

---

This item was submitted to [Loughborough's Research Repository](#) by the author.  
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

## What does “green” open access mean? Tracking twelve years of changes to journal publisher self-archiving policies

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1177/0961000616657406>

PUBLISHER

SAGE Publications (UK and US) © The Author(s)

VERSION

AM (Accepted Manuscript)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Gadd, Elizabeth A., and Denise Troll Covey. 2019. “What Does “green” Open Access Mean? Tracking Twelve Years of Changes to Journal Publisher Self-archiving Policies”. figshare. <https://hdl.handle.net/2134/21555>.

## **What does “green” open access mean? Tracking twelve years of changes to journal publisher self-archiving policies.**

**Corresponding author:** Elizabeth Gadd, Research Policy Manager, Loughborough University, Loughborough, LE11 3TU, UK. Email: [E.A.Gadd@lboro.ac.uk](mailto:E.A.Gadd@lboro.ac.uk). Tel: +44 (0)1509 228594.  
Denise Troll Covey, Emeritus Scholarly Communications Librarian, Carnegie Mellon University, Pittsburgh PA, USA. Email: [Troll@andrew.cmu.edu](mailto:Troll@andrew.cmu.edu).

### **Acknowledgements**

The authors should like to acknowledge the expert guidance of Dr Jenny Fry and Professor Charles Oppenheim who provided helpful feedback on drafts of this paper.

### **Abstract**

Traces the 12-year self-archiving policy journey of the original 107 publishers listed on the SHERPA/RoMEO Publisher Policy Database in 2004, through to 2015. Maps the RoMEO colour codes ('green', 'blue', 'yellow' and 'white') and related restrictions and conditions over time. Finds that while the volume of publishers allowing some form of self-archiving (pre-print, post-print or both) has increased by 12% over the twelve years, the volume of restrictions around *how*, *where*, and *when* self-archiving may take place has increased 119%, 190% and 1000% respectively. A significant positive correlation was found between the increase in self-archiving restrictions and the introduction of Gold paid open access options. Suggests that by conveying only the version of a paper that authors may self-archive, the RoMEO colour codes do not address all the key elements of the Bethesda Definition of Open Access. Compares the number of RoMEO 'green' publishers over time with those meeting the definition for 'redefined green' (allowing embargo-free deposit of the post-print in an institutional repository). Finds that RoMEO 'green' increased by 8% and 'redefined green' decreased by 35% over the 12 years. Concludes that the RoMEO colour codes no longer convey a commitment to green open access as originally intended. Calls for open access advocates, funders, institutions and authors to redefine what 'green' means to better reflect a publisher's commitment to self-archiving.

### **Keywords**

Green open access; self-archiving; publisher policies; institutional repository; embargo periods.

### **Introduction**

The Rights Metadata for Open Archiving (RoMEO) database of journal publisher open access policies (SHERPA, 2016a) was introduced over a decade ago to provide authors with guidance as to whether and how they might self-archive their journal articles to make them available on open access (Jenkins et al, 2007). In that time the database has grown from documenting the self-archiving policies of 107 publishers to 2,192 (as of April 2016). To aid authors in selecting journals that had permissive self-archiving policies, RoMEO colour codes publishers 'yellow', 'blue', 'green' or 'white', based on the version of the paper (pre-print or post-print, both or none) they allow authors to self-archive. Subsequently, 'Gold' was used to describe publishing in an open-access journal and 'Green' became a generic term to describe all author self-archiving, regardless of whether RoMEO coded the publisher policy green, yellow or blue (Harnad et al, 2004). Over time, a publisher's commitment to open access has been characterised by whether they are gold, green, blue, yellow or white. To avoid

confusion, throughout this paper Gold refers to publishing in an open access journal; Green refers to self-archiving an open access copy. Lower case 'green' refers to the RoMEO colour code for a publisher's self-archiving policy.

The RoMEO colour codes only refer to the version of a paper authors or their representatives may (or may not) self-archive immediately; they give no indication of the increasingly complex set of restrictions and conditions that typically surround that basic permission. This paper seeks to explore publishers changing attitudes towards so-called Green open access over a twelve year period. It does so in order to answer the research question: does the green RoMEO colour code still convey what it was initially intended to convey, i.e., full support for open access self-archiving? Or is it the case that some publisher policies give the illusion of full support (sufficient to secure RoMEO code green), but in reality the policies discourage self-archiving in a manner consistent with author preferences for discovery, preservation, reuse and increased impact (Fry et al, 2011; Spezi et al, 2013). We seek to answer this question by analysing over a twelve year period (2004-2015) the policy journey of the original 107 publishers listed on the RoMEO database in 2004, documenting not only the changes to their RoMEO colour codes over that time, but also the increasing range and complexity of restrictions and conditions their policies specify.

