the Green Route to open access

This is a totally different way of implementing the open access agenda and as has been pointed out there is occasional conflict between the two schools. Whilst the leading supporters of the Gold and Green routes are united in the ultimate aim of seeing a free-to-access system for research information, the means to achieve that end differ. And in that difference lie arguments, disagreements, conflict and confusion.

Subject-based E-Print services

‘Green’ open access parentage goes back even earlier than Gold, to 1991, when a physicist, Dr Paul Ginsparg, invited colleagues in the self-selected high energy physics community to deposit their electronic pre-publication manuscripts on his personal server at the Los Alamos National Laboratory (LANL) and allow all other physicists, worldwide, to interrogate this so-called arXiv service for free. There was limited editorial control; that was exercised by the professionals themselves who chose to use ArXiv for their information updates. The compact nature of the subject area, the strong group control and the traditional culture of sharing printed preprints in this specialised physics area, meant that the arXiv service was widely adopted by this community.

There was no payment transactions involved – supply and demand for the articles was free. It was a mechanism created and sustained by the community, and no professional publishing skills were involved.

During the following decade the service outgrew the resources of Ginsparg at LANL. As it was seen as an important national and international resource it received financial support from the National Science Foundation (NSF). It also eventually outgrew Los Alamos’ ability to support it and arXiv has since moved to Cornell University. The good news for the traditional publishers was that though the arXiv content eventually became published in the refereed journals, subscription loss was no less or greater than journals in comparable disciplines where no such subject-based institutional repository existed. There was an apparent complementarity in function between such an informal and formal publication methods. Or so it was claimed.

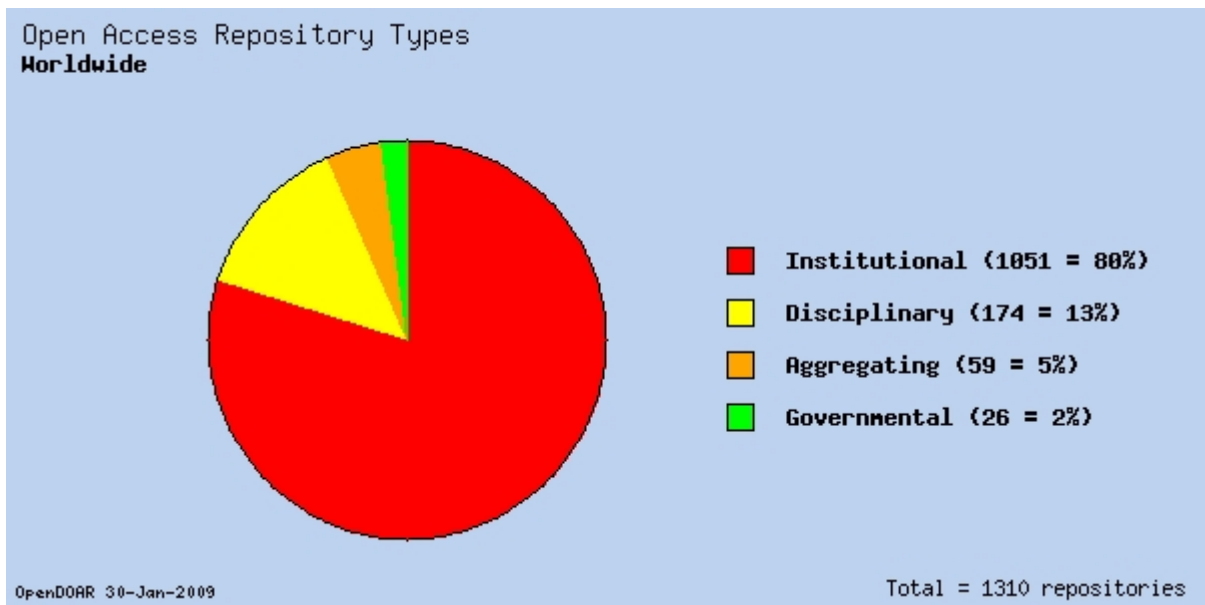
It also appeared that only few disciplines had the compactness that favoured a non-refereeing institutional-based approach. The number of discrete physics sub-subjects in arXiv increased, and mathematics took on a similar approach. Economics and cognitive disciplines also had strong advocates who created subject-based disciplines out of nothing, but the larger STM communities remained true to the primary research journal publishing model. The strength of the refereeing system encapsulated within the primary journal was seen as inviolate.

Institutional-based repositories

Much later, the idea of each research institution having its own server with a collection of research material (in whatever format) produced by researchers from within the institution became popular. It meant that each institution that had nurtured the research, in some instances co-financed the research, would be able to bathe in the glory as the results became accessible through them on the worldwide stage. The institution would become 'the publisher' – it gave a new meaning to university presses in everything but name. It enabled the institution to be seen as an eminent research centre in its own right without seeing the results distributed among the journal brands of publishers whose role in the research effort was seen as minimal.

The split between institutional and subject based repositories can be seen from the following graph:

Open Access Repository Types – Worldwide



Source: OpenDOAR, University of Nottingham

The above shows how institutional repositories have increased in numbers. It appears IRs struck a chord and there is much attention being given by funding agencies such as JISC in the UK to establish a national infrastructure to enable all higher education institutes to have their own individual institutional repository (IR). Other countries are dealing with this in a less centralised way, but even in these countries there is a groundswell of support at the local level to create such servers and having them easily accessible.

