Open Access, Almost-OA, OA Policies, and Institutional Repositories

RICHARD POYNDER

1ˢᵗ December 2015

Many words have been spilt over the relative merits of green and gold open access. It is not my plan to rehearse these again here. Rather, I want to explore four aspects of green OA. First, I want to discuss how many of the documents indexed in "open" repositories are in fact freely available, rather than on "dark deposit" or otherwise inaccessible. Second, I want to look at the so-called eprint request Button, a tool developed to allow readers to obtain copies of items held on dark deposit in repositories. Third, I want to look at some aspects of OA polices and the likely success of so-called IDOA policies, and finally I want to speculate on possible futures for institutional repositories. I am splitting the text into two. The first two topics are covered below; the second two will be covered in a follow-up piece I plan to publish at a later date.

One of the inconvenient truths that the OA movement prefers not to discuss is the fact that a large amount of the content in the circa 4,125 institutional repositories created by research institutions in order to provide open access to their research output is not actually freely available but on "dark deposit", or otherwise inaccessible. In other words, it is not open access.

There are several reasons for this. First, publishers of subscription journals now insist that papers they publish must undergo a period of embargo before the full text is made freely available.

Second, as the evaluation and monitoring of researchers grows in both breadth and depth so repositories have become vulnerable to being "captured" by university managers, who increasingly want to use them for other purposes. This can see access provision subordinated to goals for which full text availability is not necessary. Partly for this reason, many of the records in open repositories consist only of bibliographic details, not the documents themselves. While these records may include a link to the full text document hosted elsewhere, this will likely be behind a paywall.

Third, even where the full text has been deposited in a repository, the author (or their institution) may — for reasons that are not always clear — insist that a work is placed on dark deposit. Thus although the full text may be present in the repository, it may only be accessible to members of the institution — i.e. behind what we might call a "login wall".

After studying these matters, and reporting on their findings in a 2014 paper published in *D-lib Magazine*, Hélène Prost and Joachim Schöpfel concluded, "Open archives are less open than they should be."

It is, of course, perfectly legitimate (and sometimes unavoidable) for documents to be placed on dark deposit, but as Prost and Schöpfel point out doing so is "not in line with the underlying principles of the open access (OA) movement that defines open access 'as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community' and that requires freely available scientific literature'."

In the dark

But exactly how much content in open repositories is dark today? Unfortunately, we don't really know. To try and get a clearer picture, Prost and Schöpfel conducted a survey of 25 institutional repositories and over 2 million records, 53% of which were peer-reviewed papers. Their conclusion: the rate of open access varies from repository to repository from just 4% OA to nearly 100%.

Of those repositories they surveyed, they added, the median figure for items that provide open access to the full text is around two-fifths.

The amount that is OA also varies by content type, they explain: "The overall degree of openness of working papers is 0.96, which means that in the entire sample all but 4% of the working papers are freely accessible,

followed by PhD theses (0.76) and reports (0.63). Significantly less open are journal articles (0.31), communications (0.21) and books/book chapters (0.17)."

To put it another way they add, "articles are half as open as reports, and PhD theses are over four times more open than books or book chapters. The median degree of openness per repository confirms the overall statistics. The median is high for working papers (0.98) and theses (0.92), medium for reports (0.63), and low for articles (0.38), communications (0.29) and books (0.13)."

What is noteworthy here, they suggest, is that "the number of items and their openness are inversely correlated, in that the more important categories (articles, communication and books/book chapters) are less open than the less important ones."

Interestingly, while OA advocates routinely blame dark deposit on publisher embargoes, Prost and Schöpfel did not think that this is currently the primary reason. "From our results we can only make a cautious guess that embargo periods represent a small part of access restrictions (in our sample only 2%) and that most of the lack of openness is caused by on-campus only access and by the deposit of metadata without a corresponding document."

One reason why so many records are only available as metadata, they suggest, is because "institutions have seized the opportunity offered by institutional repositories to gain control of their own scientific output. Large and exhaustive repositories allow for scientometric evaluation of research results and productivity; here metadata are important while access to full text is secondary, marginal."

Prost and Schöpfel's paper is just one study of course. Nevertheless, others have reached similar conclusions. In a paper published in *Information Research* in 2015, Jongwook Lee *et al* surveyed 170 articles deposited in the Florida State University repository DigiNole Commons, and found only 60% were open access.

Elsewhere, the UK's Open University has reported that only around 38% of the documents in its ORO repository are open access. Even the poster child of the self-archiving movement — the University of Liège's ORBi repository — contains a large number of records without the full-text. At the time of writing, ORBi contained 128,353 items, of which only 80,090 (62%) included the full text.

When the openness of recent research is considered the picture is still gloomier. In 2011, for instance, Danny Kingsley noted that Primary Research Group's 2011 'Survey of Database Licensing Practices' (page 29), reported that when libraries were asked what percentage of their institutions' published scholarly output from the previous year was deposited in their digital repositories they found the mean figure to be just 23.23%.

The reduced openness of *recent* research is important to note because when reporting deposit numbers repository managers will publish the sum total of items deposited. They do not generally distinguish by type of content, or by date. Since older works — which are invariably less useful — will predominate, the published figures will be an overestimate so far as relevance and timeliness are concerned.

Looking again at ORBi, a EU-funded study by PASTEUR4OA (which I will discuss in more detail in part two) found only 37% of recent papers in ORBi were open access, with 50% on dark deposit (and 12.9% not deposited at all).

The difference between this and the higher figure published on ORBi's home page is presumably explained by the fact that PASTEUR4OA extracted bibliographic data from Web of Knowledge (WoK)-indexed journal articles published in 2011-2013 and then looked to see if they were freely available in ORBi. As such, it sought out only papers (not other types of documents), and only recently published research.

The PASTEUR4O figures would seem to confirm a suspicion I have long had that repository managers are depositing a lot of historical data. While this will boost numbers it will not greatly increase the usefulness of the repository. It might also seem to support a point made by Prost & Schöpfel cited above: "the number of items and their openness are inversely correlated, in that the more important categories (articles, communication and books/book chapters) are less open than the less important ones."