## **Background**

Open access to the "royalty-free" literature (Suber, 2003) was advocated as a solution to two related problems facing the world of scholarly communication: access and impact (Harnad, 2008). The access problem was the result of spiralling journal subscription costs, leading to cancelled library subscriptions. Scholars as readers could no longer access all the literature they needed, and consequently could not read and cite relevant papers. This led to the impact problem: scholars as authors have unread and uncited papers, thus reducing their impact. Open access was given a firm foundation in 2002 through the Budapest Open Archive Initiative (BOAI) (2002) Declaration. The Declaration provided the first community-agreed definition of open access and was closely followed by the Berlin declaration and Bethesda definition, both in 2003. The key elements of these definitions were similar (but not identical as discussed below) leading Suber (2004) to refer to the three as the "BBB definition". The common elements were that the results of scientific enquiry should be freely available to anyone with an internet connection, and subject to permission-free scholarly re-use. The definitions were neutral as to whether this goal should be achieved through Green or Gold open access. However, at the time the development of Gold and hybrid open access journals was in its infancy. For authors who wished to make their papers available via open access, the only options were to post their open access copies to personal or departmental websites from which their work disappeared when they left the institution.

From 2005, institutions and organisations began in earnest to establish repositories to address the problem of the discoverability and preservation of self-archived open access copies of papers. The Directory of Open Access Repositories (OpenDOAR, 2016) listed 128 repositories extant in December 2005 and 3,018 just over a decade later in 2016. Despite the growth in the number of repositories, the growth in the number of deposits of open access papers has been slower. Slow initial growth in the number of repository deposits improved considerably when institutions and funders began to mandate open access (Xia, J, et al 2012; Harnad, 2011; ROARMap, 2016), most of which mandates

championed Green open access over Gold (Archambault et al, 2014; SPARC, 2016). The Science-Metrix report into *The proportion of open access papers in peer-reviewed journals* (Archambault et al, 2014) puts most of the early growth of self-archived papers down to “background growth” in scholarly publishing. Growth of self-archived papers plateaus from 2004 onwards. Between 1996 and 2012 the authors of the report estimate the growth rate of Green open access to be 8.8%. By contrast Gold open access has shown a growth rate of 24% over the same period. Laakso (2014) and Troll Covey (2009) have both demonstrated the large gap between the potential for self-archiving permitted by publisher policies and the actual self-archiving practice by scholars. Bjork et al (2014) estimated “realized [G]reen open access” at about 12% of total annual scholarly output which is a fraction of the 81% Laakso (2014) estimated was actually possible after embargos of up to 12 months.

A number of user studies such as those by Swan and Brown (2004), Nicholas, Huntingdon and Rowlands (2005) and the PEER study (Fry et al, 2011) shed some light on the reasons for this slow adoption of self-archiving. The most recent study from the PEER project showed that “researchers who associated Open Access with ‘self-archiving’ were in the minority” and those that did so were more likely to belong to the Physical sciences & mathematics and the Social sciences, humanities & arts, than those in the Life sciences and the Medical sciences who are more likely to associate Open Access with Open Access Journals (Gold Road)”. They deduce that authors prefer Gold because their first choice is to read and cite the ‘version of record’ complete with publisher formatting and pagination. They also found anecdotal evidence that authors thought making journal articles available via open access was “beyond their remit” and that “checking publishers’ OA [open access] policies is the main barrier to authors managing the self-archiving process themselves”. As a consequence, the task now usually falls to repository managers within institutions. An investigation by Research Consulting (2014) into the cost of complying with the UK funders’ open access mandates estimated that the time taken to ‘triage’ each Green open access paper (including checking publisher policies) was 20 minutes at a cost of £12. They concluded that “efforts to cut the costs of Green OA are perhaps best focussed on two areas: making the deposit process as quick and easy for authors as possible; and working to achieve greater clarity in publisher policies.” These are indeed worthwhile goals. However, authors who do self-archive often do so in breach of publisher policy, either unknowingly, through ignorance of the policy, or knowingly, in the belief that publisher reliance on their free gift of content will deter them from pursuing the breach (Troll Covey, 2009).

The RoMEO database of publisher open access policies was created to assist authors with the challenge of discovering and interpreting publisher policy. The original publisher listing was an outcome of the Joint Information Systems Committee (JISC) funded Rights Metadata in Open Archiving (RoMEO) project’s analysis of 80 publisher copyright transfer agreements (Gadd et al, 2003). At the end of the project, maintenance of the listing was taken on by the JISC-funded Securing a Hybrid Environment for Research Preservation and Access (SHERPA) project and developed as a service (SHERPA, 2016a). The database was recently assigned a “4/5” rating in the *Technical Services Quarterly* (Sullivan, 2014), though in a recent survey of the usage of scholarly communication tools, only 9% of 11,250 respondents said they used the RoMEO database to select a journal to publish in (Kramer & Bosnen, 2016).