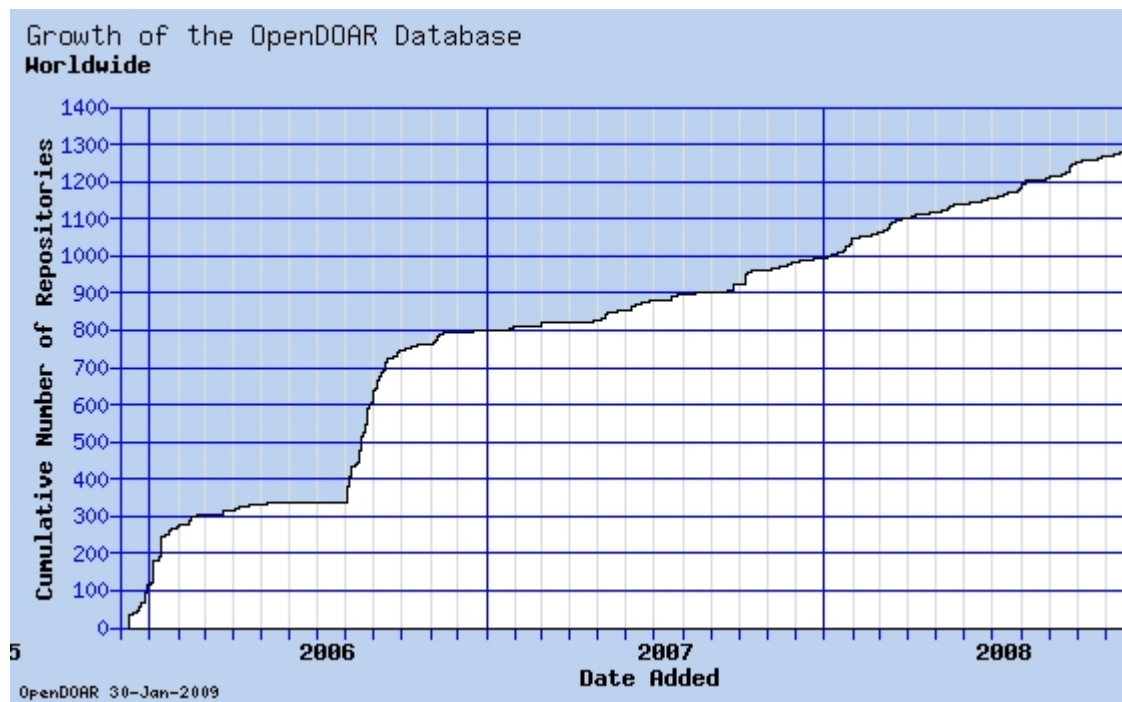
OpenDOAR

OpenDOAR is a project developed and maintained by the [University of Nottingham](#) in the UK and funded by JISC, SPARC Europe, OSI and CURL to provide quality-assured listing of open access repositories around the world. OpenDOAR assigns metadata to allow categorisation and analysis to assist the wider use and exploitation of repositories. Users of the service are able to analyse repositories by location, type, the material they hold and other measures. This key point about OpenDOAR is that this information is of use not only to users wishing to find original research papers but also for third-party service providers, such as search engines or alerting services, who need easy-to-use tools for developing tailored search services to suit specific user communities.

SHERPA currently runs a suite of services for the Open Access community. SHERPA maintains the [RoMEO](#) service, which gives summaries of the archiving rights that different publishers allow authors to retain. To complement this, SHERPA is also developing the [JULIET](#) service, which summarises the archiving responsibilities and requirements that funding agencies give as a condition of funding grants. OpenDOAR is the third part of this repository service, listing available open access repositories.

Given the ability to identify, sort and locate different repositories it is expected that new services and uses will develop. One example of this is the development of overlay journals; such emergent capabilities will be facilitated by the use of a comprehensive, structured and maintained list. From a very low base it does look as if the growth in IRs has taken off. Whilst the “tipping point” may not have been fully reached as yet, the following chart does show that there has been growth in the author archives of material on their local IRs.

Growth of the OpenDOAR Database - Worldwide

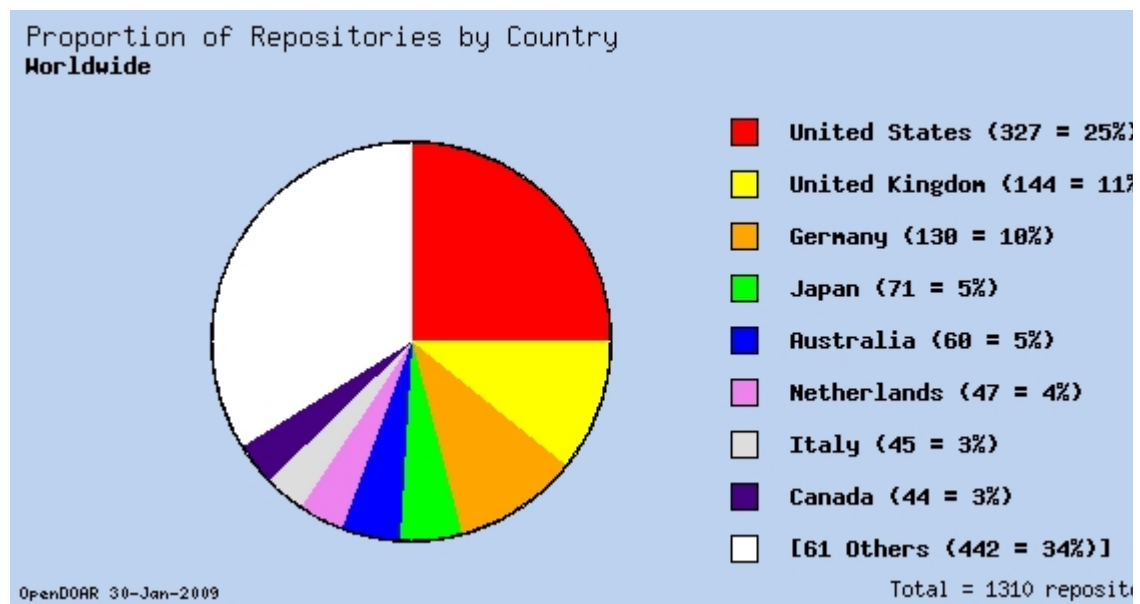


The shape of the chart in 2006 reflects the work of *OpenDOAR* rather than the

growth of the number of actual repositories. A backlog of new records built up while the database was being redeveloped during mid-2006, and clearing this backlog created the step in the graph. The chart better represents the true growth in the number of repositories from 2007 onwards.

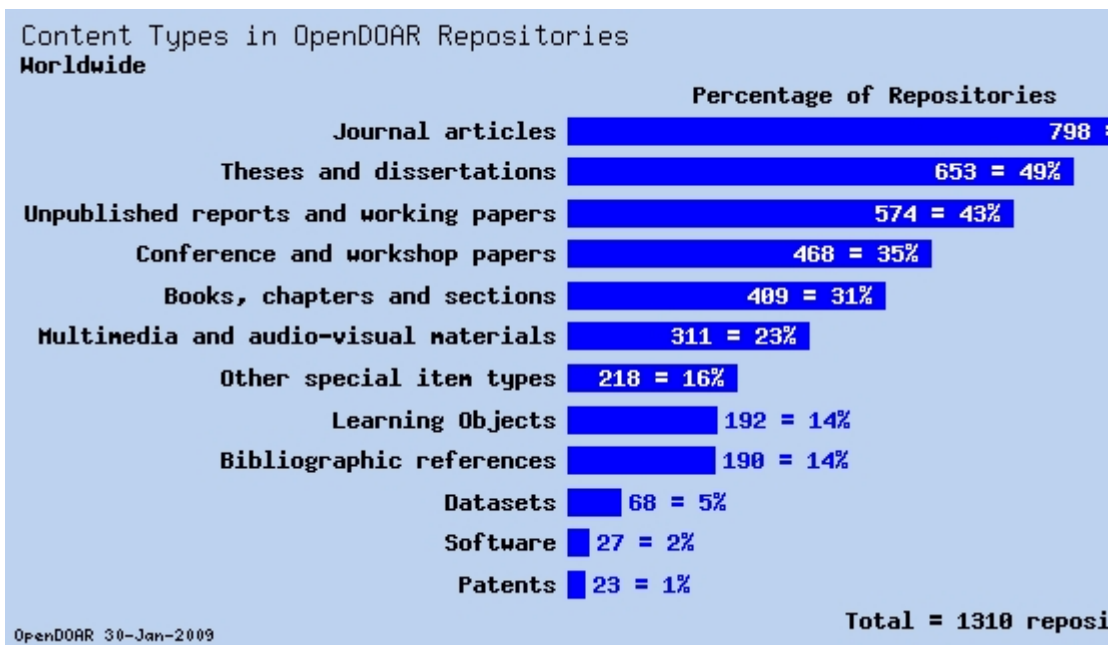
Not all countries are as far down the road to IR implementation as others. The following chart shows the distribution of the 1,200 IRs across the globe as of the Summer 2008.