But while we can speculate as to why so much content is on dark deposit, we cannot know for sure, due to a paucity of data. Nor, point out Prost and Schöpfel, do we know whether it is a short-term or a long-term phenomenon. "Often the real nature of access restriction remains uncertain. Are the documents under embargo and will they be released and openly accessible in the future? Are they restricted to on-campus only access or is it both of these? And what about missing full text, records without documents?

Nevertheless, it might not be unreasonable to assume that — so far as journal articles are concerned at least — the situation is likely to deteriorate rather than improve going forward. As funders and institutions introduce more mandates, and publishers impose new embargoes, as more authors request dark deposit for their documents, and as institutional repositories begin to merge with CRIS systems in order to allow repository content to be used for research evaluation purposes, the lights could dim further. In 2013, for instance, Danny Kingsley pointed out that the percentage of OA content in Australia fell as a result of the Excellence in Research for Australia (ERA) exercise.

She explained: "In Australia, CAIRSS holds an annual survey of activities in Australian (and now New Zealand) repositories … The results of the CAIRSS Repository Managers Surveys conducted over the past three years (2009–2011) show that the 'percentage of material in repositories that is open access' across Australian universities has fluctuated. It started strongly with 44% in 2009, dropped to 33% in 2010 and rose back up to 37% in 2011."

She added "Libraries are under-resourced, and many universities have a cross-over of repository and reporting staff, so the focus of these people was, by necessity on ERA reporting not open access during 2010."

Overall, concluded Kingsley, "the evidence seems to indicate ERA has been detrimental to the promotion of open access in Australia."

And in a 2014 paper called *Back to Grey: Disclosure and Concealment of Electronic Theses and Dissertations*, Prost and Schöpfel concluded that electronic theses are becoming darker. As they put it: "Our paper describes a new and unexpected effect of the development of digital libraries and open access, as a paradoxical practice of hiding information from the scientific community and society, while partly sharing it with a restricted population (campus)."

It turns out, they add, that the Internet "is not synonymous with openness, and the creation of institutional repositories and ETD workflows does not make all items more accessible and available. Sometimes, the new infrastructure even appears to increase barriers."

Mandates, embargoes, and Almost-OA

So there are a number of reasons for dark deposit, and its incidence depends on type of content. Below, however, I want mainly to focus on research papers and publisher embargoes, and to look at the way in which OA advocates have responded to these embargoes.

Let's start by reviewing how we got where we are today. As OA advocates frequently remind us, in 2001 the Budapest Open Access Initiative (BOAI) outlined two strategies for achieving open access (green OA and gold OA). I am going to leave gold aside here and concentrate on green OA.

In the wake of the BOAI, green OA advocates devoted a great deal of time and effort to trying to persuade researchers to self-archive copies of the papers they publish in subscription journals in open repositories.[1] By doing so, it was said, they would ensure that their work was available outside subscription paywalls, and so free for all.

However, researchers did not respond in great numbers, despite (or perhaps because of) the increasingly strident nature of the exhortations to deposit their papers. The number of researchers that spontaneously

---

[1] The first such exhortation was made by Stevan Harnad in 1994.

self-archive has stayed relatively stable at around 5-25% (depending on the field of study). Faced with at best indifference, at worst active resistance, to their calls OA advocates shifted focus and began instead to lobby institutions, funders, and governments to make open access compulsory, by means of open access policies, or mandates[2].

Yet most researchers continued to ignore calls to self-archive, even when they were subject to an OA policy. In response, OA advocates began to call for increasingly punitive policies, and today they insist that the only way to make open access a reality is to tie OA polices to performance evaluation. That is, researchers must be told that only those research outputs that have been deposited in an open repository will be taken into account when they are evaluated for promotion and tenure.

While most researchers objected to the principle of having open access forced upon them, OA advocates did begin to have success with funders and governments, who started introducing OA policies.

Publishers too resisted open access, especially green OA, which they believe poses a threat to their revenues. However, while the logical response for publishers would be to impose embargoes before papers can be made freely available on the Internet, the general lack of interest in self-archiving amongst researchers, and the haphazard way in which it was occurring, saw few publishers initially do so. They assumed that it would never amount to anything more than a niche activity.

But as more and more funders introduced mandates, publishers began to worry that if a sufficient number of papers in a journal became freely available on the Web librarians would cancel their subscriptions. It was in order to try and prevent this that they eventually did begin introducing embargoes.

It should be noted that there is no consensus on whether self-archiving is indeed a threat to publishers. Open access advocates insist there is no proven need for embargoes. In any case, they add, they should never be more than six months (certainly for STM papers).

Whether OA advocates are right is not clear as there is no reliable evidence to support either side of the debate. What research has been undertaken has been driven by advocacy (on both sides), rather than by a desire for enlightenment. As such, the findings of the few studies that have been undertaken are open to question.

Nevertheless, when introducing open access policies funders generally veer on the side of caution and permit publisher embargoes, although they invariably specify a maximum acceptable period. In practice, they generally start out proposing a six month embargo, but subsequently lengthen it in response to publisher lobbying. Today the norm for funder embargoes seems to be moving towards 12 months for STEM subjects, and two years for HSS subjects.

The OA movement has therefore had to accept embargoes as the price of having OA policies. In recognition of this reality, for instance, in 2011 the developers of the DSpace repository software added the ability for deposits to be held on dark deposit, as Kingsley noted at the time.

Today publishers use a number of different strategies to discourage researchers from self-archiving. They are, for instance, introducing more and more gold OA options for their subscription journals, and then pushing authors in that direction. And as noted, they have introduced embargoes where they did not previously exist, and then continuously sought to lengthen them — see for instance here, here and here. While a few publisher embargoes are still as short as six months they can be up to four years and beyond (See Elsevier's *Energy Economics* for instance). They have also sought to make their self-archiving rules as complicated and obfuscatory as possible, and they change them repeatedly.

---

[2] There has been much discussion about the difference between an OA policy and a mandate. While mandates sound more compulsory, in practice the terms tend to be used interchangeably — despite attempts to make a distinction between a policy "requesting" or "requiring" open access (as with the NIH Public Access Policy and the PASTEURO4OA study).

The latter is clear to see if we consider that the largest STM publisher (Elsevier) has amended the wording of its self-archiving policy on frequent occasions, gradually tightening and increasing the complexity of the rules in the process. This has attracted regular complaints from the research community — e.g. here and here.