One of the earliest activities undertaken by the SHERPA project team was the creation of a controlled vocabulary to define the various publisher positions and (in theory) to aid the future analysis of publishers' open access positions (Jenkins et al, 2008). The controlled vocabulary took the Bethesda definition of open access (Suber et al, 2003) as the foundation for analysing publisher agreements and allocated each publisher a colour according to their adherence to that definition (Jenkins et al, 2007). The key elements of the Bethesda definition of open access are that: i) a complete version of the work is archived, ii) with all supplementary materials, and iii) deposited immediately in iv) at least one online repository that is v) supported by an academic institution, scholarly society, government agency, or other well-established organization that vi) seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving. Furthermore, the terms of use should vii) "allow all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use (Suber et al, 2003)."

Whilst the RoMEO controlled vocabulary made use of many of the elements of the Bethesda definition in order to describe a publisher's open access policy, the colour codes made use of only one: the version of the work. Perhaps for simplicity, only two versions of scholarly papers were taken into consideration by the colour codes, namely, the "pre-print" and the "post-print" as defined in Figure 1. A few years later, the NISO/ALPSP Journal Article Versions (JAV) Technical Working Group (2008) defined seven different versions of "scholarly journal articles that typically appear online before, during, and after formal journal publication". One of the key omissions from the versions addressed by the colour codes is the 'publisher version'. Indeed Bjork et al (2014) found that authors most commonly uploaded the "accepted manuscript [post-print] and published articles in almost equal proportions, whereas submitted manuscripts [pre-prints] constitute a smaller proportion of [G]reen OA copies." Arguably, the 'greenest' publishers are those that allow self-archiving not only of the author's accepted manuscript, but of the fully formatted and paginated publisher PDF. Perhaps to make up for this, the RoMEO database does now provide a separate listing of publishers that allow the self-archiving of the publisher PDF (SHERPA, 2016a).

RoMEO Colour	Archiving policy
<a href="#">Green</a>	Can archive pre-print and post-print or publisher's version/PDF
<a href="#">Blue</a>	Can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
<a href="#">Yellow</a>	Can archive pre-print (ie pre-refereeing)
<a href="#">White</a>	Archiving not formally supported

Figure 1 SHERPA/RoMEO colour codes and their interpretation. (SHERPA, 2016a)

The other aspect of the Bethesda definition that the RoMEO controlled vocabulary took into consideration was the requirement that papers should be deposited immediately to qualify. As Jenkins et al (2008) put it, "embargoes are not compatible with the spirit and meaning of OA." They explain, "if a publisher allows the self-archiving of pre-prints, but also stipulates that the post-print can only be self-archived in a year's time, i.e., after a one year embargo, then that publisher is

awarded the 'yellow' colour, but not the 'blue' or resulting 'green' colour." (Jenkins et al, 2007). In an undated document explaining the RoMEO colours, SHERPA Manager Bill Hubbard makes a similar point: "Since an embargo effectively removes the right of an author to archive their work, even if temporarily, the publisher cannot be categorised as green or blue or yellow." The increase in the use of embargoes is one of the issues cited by the Science-Metrix report as a factor slowing down the growth in Green open access papers (Archambault et al, 2014). It is therefore a critical component of any publisher open access policy colour coding.

There are two other key characteristics of the various definitions of open access that the colour codes never set out to address. The first, specified by the Bethesda definition and repeated by the Berlin declaration, is the requirement to deposit in "at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organization". The BOAI agreement speaks only of the need for "free availability on the public internet" (BOAI, 2002). At the outset of open access developments, freeing the literature was the primary concern and the importance of discoverability and preservation were secondary issues. As the area has matured, the importance of these latter issues has become well understood, and the part played by institutional and subject repositories in ensuring long-term preservation and discovery is agreed to be critical (Bailey, 2001; Harnad, 2015). Bjork et al (2014) found that 80% of papers in institutional repositories and 95% of those in subject repositories were still available after three years. Availability of papers on personal web pages after three years was much lower. The original RoMEO controlled vocabulary referred only to two different locations of deposit: a web page, or a non-commercial server. As will be seen in the Results section of this paper, the specificity of locations in publisher policies recorded in the RoMEO database increased dramatically over time.