Proportion of Repositories by Country – Worldwide



This chart is based on the number of repositories in each country. However, some organisations have two or more repositories - over 20 in some cases - and this arguably skews the results. But not significantly. It still shows that the USA, UK and Germany constitute almost 50% of the world's active IRs.

Content Types in OpenDOAR Repositories Worldwide



The above shows how journal articles represent the core ‘open access’ material available in Green OA form. However, PhD theses and ‘grey literature’ generally (such as re-postings, conferences proceedings, etc) as well as multimedia items are also important material which institutions want to see included in their local repositories as these too also reflect the achievements and quality of research and academic activity within these organisations.

Despite these numbers, and the growth over recent years, the number of quality articles available for harvesting through the institutional repository system remains a small proportion of the annual output. ‘Tipping point’ has not been reached in the Green open access movement either. However, ICSTI members may well need to consider the growing interest and support for the Green open access movement and what this switching of the business model, from user payment to author payment, will have on their operations.

C. Author maintained web sites

An early example of the open access movement derives from the traditional habit of authors to exchange their initial 'first stage' reports with other colleagues and peers as exchange of the printed preprints. It led to a significant preprint exchange traffic and culture, most noticeably in physics. With the emergence of the Internet and the web, authors were able to provide the same preprint exchange facility through letting people know of the access route to their personal web site where the electronic version of their results were being held. This saved on speed, administration and mailing costs alone.

Many of the more eminent and established leaders in their subject fields have their own web sites where the complete catalogue of available electronic versions of their research results in recent years are available for free access. It is often considered as part of their contribution back into their scientific discipline to enable easy access to their research output. It enhances their own visibility and reputation in the process. According to an interim report from the University of California, Berkeley ("Assessing the Future Landscape of Scholarly Communication: and in-depth study of faculty needs and ways of meeting them"), in August 2008, "personal websites are ubiquitous, even if they only post a short bio or CV, course lists, working papers, and links to published papers".

It is an often ignored aspect of the 'open access' movement, but one which could be built on with some energy and enterprise. Elements of the exploitation of this can be seen in such ventures as the 'Cream of Science' coordinated by SURF in Holland, where a rich harvest of some of the best papers written by Dutch authors can be accessed. In the UK the JISC-funded 'Digital Lives' project which includes data on leading people is also moving in this direction.

Bringing all three strands together – Gold, Green and Grey – it can be seen that there is support for the principle of free access in many different quarters. The real focus for Open Access is in the electronic delivery of information, and the scale and speed of the combined movement is now unstoppable. And with it comes new ways of conducting business and generating content. ICSTI members should be aware of the implications for their own operations, and how they may need to be adapted to remain 'modern'. But it is all contingent on whether the authors want to play in this arena.

Author Participation rates

Whilst the infrastructure may be in place, authors still need to change their habits and submit their material to these new movements. The choice is to place their research results with the established publishers and see their article benefit from the impact factor assessments which underlie research assessment exercises, promotion, tenure, etc, or else merely see the *potential* for wider distribution of their articles from their own web site, the local server or from a new OA publisher. So far there has been no contest: despite the power of the scientific ethic - when it comes to perceived personal opportunity the established journal publishing system wins hands down.

So in terms of participation in IRs by the scientists the indication is that the number, as well as the percentage of total academics, is still very low. In a D-Lib article published in 2005, it appeared that the Netherlands estimated about 25% of the national research output, across a wide range of disciplines, was going into its institutional repositories. Belgium gave larger estimates – 33% in humanities and social sciences, 39% in life sciences, 16% in natural sciences, and 11% in engineering. With the exception of a 10% estimate for natural sciences in Germany and a 15% estimate for engineering and computer science in the UK, the other estimates (when supplied during a 2006 survey) were negligible. The above proportions seem to give a rosy impression of how well open access has penetrated in scholarly communications in a few select countries only.

This demographic issue has been explored in a number of recent studies which have tried to provide hard data on the attitudes and motivations of authors in submitting their material for open access. Most but not all of these studies have been initiated by publishers and publisher associations (see later).

Concerns of publishers

The traditional scholarly publishing industry has felt challenged by open access. It potentially diverts revenues away from the established, and let's face it often lucrative, subscription, site licensing and Big Deals, to something less proven. And with that uncertainty publishers have been adopting policies which range from dog-in-the-manager to experimental. Some new publishing houses have emerged which have totally embraced the new open access business models, and are structured differently in their operations to cope with the business challenges which open access offers.

It is clear that there is confusion, uncertainty and fear about intellectual property issues (not just getting copyright permissions to deposit, but questions about who will use material that has been deposited, how it will be

used, and whether it will be appropriately attributed). It is a confusion which also involves issues about impact factors and scholarly credit. There seems to be a feeling that material in institutional repositories is of low quality. There are multiples types of content some of which are more mainstream than others, and each has its own level of credibility within the research community. It is also evident that cumbersome and time-consuming submission procedures are a major barrier to open access, and that efforts need to be made to minimise the amount of work faculty must do to submit their work into the institutional repository.

But the real fear of traditional journal publishers is that it will destroy the current market for prepaid article information delivered through a subscription or related payment service.

The PEER project

In order to test this, a programme has been agreed, to be co-funded by the European Commission's eContentplus programme and the publishers, which will monitor over the next three years the actual impact of repositories on the subscription system.

PEER (Publishing and the Ecology of European Research) will investigate the effects of the large-scale systematic deposition of authors' final peer-reviewed manuscripts (so called Green Open Access or 'stage-two research output') on reader access, author visibility, and journal viability, as well as on the broader ecology of European research. A set of articles submitted to repositories by publishers and by authors will be compared against a control set to see whether the commercial effects of IR deposition are quantifiable.