The result today is that — where they have the necessary funds — researchers inevitably find gold OA to be a more attractive option, not least because it is a much easier way of complying with OA policies.

Embargoes are a particular source of frustration for green OA advocates, not least because subscription publishers routinely insist on acquiring the copyright in the papers they publish, and so have a great deal of power to set the rules, and to change them at will.

In 2006, therefore, green OA advocate Stevan Harnad proposed a strategy that he predicted would neutralise publisher embargoes, by allowing any self-archived paper to be freed on request.

First, he said, funders and institutions should introduce what he dubbed the Immediate-Deposit/Optional-Access (IDOA) Mandate. With such a policy, Harnad said, researchers would be required to deposit all their papers in an open repository at the point of acceptance by a journal. In order to accommodate any embargo, however, the policy would allow authors to specify that their work is made available on a closed access basis. Such papers, said Harnad would be "Almost-OA".

As he explains: "[W]hether access to that deposit is immediately set to Open Access or provisionally set to Closed Access (with only the metadata, but not the full-text, accessible webwide) is left up to the author, with only a strong recommendation to set access as Open Access as soon as possible (immediately wherever possible, and otherwise preferably with a maximal embargo cap at 6 months)."

We should note that such a policy not only accepts the inevitability of publisher embargoes but it legitimises dark deposit.

Importantly, however, there is a second *and vital* component to the IDOA mandate. All institutions, said Harnad, should install a "copy request" or "fair dealing" Button into their repository. When implemented this would allow closed access papers to be released to anyone requesting a copy. All the repository user need do is click on a link in the bibliographic record and the repository software would automatically forward the request to the author. In response, explained Harnad, the author need only "click once to comply with the request".

If the author approves the request the repository software will then send the requester a copy of the document, or more usually a private link to it.

The IDOA mandate **combined** with the "copy request" Button, argues Harnad, was "specifically formulated" to ensure that providing immediate access to papers is "immune from any delays or embargoes (based on publisher policy or copyright restrictions)."

In other words, the Button was designed to allow anyone wanting to read a research paper held on dark deposit in a repository to quickly and simply obtain a private copy, regardless of any embargo — so long as the author approved the request.

Clearly, the Button can also provide access to the full text of other documents on dark deposit in a repository. What it cannot do, however, is provide access to a document where the repository only holds the bibliographic details.[3]

The combined objective of the IDOA policy plus Button is thus twofold: to ensure that all research papers subject to it are deposited in an open repository, while *also* mooting any attempt by publishers to prevent those papers being made freely available. The assumption is that the widespread introduction of such policies

---

[3] Unless the repository can persuade the author to find and post the full-text in order to fulfil the request.

will see the number of papers that are freely accessible grow rapidly, with dark (Almost-OA) papers capable of being freed on request.

In the past few years green OA advocates have been advocating tirelessly for institutions and funders to introduce IDOA policies, and for repository managers to install "copy request" Buttons to support these policies. And as we noted, in order to ensure that researchers comply with such policies, they have lobbied to have compliance tied to the evaluation process used by the institution's promotion and tenure committees.

In light of the calls for compliance to be tied to evaluation we might want to introduce a new term — the Performance Evaluation Mandate, or PEM. The first IDOA/PEM mandate was introduced in 2007 at the University of Liège — by OA advocate Bernard Rentier (then Rector of the university)[4]. Since then the model has been copied by a number of other institutions although (as we shall see in part two), the Harvard-style policy has been more widely adopted, notably in North America. Importantly, however, IDOA/PEM is the model that green OA advocates routinely now recommend.

We have noted that Almost-OA legitimises dark deposit. It also seems reasonable to assume that as more and more papers are posted in repositories as a result of OA policies, publishers are inevitably going to seek to impose ever stricter rules on self-archiving. The upshot will surely be that more and more papers in repositories will be on dark deposit. As such, dark deposit will prove an unwelcome and contradictory consequence of the success of green OA advocacy.

The Button in practice

Given this, it is surely vitally important that the copy request Button proves successful. So how effective is it?

As we shall see, this is not entirely clear. What is clear is that there are a number of issues working against its success. For instance, not all repositories have a request eprint Button installed. In fact, it may not be possible to implement the Button in many of the 32 different platforms listed in the Registry of Open Access Repositories (ROAR). While we know that the most popular platforms — EPrints and DSpace — do offer the functionality, the Fedora software (which is available through a number of different front-ends, including Islandora, Hydra and Fez) apparently does not; nor does the increasingly popular Digital Commons platform.[5] We also know that the open source software Open Journal Systems does not.[6]

What of the other 20+ repository platforms? This information does not appear to be available.

The second problem is that even where the Button has been built into the repository platform, repository managers may not implement it. How common is this? Again, there appears to be no data available.

Third, and most importantly, even where the Button functionality is available in the underlying software, and even where the repository manager has implemented it, it is far from clear that researchers routinely approve requests for access to their papers. In fact, it appears that many do not respond to requests at all. In other words, it is far from clear that the Button is efficacious even when it *is* implemented.

One sceptic is German archivist and historian Klaus Graf who has questioned the effectiveness of the Button on a number of occasions (here and here for instance).[7] As he put it in 2009, "most scholars in my several tests haven't reacted on my request Button tests."

Of course, Graf's testimony is personal experience; it is not based on a disciplined study.

---

[4] Bernard Rentier was Rector of the University of Liège between 2005 and 2014.

[5] There is a Button of sorts, but it simply loads the lead author's email address into the requester's email software — see here for instance.

[6] While it was designed as a journal publishing platform, OJS is used by some institutions to run their repository.

[7] Graf has also drawn attention to the large number of papers in repositories that are on dark deposit.

Nevertheless, I too have some personal experience I can share. Until last year I had given little or no thought to how effective the Button is, since I had never had occasion to use it. Instead I relied on the #icanhazpdf hashtag on Twitter (with only occasional success it has to be said). I suspect that OA advocates who regularly laud the virtues of the Button do not use it much if at all either, since they likely have access to most of what they need via their institutional subscriptions, or via ILL.