The second key characteristic of the BBB definition of open access, and one highlighted by Suber (2004) as often overlooked, was the importance of liberal re-use rights. In fact, Suber (2004) coined the term *libre open access* to distinguish free-to-re-use papers from those that are only *gratis open access*: free to read. The BOAI (2002) defined the need for "users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose". The Bethesda Statement (2003), adopted as the foundation for the RoMEO controlled vocabulary, went one step further and included the making and distributing of derivative works. Unfortunately, a publisher's position on re-use rights is not covered by the RoMEO database unless a publisher has specified restrictions on re-use as part of its self-archiving policy. Perhaps in an effort to address this need, the Scholarly Publishing and Academic Resources Coalition (SPARC) and the Public Library of Science (PLOS) have created an Open Access Spectrum Evaluation Tool (SPARC, 2016) which has scored the openness of 1,000 open and non-open access journals' policies according to various factors, including re-use rights. Only 301 of the journals listed (30%) are classed as having 'fully open' re-use rights as defined by the SPARC *How open is it?* Guide (SPARC, 2014).

Issues around the 'when' and 'where' of deposit and the re-use rights permitted by publisher open access policies are becoming even more significant with the introduction of funder open access mandates. RoMEO's sister database, JULIET (SHERPA, 2016b), shows that over half of the world's funders indexed in the database require or encourage open access deposit. The Science Metrix report found that "most [funders] favoured Green (40%), while 2% mandated both Green OA and

Gold OA, and 58% expressed no preference.... 80% of policies specified that articles should be deposited in their final accepted version or post-print (Archambault, 2014).” An analysis by SPARC found that only 20 of 79 funder policies (25%) indexed preferred Gold to Green (Proudman, 2015). The Registry of Open Access Repository Mandates and Policies (ROARMap, 2016) classify funder mandates using 19 fields. Not all funder mandates are complex enough to require all 19 fields; however, re-use specifications, embargo specifications, repository type, version type and who can deposit are fields common to most mandates. In the UK, the introduction of the Higher Education Funding Council for England’s open access policy (HEFCE, 2015) is likely to have the single biggest effect on academics’ self-archiving practices in UK universities. In addition to requirements around the version of the paper (the author’s peer-reviewed accepted manuscript), this policy specifies requirements around the when (acceptable embargo periods), the where (deposit in an institutional or subject repository), and the re-use rights they expect papers to have met in order to be submitted as part of the next Research Excellence Framework (REF) exercise which allocates quality-related research funding to UK universities. (HEFCE, 2014)

To summarize, the RoMEO colour codes address and convey the version of a paper that can be self-archived immediately, but they do not address or convey other important aspects of Green OA, such as where a paper may be self-archived immediately or with what re-use rights. Despite this, the RoMEO colours are perceived as a grading system for a publisher’s commitment to open access. Publisher policies coded green win favour from authors and funders alike, and some scientific studies use a journal’s RoMEO colour code as the sole indicator of its openness to self-archiving (see, e.g., Miguel et al, 2011). One of the main reasons for this is the simple hierarchy of the colours chosen and the language used when describing them. Early discussions around green open access describe it as “giving the green light” to author self-archiving (Harnad et al, 2004). Some have referred to “solid”, “dotted” and “pale” green as variations (Bailey, 2001). Blue and yellow are considered to be component, but lesser, parts of green (Jenkins et al, 2007). White is colourless and to be avoided. The early documentation around the colour codes talks about ‘awarding’ (Jenkins et al, 2007) or “assigning” (Hubbard, [n.d.]) publishers a colour as some kind of badge of honour. Sullivan’s (2007) review of the RoMEO service summarises the perception well when she states: “[g]reen is the most desirable...and white is the most restrictive.”

It is therefore in publishers’ interest to be seen as a ‘RoMEO green’ publisher or at least to allow some form of self-archiving. Not surprisingly, the overall percentage of all publishers on the RoMEO database allowing some type of self-archiving is increasing, and the percentage of white publishers is decreasing (see Figure 2). It should be noted that the percentage of green publishers has been steadily increasing since 2011 but has never returned to the peak of 2005 which was perhaps characterised by an influx of early adopters. Though this situation looks promising for authors and funders who seek to increase the accessibility of the journal literature, the question remains as to whether what is allowed really serves the interests of open access as originally defined or the practices preferred by authors and repository managers or mandated by funders and institutions.