On December 22nd 2008 the formal Invitation to Tender documents were published. It is intended that some 16,000 peer reviewed manuscripts due for publication in between 200 and 300 ISI ranked journals over the next three years will constitute the database. Two of the constituent three parts of the overall PEER project – the Behavioural Research and Usage Research aspects - will be undertaken between May 2009 and August 2011.

The third part of the study – a study on the Economics which derive from publishing in journals and within depositories – will be made available for tender in the summer/autumn of 2009. For more information on the PEER project see the International Association of STM Publishers web site at <http://www.stm-assoc.org/>

CEPA

Meanwhile, a thorough approach to the economics of online journal publishing in more general terms has been undertaken by the Cambridge Economic Policy Associates (CEPA) research group based in London (and with the results published in May 2008). They were commissioned by RIN (Research Information Network) to investigate the cost structure of electronic journals. This has implications on the costs of open access journals as well insofar it indicates the level of per article charges which would be appropriate. The researchers quantified the direct first copy costs of journal publishing worldwide as £3.7 billion. However this excluded the estimated non-cash costs of the refereeing process which publishers have established and manage. This adds a further £1.9 billion to the costs of publishing. Adding in other indirect cost elements produces a global industry investment in scholarly communication of £6.4 billion.

This poses the key question to publishers and authors - whether the process of self-archiving could lead librarians to cancel journal subscriptions – and thereby erode the base of publishers' revenues to support this investment. Some commentators felt that there was no risk as reflected in the claim (by some) that the arXiv free print which has existed for fifteen years in the physics area has had no detrimental effect of the sale of subscriptions. (However, no independent evidence or data has been issued which supports or challenges this assumption).

The issue is whether researchers worldwide could access articles via the author's self-archived copy rather than via the subscribed copy. And whether the growth of services which harvest the metadata about such self-archived articles (and even the articles themselves) will have a growing impact on the library's willingness to carry on with journal subscriptions. One of these services being Google.

Earlier work sponsored by ALPSP has concluded that open access repositories are not currently seen by librarians as a substitute for properly managed journal holdings. (Ware M. "ALPSP survey of librarians on factors in journal cancellations", 2006). However a key conclusion from this report was that librarians believe that 53% (rising to 81% in the following five years) thought that the availability of content through open access and self-archiving systems as being an important or very important factor in determining a cancellation.

To see how relevant the author archives were in determining whether a journal cancellation would be made the Scholarly Information Strategies Ltd

was commissioned by the Publishing Research Consortium to investigate this. The report entitled “Self Archiving and Journal Subscriptions: Co-existence of Competition?”, by Chris Beckett and Simon Inger, was published in 2007. It was based on a model which used conjoint analysis to separate interacting variables which went together to make the final cancellation decision by a librarian. There were six variables considered, including cost, quality, currency, version and percentage of the journal’s articles being already available.

The study confirmed the importance of content quality and cost as the two most important factors in determining selection choices. The version was one of the lowest factors. However, there was also a strong preference by librarians for content that had undergone peer review. The author’s un-refereed original manuscript was seen as a poor substitute for any post-refereed version of an article.

Only 38% of the librarians responding to this survey believed that publishers should not worry about libraries cancelling subscriptions because of open access repositories, and as many disagree (or think that publishers should worry). The authors of the study also believed, on the basis of the evidence available to them, that mandating self-archiving within six months of less of publication will undermine the subscription-based peer review journal.

Studies on Authors

Irrespective of the findings which will emerge from PEER and similar projects over the next three years, there is powerful evidence from elsewhere that author deposition in subject and institutional repositories is not a process for which authors are crying out.

Four studies in particular are highlighted here. The first was undertaken in 2004 by CIBER, a research unit attached to University College London, and entitled “Scholarly Communication in the Digital Environment – What do Authors Want?” (Nicholas, Rowlands and Huntingdon). This report which includes the views from 4,000 respondents indicates that:

- 32% of authors have used their web pages or departmental sites to deposit their publications. A further 53% claimed that they might do so in future. A small minority dismissed this as a possibility.
- 21% said they deposited scholarly material in institutional repositories, and a further 55% said they might do so in future.

- Youth is more supportive of open access than the elderly – 28% of the under 35's would deposit in repositories, compared with only 23% for those over 45 years

A second report, also prepared by CIBER for the main UK publisher trade associations, and published in September 2005, investigated the behaviour of 5,500 authors who published in ISI indexed journals (see “New Journal Publishing Models – an International Survey of Senior Researchers”, Rowlands and Nicholas).

There was according to CIBER a lack of support given to ‘alternative’ publication systems such as institutional repository deposition as reported by 5,500 authors

- 10% of authors claimed they knew a lot or quite a lot about IRs
- 38% declared a clear unwillingness to deposit their articles in IRs
- 3,089 of the total respondents knew ‘nothing at all’ about IRs.

A third study, undertaken by Key Perspective Ltd (KPL), has also looked at open access with a more favourable agenda in favour of openness – it was commissioned by JISC. Entitled “Open Access Self-Archiving: An Author Study” (May 2005) some 1,300 authors’ opinions were received.

- 49% of respondents claimed that they had self-archived at least one article during the previous 3 years in either IRs, subject repositories and on their web sites – particularly on web sites.
- 20% had used IRs; 12% subject repositories
- 81% of authors would willingly comply with a mandate; 13% would comply reluctantly. 5% said they would not comply.

KPL suggested that open access does not destroy the subscription system, and drew on the physics situation to reflect this. The US and UK institutes of physics both claimed that their subscription cancellation rates were not any different from other subject areas even though physics faced the arXiv alternative. Nevertheless, the jury is still out on this as so far the physics publishers have not revealed all the relevant facts.

Finally, the Biosciences Federation (BSF) is a body representing many of the UK's leading biosciences learned societies. Its 35 learned society members have a total membership of most 40,000, most of them charities. 27 of the societies publish journals, producing a total of 75 relatively low-price titles as part of their charitable educational remit, with a substantial proportion of the content being available online free of charge.