So what is my experience with the Button? In October 2014 I wanted to read a paper called *Institutional repository as an important part of scholarly communication*. To read it on the publisher's site I would have had to pay Emerald £20.00 for 30 days access. Unwilling to do this, I did a web search and discovered that a copy of the paper had been self-archived in the author's institutional repository at the University of Ljubljana. Great, I thought, and clicked the download link. But I immediately hit a problem: the self-archived version was sitting behind a login wall. And while there was a request eprint Button, when I hit the relevant link I received a message saying that my message could not be sent (see image below):

Image 1.

This was not encouraging. Moreover, it was not clear whether it was just a temporary problem, or an inherent fault in the way the Button had been implemented in the repository. Since then I have tried to request the paper on a number of further occasions, each time with no greater success — the last time in November 2015, over a year after my first attempt.

I assume the reason why this paper is on dark deposit is because at least one of the authors is in a faculty that has an OA policy and Emerald's current rules state that where an author is subject to a policy, a 24-month embargo applies. Specifically, one of the authors of the paper is in the Faculty of Civil and Geodetic Engineering, whose policy can be read here.

Looking again at the message I received, it occurred to me that the lack of response to my request may have been a problem with the repository email system. This seemed to be confirmed when I tried to clarify the situation by emailing the administrator via the address published in the repository (drugg@fgg.uni-lj.si). My message bounced with a note saying, "The email address that you entered couldn't be found."
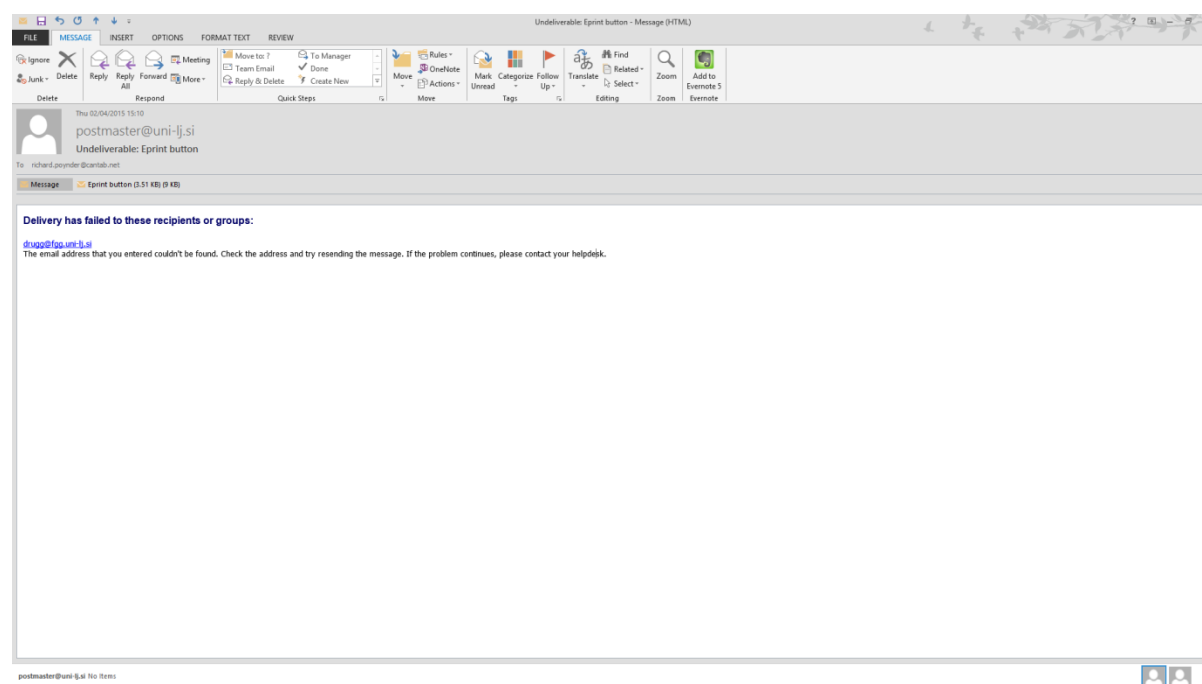

Image 2.

This suggests a fourth potential problem with the Button: even where the repository platform has the functionality installed, even where the Button has been implemented, and even where the author may be willing to respond to copy requests, other technical problems may intervene to subvert the process. And it would seem that the repository manager may not even be aware of the problem (and cannot even be told about it).

Lost in an email stream

But my adventures with the Button were to become even more puzzling. The same month (October 2014) I tried to obtain another paper in the same way. This paper (written in 2007) was called *A Longitudinal Study of the Practice of Self-Archiving* (see below) and appeared to be on dark deposit in the University of Southampton's repository.
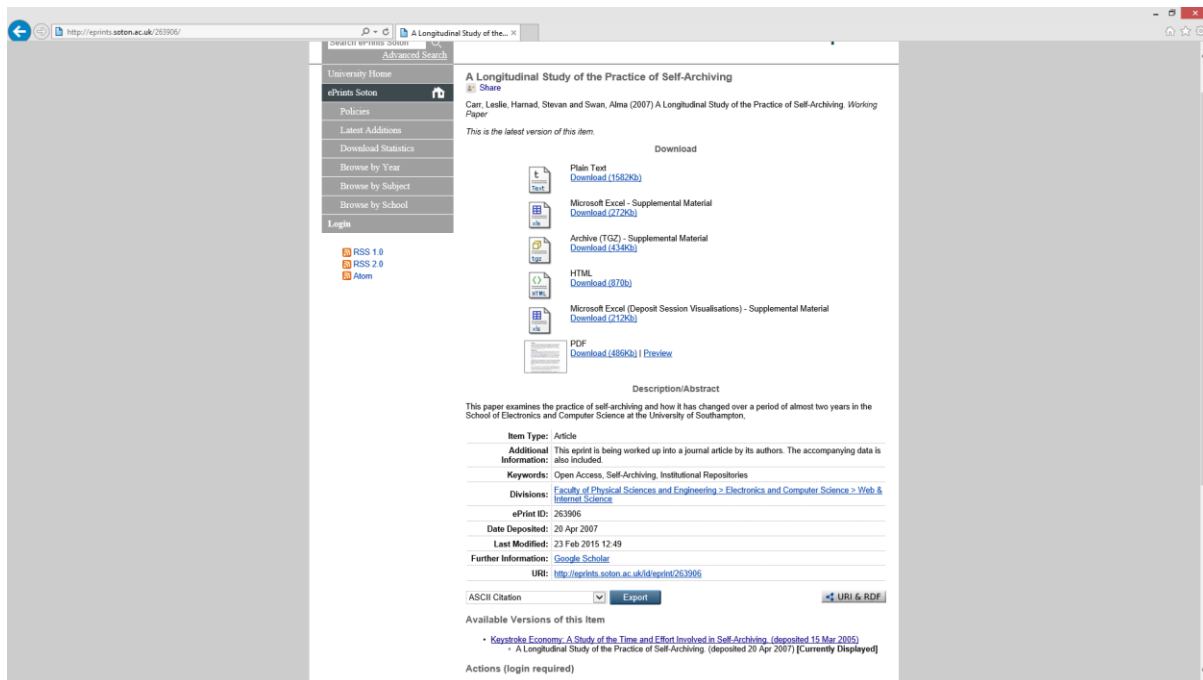
Image 3.