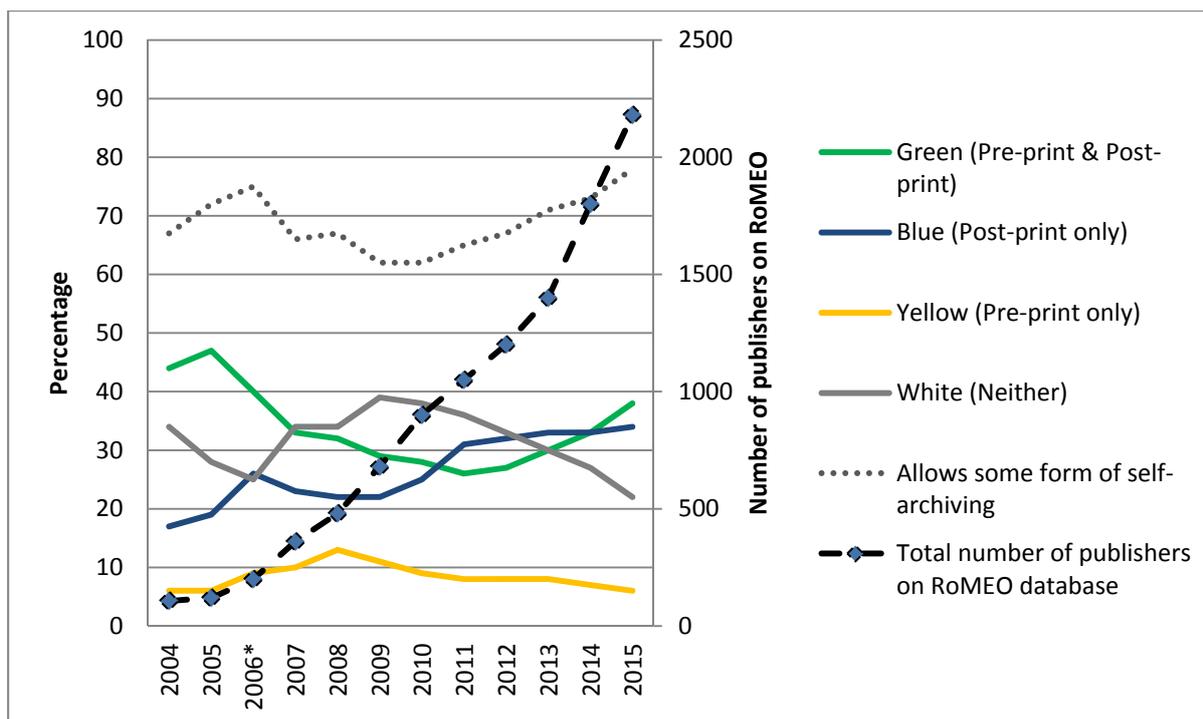


Figure 2. Change in the percentage of *all* publisher open access policies RoMEO colour code over time.  
 \* No data was available for 2006 so the 2006 data points were taken from January 2007.

This paper examines the changes in publisher open access policies over a period of twelve years, since the inception of the RoMEO journal publisher policy database, to try to discern publishers' commitment to the self-archiving of open access copies in a manner consistent with institutional and funder mandates and author preferences for discovery, preservation, reuse and increased impact. It does this by tracing the policies of the original 107 publishers listed on the database in 2004, plotting not only their RoMEO colour codes, but also any restrictions and conditions placed on author self-archiving over time. A full list of the publishers included is given in Appendix 1.

## Methodology

The only publicly available source of historical publisher open access policy information is the archives of the RoMEO database created by the Internet Archive "Wayback Machine" (Anon, 2016). The Internet Archive is a non-profit organisation founded in 1996 to build an Internet library. As such, it periodically crawls for updates to web pages. There are no documented selection criteria for, or pattern to, the crawling activity so inclusion in the Archive is not guaranteed. The RoMEO database provides both colour codes and policy summaries (listing the restrictions and conditions) for each journal publisher. Until 2010, these both appeared on a single page A-Z listing. However, from 2010 onward, the single page A-Z listing only provided the colour codes and a link to a policy summary page for each individual publisher. The Internet Archive was searched for archived full A-Z listings from the <http://www.sherpa.ac.uk/romeo> website. At least one full list was available for each year from 2004-2015 giving the overall RoMEO colour codes. However, the changes in 2010 meant that no policy summary data was available for that year and was very patchy during 2011 and 2012. Where no summary was available for a particular publisher in a particular year, the last known policy summary for that publisher was carried forward. Of the 1177 publisher policy summaries

(from 107 publishers over the 11 years available), 118 (10%) were missing, mainly from 2011 and 2012. In 39 of these cases, the policy summaries pre- and post- the missing years were identical or revealed only one change (e.g., the introduction of paid open access). This reduces the unknown policy summary information to 6.7% and increases confidence that no significant policy changes were missed in the intervening period. Nevertheless, the results presented here slightly under-report the volume of policy changes rather than over-report them. During the period studied, five of the original 107 publishers merged with, or were acquired by, larger publishers. Whether or not they adopted the policy position of the new publisher they all retained their own entry on the RoMEO database, the contents of which were used for this study. Fortunately, none of the original 107 publishers ceased trading during the studied period.