There were 1,368 usable responses to the BSF questionnaire. The analysis has highlighted a number of important points:

1. Societies make an important financial contribution to UK universities. Only 10% of the societies' journal subscription and licensing income comes from UK institutions (a total of £1.8M for those societies analysed) but almost 80% of the benefit from their grants and events support goes to UK scientists, students, and their institutions (an annual total of almost £3.9M for those societies providing figures). Thus, they contributed more than twice as much to UK HEIs as they received from them. Any moves that threaten the financial viability of learned societies would therefore have a significant deleterious effect on the funding of the UK science base.
2. All provide free access to most of their material. Of the 17 societies analysed which publish journals, all provide some form of free access to most of their online journal material. This is usually in the form of delayed free access, usually after 12 months. Some also offer optional immediate open access on payment of a fee, but they report very low take-up so far.
3. Amongst the 1,368 respondents to the online survey, there was considerable confusion about what Open Access journals actually are. Almost half of the Open Access journals they said they read, and a third of those they said they published in, were not Open Access journals at all. There seemed to be confusion between online journals (whether providing material free or not) and journals where all material is available free immediately on publication. Thus it is unclear how many of the 74% who said they supported Open Access really understood the issue. Nonetheless, there seems to be substantial support among researchers for the principle of Open Access.
4. Only around 15% of survey respondents said they had tried to access OA publication funds from their institutions or research funders to pay for author-side charges. Of these, 53% had found this fairly difficult or very difficult. This adds weight to the BSF's earlier call for universities to set up ring-fenced funds, and to provide researchers with simple information on how to access them.

5. In general, researchers prefer publishing in established journals to self-archiving

Although almost three quarters of researchers responding to the questionnaire said they considered OA journals a good idea (with the caveat about lack of clarity on the definition), only about one third thought self-archiving (deposit of one of a variety of versions of the paper into university or subject repositories) was a good idea and there was considerable concern about self-archiving. Again, many respondents were confused about what was or was not a repository of self-archived material.

- Three-quarters of respondents are happy to read the final published journal article, but less than 20% said they were happy with the author's final version (ie before it is copyedited and laid out by the publisher, but after peer review). This is the version commonly available in repositories.
- Only 3.5% said they accessed the self-archived version where possible if they also had access to the published version, and 67% never, or rarely, accessed the self-archived version, even if they did not have access to the final published version.
- Only 12.5% of respondents self-archive whenever possible and 71% never do so. Many cited fears about multiple versions and unedited versions as their reasons for this. These responses show that, even where some form of OA publication is required by funders or institutions, researchers still prefer to use the final version, as it appears in peer-reviewed journals, to earlier versions in institutional or subject repositories.

It appears from the survey that researchers are sympathetic, at least in principle, to funded Open Access publishing, although this is not fully borne out by their practice to date, and there is substantial confusion about what Open Access actually is. Researchers are more worried about self-archive repositories.

According the authors of the report, provided it is adequately funded, Open Access publishing could be a viable alternative to the current subscription model in some disciplines. However, there are areas where it is unlikely to work without new funding streams being introduced. This includes subjects such as clinical medicine and systematics, where most research is not supported by grant funding. It also includes review papers, which are often the most highly cited (and by implication most widely read), but which are also not supported by grant funding, and papers from parts of the world where funding would not be available to authors.

If there is a continued expansion of moves by funding bodies and universities to mandate self-archiving with access becoming free within a period that is less than the journals' current time frames for making material free to all, then a point will come at which so much of the material will be free to readers that the current model of library subscriptions is logically likely to collapse. If peer-reviewed journals are not to cease to exist, this implies a move to author-side payments for journal publication, which will simultaneously achieve funders' objectives of making articles immediately freely accessible. This could be achieved if funding bodies were to make the money for this available to researchers via their host institutions, and if institutions were to have robust and clear systems to allow researchers to access these funds.

Not only will this assist the effectiveness of researchers by reducing time spent dealing with confusing and complex systems, but it will also help prevent the potential collapse of the journals currently published by learned societies, which provide the peer-reviewed definitive version of research outputs. In addition it will protect the substantial financial contribution that learned societies make to UK universities, especially in terms of support to postgraduate students and early-career scientists. The report's authors claimed that the UK science base, and in particular bioscience departments in universities, would be seriously disadvantaged if this funding stream were to cease.

In essence these various studies on the authors' attitudes to self-archiving in repositories showed a similar lack of appeal to that which publishers pursued in their support for repository deposition. The studies indicated no great upsurge in demand for this as a requisite for scholarly communication.

Journal Authors' Rights - perception and reality

On a related theme, the Publishing Research Consortium (PRC) will shortly be publishing its fifth Summary Paper, which compares the findings of a specially commissioned study on what authors want to do with their journal articles, and what they believe they are permitted to do by their publisher agreements, with re-analysed data from the recent ALPSP "Scholarly Publishing Practice" survey on publishers' policies compiled by John Cox Associates and Frontline, September 2008. To mark the Academic Publishing in Europe conference in Berlin in early January, a preliminary report has been issued summarising the findings.

From the findings circulated at APE it appears authors underestimate what their publisher agreements allow them to do with pre-publication versions, and

their rights to re-use the published version in their own publications, in course packs, or to send to colleagues; however, they also significantly overestimate their ability to self-archive the final published version. This is based on 1,163 online responses from established authors. Their main wants are to provide copies of their articles outside their own institution (the visibility goal), but their desire to do this through repositories – be they institutional or subject based – is much further down their wish lists. There is a slightly greater preference by authors to post their articles to their own or their departmental web sites, rather than to institutional or subject repositories.

In comparing this with what publishers allow authors to do (as reflected in the ALPSP study) it is apparent that over 80% of publishers allow authors to provide copies to colleagues, incorporate into their own works and to make use in course packs. Only in the cases of posting the published article to repositories and web sites is there publisher resistance (though there is greater acceptance – by 60-80% of publishers – to allowing ‘submitted’ and ‘accepted’ manuscripts to be posted in these services).

The final report highlighting the results from the author and publisher surveys on authors rights will be made available through Publishing Research Consortium in due course (see <http://www.publishingresearch.net>).

Voluntary or Mandatory?

However, the open access advocates are insisting that there is a need to force the issue. In the best interests of the scholarly community, they claim, a more benevolent dictatorial approach may be required.

A point frequently cited is the lack of mandatory provisions in the policies of institutions or funding organisations to deposit the outcome of academic research into repositories such as IRs. However the establishment of such policies, particularly at an institutional level, continues to be controversial. Related to this is a trend towards greater accountability and evaluation of research (such as the Research Assessment Exercise in the UK, which will become the Research Efficiency Framework for the next assessment) and the competition for funding. To the extent that IRs are directly linked to research funding and research evaluation (at the individual or institutional level) faculty have a very compelling reason to deposit material into them.

The real stimulus to an IR publishing model is if it is made mandatory for the author to participate – if criteria are set by the funders to have their published work appear in nominated services, it is likely that the authors will accede to

such demands. Particularly if there is an implied stick in that future funds may not be so readily available if the author ignores the request.