When I hit the eprint Button in order to obtain a copy I received the response below:

> *Please note that this request goes directly to the author. We try to make an open access copy available where possible. However there is no guarantee that there is an electronic copy which can be distributed, or that the author is currently available to respond to your request. You may need to contact your local library for an inter-library loan.*
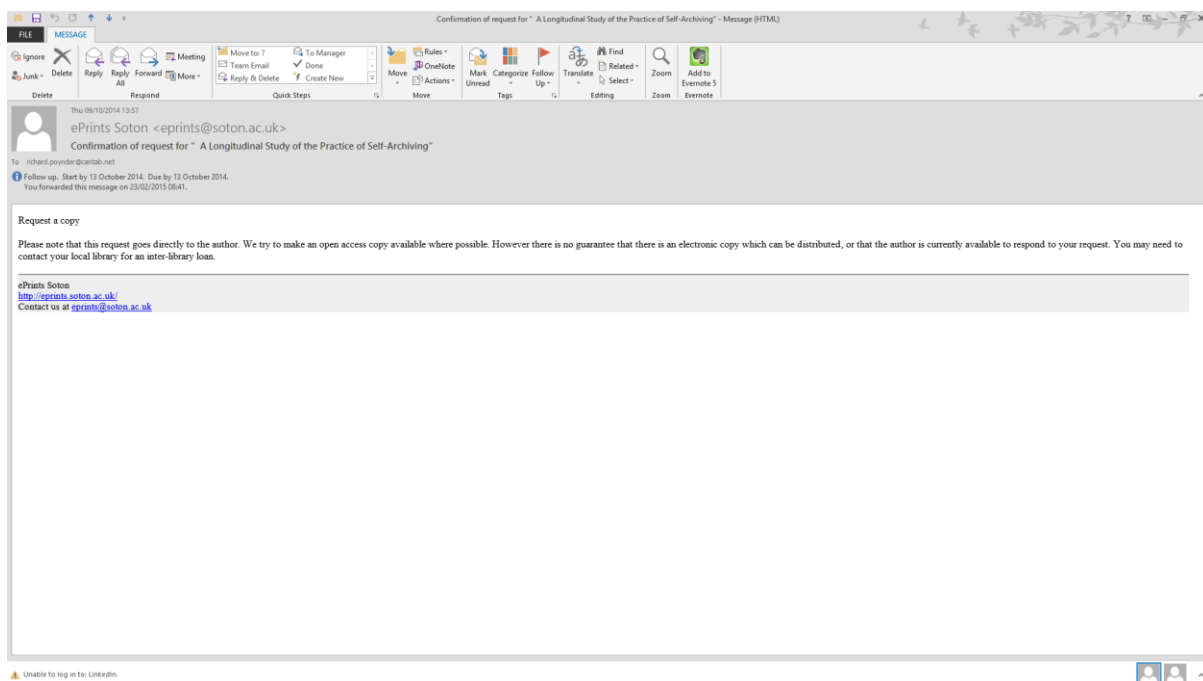

Image 4.

So far so good, I thought, and sat back and waited to receive the paper, or at least to receive some communication from the authors or the repository manager. But my waiting was in vain. To my surprise, I not

only did not receive the paper but I received no explanation as to whether or when I could expect to receive it. My request just floated off into cyberspace, with no more than the message above as an echo.

I was also surprised that the paper was behind a login wall in the first place. Not only had it apparently been written 7 years previously, but the named authors are three OA luminaries: Stevan Harnad (who came up with the idea of the eprint Button), Leslie Carr, technical director of the EPrints Repository software team (which developed the eprint Button), and Alma Swan, at that time director of European Advocacy at SPARC, the most high-profile and proactive OA advocacy organisation in the world.

However, my surprise did not end there. When I went back to the record I saw a note indicating that there was an "available" version of the paper elsewhere in the repository. But when I followed the link I was presented with a paper called *Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving* (below). This paper was dated 2005, and listed two of the authors of the paper I had requested (Harnad and Carr), but not Swan. It seemed clear to me, therefore, that it was not the same paper.
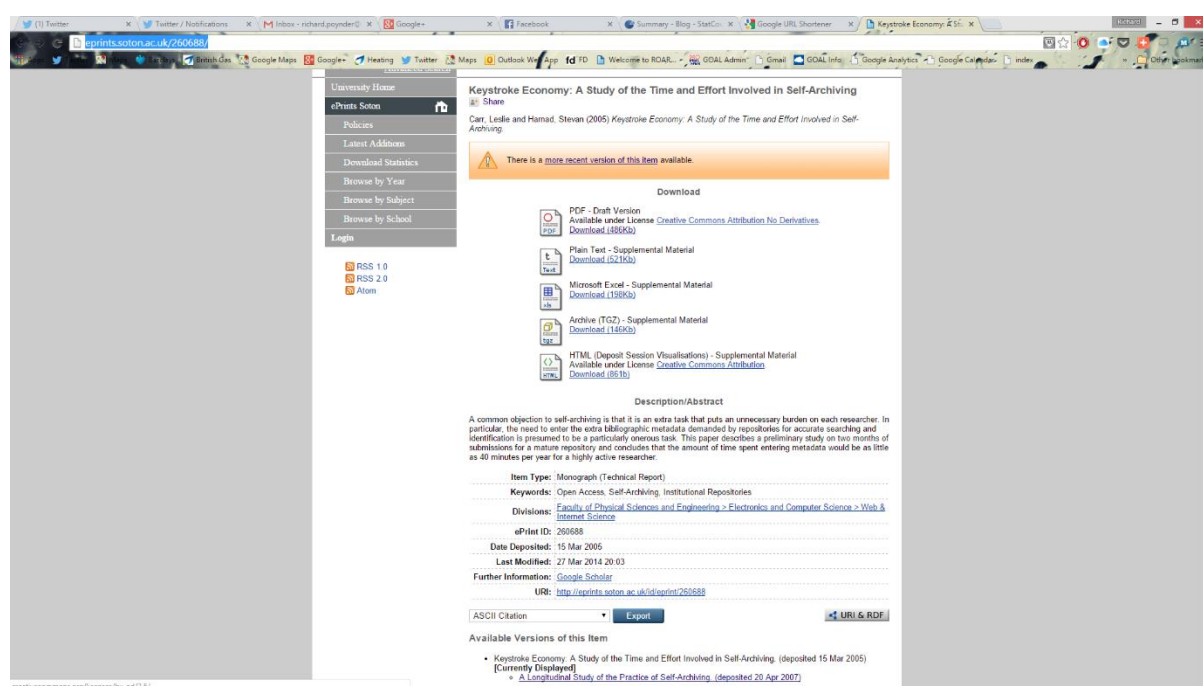