The reason for studying the original 107 publishers on the RoMEO database rather than taking a sample of publishers from each year was because the original set of publishers is the only group for which twelve years-worth of data was available, enabling us to trace changes to their policies over a significant period of time. The other benefit of this method was that, unlike other studies (e.g. Miguel et al (2011) and Laakso (2014)) which only examined the policies of larger publishers (i.e., mainly commercial publishers based in the UK or USA), this group represented a range of publisher types, sizes and nationalities.

Investigations were undertaken to assess how representative the 107 publishers were in terms of size, nationality, and publisher type. Using the *Ulrichs Global Serials Directory*, the number of current active titles for each of the 107 publishers was determined in February 2015. Where a publisher had merged with another and became an imprint of the larger publisher, the number of titles with the imprint was counted. Where a publisher had merged and no record remained of its former title list, the number of titles belonging to the new host was counted. If a publisher wasn't listed on Ulrichs, a web search was performed to determine the number of titles owned by that publisher. Clearly this does not represent the total number of journals each publisher had during the ten year period, but serves as an indication of size for the purposes of analysis.

As Ware and Mabe (2009) have noted, journals are not evenly distributed across publishers. They write, "at one end of the scale, 95% or more publish only one or two journals, while at the other end, the top 100 publish 67% of all journals. The top 5 publish nearly 35% of journals, while three publishers (Elsevier, Springer, and Wiley-Blackwell) have well over 2000 journals each." The 107 publishers studied here publish a total of 20,815 titles – 31% of the 66,951 active, scholarly, peer-reviewed titles that Ulrichs listed in February 2015. Figure 3 illustrates the spread of titles across the publishers in this study using a logarithmic scale. It follows the expected skewed distribution, however does not have the long tail of publishers with one to two titles each. In our sample of 107 publishers, the top 5 publish 53% of the 20,815 titles, the top 10 publish 74% of the titles; 11% publish only one or two titles. The data suggest that the early adopters of open access self-archiving policies were more likely to be or to become rather prolific publishers.

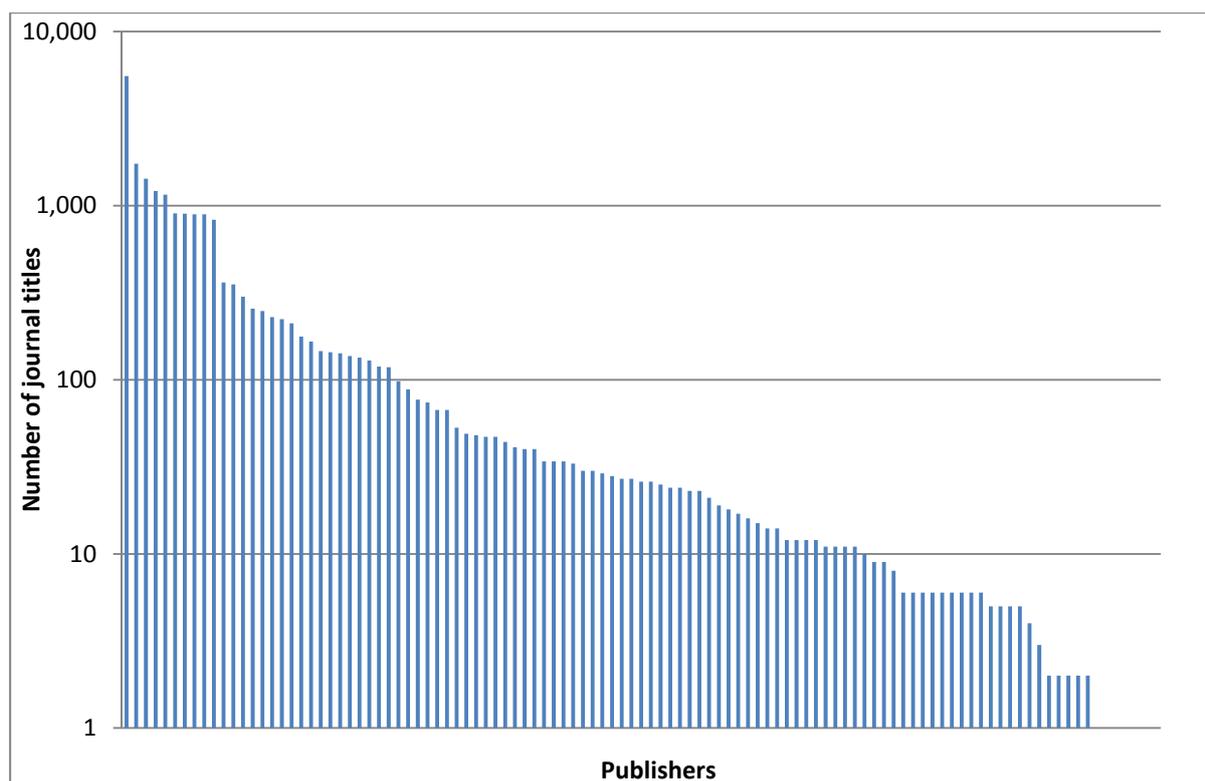


Figure 3 Distribution of journal titles across the 107 publishers studied.