According to the KPL report (see earlier) between 77-96% of authors claim they would willingly accept a mandate policy. This suggests that enforcement is the way IRs would become successful. Voluntary submission does not seem to work. In a classic study of a voluntary system operated by the National Institutes of Health in the US (PubMed Central) only 3% of research was made available through PMC.

However, there is now some evidence that the large funding bodies are willing to commit to a mandatory policy. There is much regional variation about the implementation of mandates.

Regional OA Developments

North America

- The US Congress invited the National Institutes of Health to come up with a policy that would mandate the output of all research wholly or partially funded by the NIH to be included in NIH's PubMed Central. However, the latest proposal from NIH appears weaker than it might have been: it recommended a compulsory code of deposit but with 12 months (rather than six) as the period within which deposit should be made. This became part of the US Appropriations Act legislation signed into law at the end of December 2007. Since then a strong publishing lobby at Congress has produced a proposal to limit the scope of the NIH mandate.
- NIH estimates that approximately 80,000 research papers arise from NIH funds each year. During the voluntary policy, from May 2005 to December 2007, NIH was able to collect a total of 19% of targeted papers, from all sources. Under the first five months of the Section 218 requirement (April to August 2008), the rate jumped up to an estimated 56% of papers per month. While NIH expects to post all of the estimated 56% of these NIH papers, most of them will not be publicly available until 2009 due to the respective embargo periods applied by publishers to the publication.
- Some prestigious universities are adopting open access on behalf of their researchers, notably parts of California and Harvard universities.

European Commission Developments in OA

There are two main centres for open access development at the European Commission. Developments are being pursued through the Research Directorate (under Janez Potocnik), with some being coordinated through the Information Directorate under Vivienne Redding.

The EC has gone through the stages of compiling background information on open access (a study on the economic and technical evolution of scholarly publishing was published in 2006, followed by a stakeholders conference in Brussels in February 2007 and a Green Paper on the future of the European Research Area). This was followed by a decision-making phase with a number of council decisions being promulgated in 2007 and 2008, and is now entering the phase of implementation. Much of the EC work in OA stems from the so-called “Fifth Freedom” which Commissioner Potocnik hailed as meaning the free circulation of knowledge and technology by 2020 within Europe

- This has resulted in a plethora of research projects being funded by the EC many of which have an open access agenda. These include Communia, EuroCancerComs (a forthcoming project in cancer research), LiquidPubs (iterative and live articles), Oopen (in social sciences and humanities), Parse (which deals with aspects of preservation), and Soap which will look at models. Most of these are so-called ‘capacity building programmes’.
- One project particularly relevant to open access is Driver I and II. This is to enhance repository development. Its main objective is to build a virtual, European network of existing institutional repositories using technology that will manage the physically distributed sites as one large scale content source. DRIVER II is funded under the 7th Framework Programme of the European Commission and launched in early 2008.
- Equally relevant is a new programme about to be launched entitled Peer (see above)
- Irrespective of the results of the projects, the European Research Council has indicated that it has a preference for ‘Green’ over ‘Gold’ open access provision. An Open Access pilot has been launched in August 2008 whereby 20% of research funding will require the publication to be made available through IRs. This pilot effects specific disciplines in this initial phase.

A new EC call for proposals is to be issued on 29 January 2009, with a budget of Euros 25 million, which will include open access provision within digital libraries (part of the Competitiveness and Innovation Framework, the old e-Content Plus programme). There is another call for tenders outstanding at present which seeks provide support for scientific information repository building within the EC's current 7th Framework Programme, with a Euros 4 million budget line.

Other European Developments in open access

- The UK Select Committee of Enquiry on Scientific Publications (June 2004) suggested that the UK government should enforce a policy of local institutional repositories at all the key HE/FE centres. The government response was less supportive of this prescriptive route (the business interests within the then powerful Department of Trade and Industry holding sway), and the outcome was instead to allow market forces to determine whether OA journals, IR's, the traditional journal publishing system, or combinations thereof, should operate in the UK.
- Nevertheless, JISC in the UK is helping to create the infrastructure without becoming directly involved in the issue of mandating. JISC is assisting individual UK universities to fund their own institutional repository, and as a backstop, to create a Depot which other higher education establishments could use if they were unable or unwilling to create one of their own.
- Research Councils UK (RCUK) has basically come out in favour of open access with mandates of about six months from date of publication being mandated for its grant recipients.
- A strong lead has come from the Wellcome Trust. This charity has stuck to a firm policy of insisting that those researchers who benefit from Wellcome's research budgets (over £300 million per annum) should make their articles publicly available within six months.
- The Max Planck Society hosted the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities in October 2003 which was a key event for open access in Europe. MPS was also involved in bringing together a German Alliance of Research Organisations which have adopted a common policy on open access. The Max Planck Society's Digital Library initiative in Berlin is also closely involved in the PEER project with STM and EC funding.

- Also in Germany, the Deutsche Forschungsgemeinschaft (DFG), was one of the first signatories to the Berlin Declaration. DFG has also provided a legal footing for free-of-charge access to publicly funded German research in 2006 using either Gold or Green routes.
- France has been particularly active in creating a more national approach to the institutional repository concept with its HAL database of French research documents. This centralised collection of cross disciplinary research has been a particularly unique development and fits in well with the more centralised funding mechanism.
- In the Netherlands, the SURF network has been particularly pioneering with its DARE developments which include the Dutch universities.

Impact on Intermediaries

Library Operations

The role of the library is coming under scrutiny in a number of areas (a possible topic for a future ICSTI Insight?) Of particular importance is the effect which Google's digitisation of books programmes (with both publishers and libraries) is having, but if this is taken in tandem with the growth of the open access movement affecting research articles, which Google and other search engines will be able to 'harvest', then it gives one organisation a dominant position in providing free access to scholarly information. Whilst Google claims that it will not become a 'publisher' (APE, Berlin, 2008), the fact that one organisation will sit on a vast and growing mountain of scholarly material may raise future questions about Google's future role as the gatekeeper in the information dissemination process. Text and data mining are technical processes it could apply to such vast published resources for the betterment of science and society.

Many libraries have fought against the idea of them becoming a 'museum of the book', but unless they shake off the role as being mainly a dominant collection development agency in favour of striking a closer relationship with its clients in providing bespoke and customised/personalised information services it may lose out to the gargantuan search engines in fulfilling their mission.

Publishers Operations

The challenge facing publishers, as indicated earlier, is that they face an insecure future if the open access business model takes over from the subscription business model. The open access business model is currently not robust enough to generate the margins which large commercial publishers in particular have become used to. Although one of the largest commercial publishers has acquired a large open access publisher (Springer S+BM purchased BioMed Central in September 2008) this is seen as a strategic/financial move related to achieving corporate investment appeal with little operational consequence for the near future.