Image 5.

So I waited and waited for four months, until February 2015. At that point I emailed all three authors to ask if they had received my request. Harnad replied that he had not, but that Button requests only go to one of the authors (a less than optimum approach perhaps). In this case the Button request should have been emailed to Carr.

Harnad explained further: "There are two pieces of research. The one a 2005 (unpublished) paper by Les and me and the other a 2007 data-set and analyses by Les, from which we had planned to write a paper, but did not. It is listed as a 'working paper,' which is exactly what it was. Neither of these papers are published articles, let alone peer reviewed."

In other words, although it had been (and remains) indexed in the repository, the 3-author paper I had requested had never been written. This was puzzling not least because all three authors had added the unwritten paper to the publication lists on their web sites. And the paper was still listed by all three authors at the time of writing, many months after I emailed them all, and after Harnad had reported that the paper had never been written (see here, here and here).

Essentially, eight months after I contacted them all three authors were still publicly claiming to have written a paper that does not exist (and that was still the case the last time I checked).

Further complicating the situation, when I had contacted the authors Harnad had responded by attaching the earlier paper (*Keynote Economy*) to both records in the Southampton repository, even though the two records have different names (here and here), and even though Swan is not an author of the paper that had been written.

To be clear, there is no suggestion here of anything untoward. Rather I think my experience demonstrates two things. First, even leading OA advocates do not necessarily respond to Button requests for copies of their papers.

Second, publishers would appear to be right when they argue that self-archiving creates a certain degree of chaos in the scholarly record, not least because there are likely to be any number of different versions of any paper floating around in cyberspace, including unpublished and non-final versions.

Moreover, repositories appear to do too little to flag whether the papers they host have been peer reviewed, or officially published. I could see no indication on the repository record of the *Keystroke Economy* paper, for instance, indicating that it had not been peer reviewed.[8] For this reason, perhaps, the paper has been widely cited as though it were a peer-reviewed paper.

My experience also reveals a further problem in the way the Button works. As noted, my request for a copy of the paper in the Southampton repository had simply disappeared into the blue, and I received no response from any of the authors until I emailed them. Evidently the Button functionality is not designed to provide feedback to users. So how long should one wait before giving up?

Nevertheless, I was still left wondering why the very person who had developed the Button had not responded to a copy request made by means of it. When I contacted the authors in February 2015 Harnad had replied, "My guess is that the fact that the paper in question here never actually got written is the best explanation of why the eprint request was not fulfilled by Les."

I did not find this explanation entirely convincing. After all, if that was the reason, then why did Carr not email me to say as much? My scepticism appeared to be vindicated when Carr later replied to my earlier email confirming that he had received my request on 9th October 2014. While he did not say as much, he seemed to imply that he had simply ignored it. When I asked him for confirmation of this, he did not answer directly, but in a somewhat cryptic response suggested that eprint requests can easily get lost in a researcher's "email stream".

Intrigued by my experiences I decided to test another repository, requesting a copy of a paper on dark deposit in RepositoriUM, the repository of the University of Minho in Portugal. The paper was called *Academic job satisfaction and motivation: findings from a nationwide study in Portuguese higher education* and had been published by Taylor & Francis in 2014. Those without subscription access can purchase the paper from the publisher for £25. Rather than do that, I hit the request eprint Button in the UMinho repository.

In response I received the message below, but nothing further. I did not receive the paper; nor did I receive any message from the author or anyone at the University explaining what had happened to my request. Did it get lost in an email stream, was the email system (as the one at the University of Ljubljana apparently is) broken, or was there some other reason for my lack of success? I do not know.

---

[8] While the record for the *Keynote Economy* paper now points to the record for the *Longitudinal Study* paper (flagging it as "a more recent version"), the record for the *Longitudinal Study* is in fact now also linked to the *Keystroke Economy* paper — the user gets stuck in a loop.
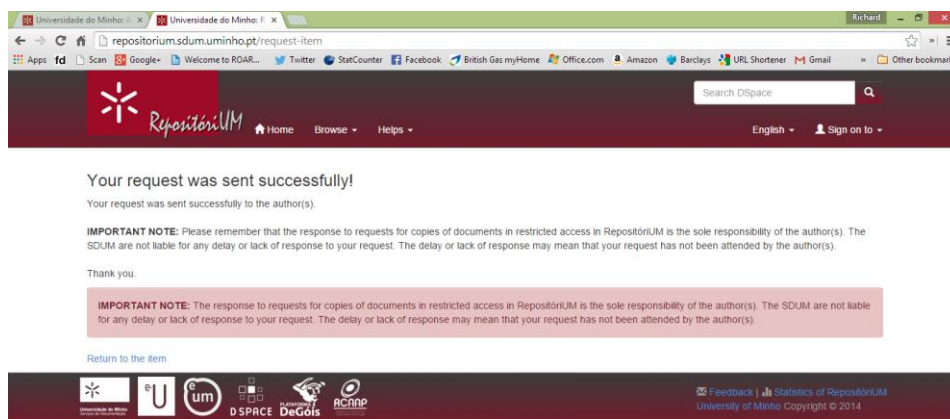
Image 6.