Publishers were categorised according to whether they were commercial, learned society, university presses or 'other'. Table 1 indicates the distribution of titles by publisher type and compares the percentage of title distribution with Ware and Mabe's analysis (2009) of title distribution in the Thomson-Reuters Journal Citation database. Unfortunately the number of publishers is not provided by Ware and Mabe's study. However, it can be seen that the title distribution among the publishers studied here is different. The learned societies in our sample published only 10% of the titles covered, although they represented over half of the publisher sample. As Ware and Mabe noted, 95% of publishers only publish one or two titles and are likely to be learned society publishers. As many studies to date have only focussed on the larger journal publishers, the inclusion of so many learned societies in this study sample provides a more thorough representation of publisher self-archiving policies.

	This study				Thomson-Reuters
	# pubs	% of pubs	# titles	% of titles	% of titles
Commercial	23	21.5%	14,997	72.0%	64.0%
University Press	17	15.9%	2,421	11.6%	4.0%
Learned society	62	57.9%	2,081	10.0%	30.0%
Other	5	4.7%	1,316	6.3%	2.0%

Table 1 Distribution of 107 publishers across publisher types

The geographical locations of the 107 publishers are listed in Table 2 alongside the proportion of publishers in the current RoMEO database with the same nationality. This study categorised three publishers as multinationals (Elsevier, Sage and Taylor and Francis), although the RoMEO database

lists these as US, UK and UK publishers respectively). The majority of the remainder are based in the USA (57%) and the UK (30%). However seven other countries were represented which provides an international perspective. According to the April 2016 RoMEO statistics, the USA accounts for 24% of the publishers listed, and the UK 12% (SHERPA, 2016a), which means these countries are somewhat over-represented in this study. Forty-seven per cent of publishers currently on the RoMEO database are from countries not represented by the original 107, indicating that the international coverage of the RoMEO database has increased dramatically since its inception in 2004.

Country	This study		RoMEO database
	# publishers	% publishers	% publishers
Multinational	3	3%	NA
USA	61	57%	24%
UK	32	30%	12%
Canada	3	3%	3%
Netherlands	2	2%	1%
Australia	2	2%	2%
Germany	2	2%	4%
Japan	1	1%	1%
India	1	1%	5%
Sweden	1	1%	1%

*Table 2 Countries in which the 107 publishers were based*

The sample of 107 publishers in this study is not representative of all publishers when compared to the data sources we have. However, the preliminary analysis of the sample sheds some light on the character of early adopters of open access self-archiving policies.

As Laakso (2014) found, one of the key challenges regarding the use of the RoMEO database for studies of this kind is that, despite the best intentions of the original controlled vocabulary, metadata is not stored in any structured way. However, it is “the most comprehensive service of its kind” and the only publicly available source of historical publisher policy information. Thus, the policy position data extracted from the Internet Archive, was manually coded according to what version of a paper could be archived, where, when, under what conditions or restrictions, and in what context (e.g. whether the publisher offered paid open access options). A list of the policy positions that were coded is provided in Appendix 2. The policy information was coded three times on separate occasions to ensure the highest levels of accuracy and consistency. The dataset will be made freely available on Loughborough University Data Archive (Loughborough University, 2016) once the analysis has been completed.

## Results

### What can be self-archived: changes of RoMEO colours over time

Figure 4 shows the changes in RoMEO colour codes of the original 107 publishers over the twelve year period. In line with the overall trends illustrated in Figure 1, there has been an upward trend in the growth of green publishers from 48 to 52 (a 4% increase) and a decrease in 'white' from 28 to 23, yielding a 12% increase from 2004 to 2015 in the number of publishers allowing some form of self-archiving. To better understand the significance of the RoMEO colour codes, if a publisher was coded 'white' but allowed self-archiving in some form under a restriction or condition which prevented the allocation of another colour, this was coded into the dataset. The number of 'white' publishers allowing some form of self-archiving grew significantly over the period from zero in 2004 to 14 in 2015. The number of publishers either allowing or insisting on the use of the publisher PDF also increased during this period. In 2004, only 17 publishers permitted this (16%), two of which were the Gold open access publishers *Biomed Central* and *Hindawi Publishing*; by 2015, 39 publishers (36%) allowed this, although with increasing numbers of restrictions and conditions as outlined below. By the end of the studied period, 97 of the original 107 publishers (91%) listed on the RoMEO database in 2004 allowed self-archiving of some description.