Publishers have a difficult choice to make. How will they fund the process of managing the quality control process in future in all its aspects? If subscriptions are seen to be an outdated mode of payment in an Internet economy what else is available? Author payments are still not de rigueur; online advertising is not yet prevalent in this specialised industry sector, and subsidies and charitable contributions are not long-term options. Publishing may well have to absorb parts of the new electronic publishing developments, including Green open access systems, social networking, adoption of e-Science procedures, work bench support and provision of total integrated information systems from 'soup to nuts'. Reliance on selling articles on their own may become marginalised as Web 2 and semantic web developments continue to change the face of scholarly communication. And apparently 'free' information will become more evident and it will take many forms.

These are huge challenges publishers need to face and will require much closer identification with the real needs of users and authors, and less on churning out material in traditional formats.

Harvesting open access material

In tandem with the two main formal approaches to open access – the Gold and Green routes - a technical structure has been established to enable consistency in gaining access to the freely available research material. A common protocol – open access initiative, protocol for metadata harvesting or OAI-PMH - was developed. This would enable the header information on such articles to be 'harvested'. Though minimalist in some descriptive areas, and subject to varying degrees of quality from the authors, the basis for accessing a large corpus of free articles on demand was ostensibly put in place.

The OAI-PMH protocol was derived from the open source movement which, as has been highlighted, has become a significant force within IT. This is in principle an ideal approach, enabling a multi-institutional search for particular items. Services such as OIAster have emerged to offer such aggregated search, and GoogleScholar and similar search engines (such as Elsevier's Scirus) are picking up the references contained within the growing number of IRs.

But principle and practice are not always in synch, and there is still some suspicion that metadata processing and federated searching are not fully sustainable. A distributed approach suffers from the way it is implemented at local level, and whether full interoperability is really achievable. The combination of limited deposit controls, broad range of content, variable quality, quantity and consistency of metadata means that content in IRs will be difficult to use as a replacement for traditional journals any time soon.

One commentator (Chris Keene) reports in his blog that "this is proving to be a steep learning curve in the joys of OAI-PMH, and how the different repository systems (and the different versions on these systems) have allocated information about the fulltext in to different Dublin Core (DC) elements." There still seems to be some way to go before complete interoperability between the various IR systems and IR applications is achieved.

This is a problem which subject-based repositories seem to have largely overcome. It is therefore no wonder that subject-based repositories (such as PubMed Central) are feared more greatly than the disjointed activities of institutional repositories in the eyes of publishers.

The Copyright Issue

Fundamental to the issue of open access is whether the 'freeness' which underlay OA is a good thing for society or not. It could be argued that open access is a destructive force – it goes against the ability of individuals to claim credit and ownership over the sweat of their brows. As such it could reduce incentives to publish their research results openly.

The publishing industry believes that the interests of individuals can be protected by the publishers seeking 'licenses to publish' rather than claiming total ownership of copyright. Author licences were not seen to be a big issue until Professor Lawrence Lessig (Stanford University) launched the Creative Commons which were dropdown menus enabling authors to determine what

rights they as authors were giving up to publishers and what rights they were prepared to keep. These licences were quickly adopted by the main open access publishers (BMC, PLoS, etc) and the conditions included in the licences are being acknowledged by the major search engines (Google, Yahoo). Commercial publishers also adopt the variants of the CC licences as implicitly such licences recognise the importance of ownership of intellectual rights (as decided by the author) over a free-for-all which some open access advocates believe in and the Internet inherently supports.

The issue of ownership of intellectual property is unlikely to go away under the growing support for open access – but it is an open access regime which accepts that individual creativity should be protected and rewarded to some extent. Perhaps not as extensively as under a subscription-only business model however.

Other Stimulants towards open access

Data and datasets

More journals, both open access and subscription-based, encourage OA to link the data underlying published articles. Major publisher associations such as ALPSP and STM, which lobby against national OA policies for text, encourage OA for data. Even when these policies do not cover peer-reviewed articles, they accelerate research, demonstrate the benefits of unrestricted sharing, and build expectations and momentum for OA in other categories.

Opinions vary on sharing and publishing data sets. Generally most scholars are agreeable to sharing as long as they have finished their analysis and publication of their data. Sharing can ultimately depend on who is doing the asking and what they want to do with the data. Journals in the sciences, economics and political science in particular increasingly require that data sets be published. There is also a move for funding bodies to promote the publication of data. Whether this policy is mandatory differs among agencies.

The mass of data obtained at great expense from satellite missions, global sensor networks or large scale basic-research experiments, as well as clinical studies and statistical surveys, require traceability, plausibility and re-use by colleagues in the field. But it is not just the Big Science projects which need open access to their data; just as important but much more difficult is the plethora of small, specialised data which is available to each researcher at the end of their particular project, and which is often disposed of without consideration for its relevance for future studies. There is the rub – deciding

what data is useful to keep and what is not. And if it is to be kept, what standards and metadata attribution is required, and who does this?

Some scholars publish supplementary datasets on their websites. This practice allows for the reuse of data and provides transparency to facilitate better scholarship – in many instances, institutional support for hosting and managing such data is not provided.

Social Networking and Web 2.0 impact

Whilst open access journals and institutional repositories continue to develop, other vehicles which adopt the open access principles are mushrooming. These include blogs, wikis, ebooks, podcasts, RSS feeds, and P2P networks

In a study being undertaken by the University of California, Berkeley, they found that most of the researchers they interviewed felt “blogs were simply off the radar as a source of scholarship and are generally viewed as a waste of time because they are not peer reviewed”. There have to be filters – the refereeing system.

However, new and effective tools for collaboration are also triggering OA adoption. Social tagging, searching by tags, open peer commentary, searching by comments, social networking, community building, recruiting collaborators, facilitating work with established collaborators, following citation trails backwards and forwards, following usage-based “similar to” and “recommended” trails, open APIs, open standards, and ‘mash-ups’ – these have opened up a whole new vista in electronic publishing (EP). A new generation of digital scholars is building on the new collaboration services that also build on OA.

Nevertheless, the important impact of disciplinary culture and tradition on many scholarly communication habits is a powerful force for conservatism. These disciplinary traditions may override the perceived ‘opportunities’ afforded by new technologies, including those falling into the Web 2.0 category.

Digitisation projects

Huge book-scanning projects, particularly those from Google, the Open Content Alliance The European Library, the Kirtas-Amazon partnership, and Project Gutenberg, increases the number of print books available in some free-to-read digital form. Also, the price of book scanning is dropping quickly

as large organisations see the investment return potentials from large-scale digitisation projects.