A few months later, I undertook a Q&A with the director of the University of Minho's Documentation Services Eloy Rodrigues. Rodrigues is responsible for the UMinho repository, and another leading advocate for green OA. He is also an advocate for the Button, and in fact developed a version of it for the DSpace repository software. I put it to him that after several failed Button requests most users would likely stop using it. He replied, "Yes, I recognise that. It is not very user friendly, and people may be inclined to give up after a couple of 'non-answers'."

In passing, we could note that Rodrigues reported to me that the percentage of full-text content freely available in the UMinho repository is higher than most, at 79%. By contrast, the PASTEUR4OA study mentioned earlier indicates that only 39%[9] of recent papers in RepositoriUM are open access, with 23% on dark deposit (and 37% not deposited at all).

Again, the difference is doubtless explained by the fact that PASTEURO4OA used bibliographic data from Web of Knowledge (WoK)-indexed journal articles published in 2011-2013. As such, it sought out only papers (not other types of documents), and only recently published research.

Copy request fatigue

By now I had become a bit of a sceptic, not just about the quantity of content in open repositories that is actually freely available, but about the ability of the request eprint Button to liberate dark content.

Like Graf's, my experience is only anecdotal. So is there any more objective evidence of the efficacy of the Button? Sort of, yes. Last year Graf pointed me to a 2010 study undertaken by a group of open access advocates (including Harnad and Rodrigues).

While by no means an in-depth study (reporting as it does on the use of the Button in just three repositories) the findings would seem to support Graf's and my experiences. Approval rates [i.e. when authors responded positively to a copy request] in these institutions ranged from 27% (University of Minho), through 47% (University of Southampton) to 60% (University of Stirling).

But here is the puzzle: if researchers are (as OA advocates persistently maintain) desperate to have as many people read and cite their work as possible, why are copy request approval rates not much higher?

---

[9] Interestingly, this is higher than the ORBi figure.

We have discussed some possible reasons, including the possibility that the requests never reach the authors, or get lost in their email streams. A further possibility is that researchers quickly develop what we might want to call "copy request fatigue".

This is an explanation Rodrigues himself proposed in a blog comment he posted in 2010. As he put it, "Our experience is that authors get 'tired' of replying to copy requests, especially when requests are very frequent. The consequence is that some start not replying at all, and others ask to change to Open Access articles/ papers/ theses that were in closed/ embargoed access. We had more than 20 of those requests just on the last year…."

When I interviewed Rodrigues in March 2015 he suggested yet a further possibility: researchers, he said, worry about the consequences of self-archiving (they fear, for instance, that using the Button might infringe copyright, or that it will upset the publisher, to whom they will invariably have assigned copyright). While some are confident and courageous about making their work freely available, said Rodrigues, others are fearful, especially when their papers have been published in journals or conference proceedings that do not have well formulated self-archiving/OA policies. However, he added, this anxiety reduces over time.

The problem is that even if researchers become braver over time, and begin to ask for their papers to be made open access, repository managers cannot oblige them if the publisher has placed an embargo on their work, or at least they cannot do so until the expiration of that embargo.

What about the copyright concerns? Are they justified? Harnad and fellow OA advocates insist they are not, arguing that use of the Button is permitted under the fair use/fair dealing rules of copyright law. However, not everyone agrees. A researcher I spoke to at a major US university, for instance, told me recently that the Button has not been implemented in his repository because the university lawyer "thinks it is unlawful".

And is there any evidence to support Rodrigues' belief that researchers become braver over time? Quite the opposite, it seems. Approval rates at Minho have fallen further since 2010, from 27% to just 23% four years later. "In 2014 we had a global response rate of around 23%, with 21% sending the requested documents and 2% denying the request", Rodrigues told me.

I don't know whether Southampton and Stirling have experienced a similar drop, but if copy request fatigue is a real phenomenon it might seem likely.

Rodrigues believes that there is a simple explanation for the falling rates at Minho. The majority of failed requests, he says, are for theses and dissertations (T&Ds) not papers. And this, he says, is because the repository loses touch with the authors of such documents when they leave the institution. "Because most authors of T&Ds don't maintain any connection with the university after completing their thesis and dissertation, and they often change the email that was registered at the time the document was deposited in the repository (which is the email used to send the requests to authors), the T&Ds response rate is very low (probably lower than 10%), and that obviously affects the global response rate."

In other words, Rodrigues believes that the global response rate at Minho is lower than it ought to be because the response rates for T&Ds are disproportionately low, and so pull the overall rates down. However, we should note that most researchers move around a great deal these days, so repositories must inevitably be losing contact with the authors of work they host on a continuous basis. In any case, Rodrigues concedes, his theory can only be speculation because repository request logs do not record the type of document being sought. To confirm his suspicion, he says, it would be necessary to dig into the request logs 'manually', something he has not done. Once again, we are hampered by a shortage of data.

It is striking that so little data has been collected on the usage and effectiveness of the Button. When I asked Rodrigues why this is so, he replied: "At the time of development we really didn't consider the issues around monitoring, reporting, collecting statistics on the use of the Button, or providing feedback to requesters. And after the initial development we have really just made some minor improvements/ adjustments (like spam control through a captcha feature) and upgraded it to the newest DSpace releases."

Given the importance that green OA advocates attach to the Button I find this surprising. It is, after all, promoted as a vital, and core, part of the OA infrastructure.

We could also note that the concept of copy request fatigue would seem to challenge the claims that researchers are desperate for readers. Certainly they do not appear to be as motivated to maximise the number of people who read their papers as OA advocates claim.

What researchers clearly do want is for Search, and Promotion & Tenure committees to conclude that their work is of high quality. After all, unlike readers, the decisions of these committees have a direct impact on an author's job and career prospects. And what these committees inevitably look for is evidence of where a researcher's work has been published, not how many people have read it.[10]

This in turn means that researchers care more about what journal they publish in than their readership numbers. In practice, therefore, it is the prestige of a journal (measured by its impact factor) that counts, even though the IF is only a proxy measure of quality and designed to assess the value of a journal, not the articles published in it. In fact, since it tells us little or nothing about individual articles (and thus authors) the IF has been widely discredited.[11] Even so, researchers continue to attach huge importance to it.

We could also note that while open access advocates insist that making research freely available increases both readership and citations, the jury on this is still out. A study by Research Information Network (RIN) last year, for instance, concluded, "Overall, articles published OA appear to show a higher number of citations, though the effect is small, and the data provided does not allow us to control for possible confounding effects."

RIN added: "Similarly, any effect of OA on the timing of citations appears to be small, and we have not been able to control for possible changes such as increased awareness of the journal on the part of both readers and authors."

RIN does say, however: "[A]lthough the impact on citations is small, the impact of open access publication on HTML views and PDF downloads is large and significant, suggesting increased visibility for the open access papers."

But as we have suggested, visibility is not what researchers crave; or let's say that they crave it far less than they crave being published in prestigious journals.

It is no great surprise therefore that even young researchers (who have been reared on the Web and are widely said to be the most receptive to open access) remain somewhat indifferent to OA. A 2015 *eLife* survey of early career researchers found that "only about 25% of respondents felt that they received any individual benefits on publishing in OA journals."

And when asked to list the factors they considered important when selecting a journal in which to publish "open access" came at the bottom of the list.

The point is that as a result of the incentive and evaluation systems currently in use in academia, researchers value prestige more than readership, or even visibility. So the brand value of the journals they publish in is paramount, and being published in a high impact journal is far more important than the number of readers a

---

[10] P&T committees deny this, but we know they continue to use the IF and journal brand as a proxy measure for quality. We also know that researchers hanker after being published in high impact journals in order to enhance their own prestige too. This was well expressed in *The Huffington Post* in May: "You don't get a faculty position at Princeton by publishing something in the Journal Nobody-Ever-Heard-Of," Dr. Oransky said. Is being lead author on a big study published in *Science* "enough to get a position in a prestigious university?" he asked, then answered: 'They don't care how well you taught. They don't care about your peer reviews. They don't care about your collegiality. They care about how many papers you publish in major journals.'"

[11] See also the report, *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*.

researcher's work attracts.[12] Doubtless this helps further explain why researchers may not be overly motived to accept Button requests.

Ironically, the craving to publish in high IF journals makes green OA all the more important, since some of the most prestigious journals (i.e. *Nature* and *Science*) do not offer a gold OA option, or where they do, the cost is so high that most researchers are not able to afford to publish in them. This means that green OA is often the only way to make papers published in such journals open access.

Reimagining the button

But let's return to the question of the efficacy of the Button. Undoubtedly there is a logic to the device. In fact, it is the logical next iteration of a long-standing custom in which researchers share their papers on a peer-to-peer basis. In the print world if a researcher wanted to read a paper that they did not have access to they would contact the author(s) and ask to be sent a print copy — what was known as an offprint.

This process becomes much easier in the digital world because researchers can now simply email the author(s) and ask them to send over a digital copy, generally referred to as an eprint. The assumed appeal of the copy request Button is that by semi-automating the process the chances of a request being successful are increased. But are the approval rates of the copy request Button higher than standard email requests?

Again, the signs are not that encouraging. In a 2011 study published in PLOS ONE the authors sought to establish how easy it is to get copies of papers by emailing the authors. To this end they sent out 40 requests for 50 of the most recently published papers with the keywords "HIV vaccine". They received 24 positive replies in response (a 60% success rate).

The authors also compiled a second set of 200 articles randomly selected from those cited in the first set of papers. They explain: "For 'cited' papers, we intended to send out 65 requests, but could not locate the email addresses of eight of the corresponding authors. For the remaining 57, requests were sent out and 31 authors responded by sending a copy of the paper (54% success rate). Among the 26 emails that were unsuccessful, 8 email addresses were outdated, one author declined to provide the paper citing copyright issues with the journal and the rest did not respond."

If we set these figures alongside the approval rates Harnad *et al* found in their 2010 study — i.e. 60% at the University of Stirling, 47% at the University of Southampton and 27% at the University of Minho, it is far from clear that the Button offers any obvious advantage.

Again, the PLOS paper is only one study and not an in-depth one at that. Nevertheless, it is hard not to conclude that while the concept of the eprint Button is sound on paper, in practice it is no more effective (and perhaps less effective) than simply emailing the lead author (whose address is routinely available as a clickable link on the abstracts of scholarly papers — e.g. here) and asking them to send a copy. And in doing so, it seems, they can expect to be successful around half the time.

An additional point to make is that while the copy request Button may be able to liberate papers held on dark deposit in repositories, there is no way it can enable access to papers that are behind paywalls like ScienceDirect and Wiley Online Library if they have not been self-archived. This means that email requests also have a wider reach than the Button. As such, the email route would seem to be a better strategy. In addition, using email incurs no software development costs, and does not require repositories to implement the Button.

In short, it is far from clear that the Button can deliver on its promise. This thought inevitably occurred to Harnad *et al* when reviewing the disappointing results of their 2010 study. "Given a significant number of Button requests which are ignored or lost, one might be tempted to assume that it has not worked," they said.

---

[12] Of course if a paper is published in *Nature* or *Science* it will get more readers, but readership is still not the key issue here.

Unwilling to be downcast, however, the authors went on to suggest that this would be the wrong conclusion to reach. "The principal impact of the Button has been to enable the adoption of institutional IDOA mandates," they asserted. "Deposit is mandated immediately without legal constraints, with the Button serving to assist authors interested in the dissemination of their articles."

Here I think we see a subtle shift in focus from reader to author. Instead of ensuring that readers can get access to research papers, the Button now "assists" authors wanting to encourage peers to read their papers. The Button has been reimagined. Yet it is still being promoted as a way of neutralising publisher embargoes. As such, we could view it as a sales tool intended to persuade funders and institutions to buy into "Almost-OA" — an access model that by any definition is second-best. But like many sales tools the Button over-promises, it does not do what it says on the can.

The next logical question to ask is this: is the sleight of hand involved in using the Button to promote the IDOA/PEM mandate justified by the end goal — which is to see a proliferation of such mandates? Or to put it another way, how successful are IDOA/PEM mandates likely to prove? That is a question I want to address in part two, where I also plan to speculate on possible futures for institutional repositories.