Some of these digitisation projects include:

- Google Library and Publisher Very significant – provides a massive book digitisation programmes digital collection which can be drawn on
- Open Content Alliance Potentially strong European force leading towards digital information adoption
- Microsoft/Ingram An aggressive Ingram push could promote swing to digital over text as preferred storage system
- Internet Archive (Brewster Kahle) Private initiative but with much support and Wayback machine
- KB Koninklijke Bibliotheek and Enlightened promoter of the digital cause Safe Places along many fronts. Particularly e-journals, books and newspapers
- The British Library Undertaking some retro-digitisation with JISC funding (particularly of old newspapers)
- Others (Wellcome/NLM) Developing impact – not significant at present

Recent developments

Open Access Day

October 14th 2008 was nominated as 'open access day' by the many OA supporters globally. It was meant to show the world that open access is healthy, thriving and here to stay. It is claimed that there were 116 events celebrating 'open access' taking place in six continents throughout the world on the same day. One of these events was the launch in the UK of OASPA.

Open Access Scholarly Publishers Association (OASPA)

At a meeting held in the atrium of The Wellcome Trust building in London on Open Access Day, attended by some 40-50 people across all information sectors, the Open Access Scholarly Publishers Association, OASPA, was launched. This is to be for the open access publishing movement what ALPSP

has been for learned society publishers and STM for international publishers. OASPA will aim to support and represent the interests of open access (OA) journals publishers globally in all scientific, technical, and scholarly disciplines through an exchange of information, setting of industry standards, advancing business and publishing models, advocating for gold OA journals publishing, education and the promotion of innovation.

The founding members of OASPA represent a broad spectrum of OA publishers and include: BioMed Central, Co-Action Publishing, Copernicus, Hindawi Publishing Corporation, Journal of Medical Internet Research, Medical Education Online, the Public Library of Science (PLoS), SAGE, SPARC Europe and Utrecht University Library. Representatives from each of these publishers will form an interim board until a first General Meeting is held during 2009.

Further information including membership criteria and an application form can be found on the OASPA website, www.oaspa.org.

Summary

The emerging welter of conflicting practical and operational trends might not give comfort to either friends or foes of OA, or to anyone trying to forecast the future with any confidence. There are some key strands which stand out, but making a consistent sense of them all is proving difficult for all stakeholders. It is something of a jigsaw.

There are the many trends created by OA proponents. These include: the growing number of OA repositories, an increase in OA journal numbers, OA policies for mandating at key universities, OA policies at public and private funding agencies, and public endorsements of OA from notable researchers and university presidents and provosts.

Funding agencies are now considering OA policies because of their intrinsic advantages for increasing return on investment and by increasing the visibility, utility, and impact of research conducted at their institution.

Subscription prices are still rising faster than inflation. The economic crisis affecting many national economies is adding further constraints – some countries (such as the UK with rapidly declining currency exchanges) - are affected worse than others. Rapidly rising prices undermine the sustainability

of the subscription model. They undermine publisher arguments that all who need access can get access.

Furthermore, since the rise of peer-reviewed journals in the 17th century, most publicly disseminated works of scholarship have been refereed and distributed by publishers. Letters and lectures were exceptions. Today, the categories of exceptions, the volume of research-reporting they represent, and their integration into the workflow of ordinary research, are all growing.

The shock of the new is wearing off. OA is gradually emerging from the fog of misunderstanding. Time itself has reduced panic surrounding OA. Everyone is getting used to the idea that OA literature can be copyrighted, that OA literature can be peer-reviewed, that the expenses for producing OA literature can be recovered, and that OA and subscription-based literature can co-exist.

As such the paradigm is changing, and open access is a key driver in making this happen. How significant this will be in the long term may be questionable, but in the short and medium term it will dominate the agenda for electronic publishing and has become a crucial issue within the corridors of those who can make an impact on the publishing scene – the large research funding agencies.

Only when this impact drills down to change the behaviour of research authors, when they are either persuaded or forced to change their views on where their research results are made optimally visible (by their own criterion), will open access make the breakthrough which pundits have long claimed and advocates have long hoped for. But there is still some way to go before this happens. Changing technology is much simpler than changing a social culture and habits with respect to scholarly communication.

Open Access – What shape the Future?

To return to the opening question, what is the shape of the future for open access? Open access is inevitably something which will be part of the scholarly communication scene, and will gather pace as the administrative forces (mandating and funding agencies) dominate proceedings. It has a groundswell of support from those people who see the value to society of allowing as widespread dissemination of public-funded research as is possible.

The arguments raised by publishers are, in contrast, mainly seen to be self serving.

As such we will witness an increasing proportion of articles coming into public domain as publishers search for alternative, less viable, business models to sustain their operations.

Much depends on the support given to the open access movement by the people who really count – the authors and users of such scholarly journal articles. Changing their attitudes and perceptions will be a longer process. It is one which will keep the erosion of open access into the traditional subscription-based publishing in a steady state rather than becoming a revolution.

It is suggested that on the basis of the above evidence and impressions, open access will grow from a 15% share of the scholarly journal article business (in all its forms, but mainly Green versions) to 30-40% by 2015.

However what has to be considered is how the open access movement will be driven by other trends which will become 'ICSTI Insight' investigations in due course. Whether open access in data and datasets will take off, and subsume some of the functions of the formal articles as part of its growth, has to be considered. The future of articles per se – whether open access or formal published items – will remain as the mainstay of scientific communication or whether they will be displaced by something 'better' or more appropriate, is something which needs to be considered.

By the same token, the growth of informal communications using the Web 2 (and ultimately Web 3) developments, could impact seriously on that cornerstone of the current scholarly communication process, the refereeing system. If the 'Google Generation' chooses to be as equally cavalier about accessing quality-controlled journal articles in future as they seem to be in their general search behaviour currently ('something is good enough') then one of the main planks of both the OA and TA publishing activities disappears.

Then there is the role which the other elephant in the room will play – Google. As long as it continues to see itself only as a search engine, opening up free access to research literature (subsidised by advertising), it will be a key driver for change towards open access. What could become more intriguing is if Google changes tack and becomes a more aggressive information packager and re-packager of the mountain of digital information it is collecting in book, journal and data formats. This would hasten even more quickly the migration to open access systems.

So the argument is not about open access versus subscriptions, but rather about the optimal forms the future scientists will take to gather

their information. Open access may well become a relevant process, but the formats may become much broader in functionality than what is in general use today, and it is this new demand for new information delivery systems which will drive open access forward to becoming a dominant process in the 10-20 year scenario. The arguments between subscription and open access business models will become more of a side show.

Prepared for ICSTI by SCR Publishing Ltd, Oxford, UK

Copyright © 2009 ICSTI. All rights reserved. No part of this product or service may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written consent.