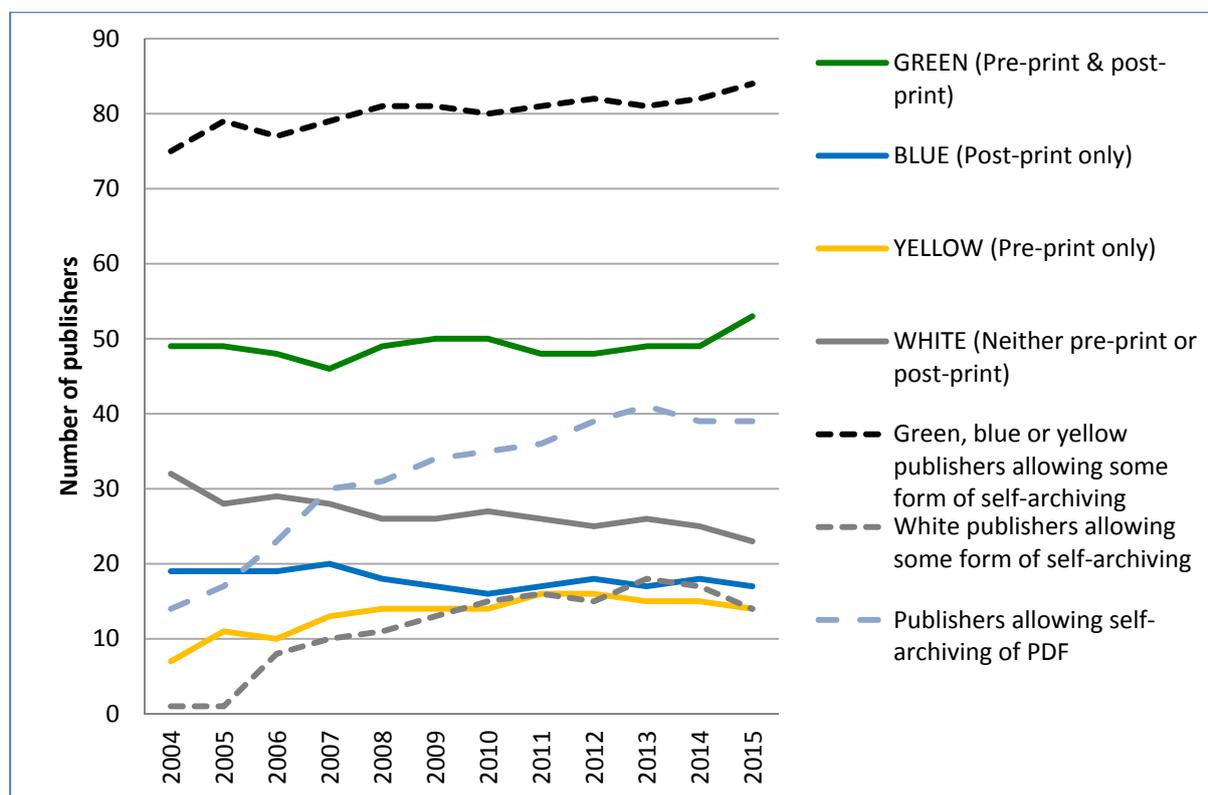


Figure 4 Changes in RoMEO colours for the publishers studied

The total number of publishers that made a change to their policy resulting in a change of RoMEO colour code during the studied period is illustrated in Table 3. It shows that 47.7% of the publishers made at least one change. Fourteen per cent made two or more changes over the twelve years. A more detailed analysis of publisher RoMEO colour code changes over time has been performed by Troll Covey (2013).

# publishers	# changes	% of total
--------------	-----------	------------

56	0	52.3%
36	1	33.6%
12	2	11.2%
3	3	2.8%

Table 3 The number of changes to RoMEO colours for the publishers studied between 2004-2015

### Restrictions & conditions – the where, when and how of self-archiving

The RoMEO colour codes refer to *what* version of a paper can be self-archived immediately (pre-print, post-print, both or neither). Figure 5 illustrates the number of restrictions and conditions listed in the RoMEO publisher policy data relating to the *where*, *when* and *how* of self-archiving. This analysis was done at a publisher level to make it viable, however in some cases (and increasingly over time as will be seen below) publisher policies varied according to the journal. The figure plots the number of restrictions and conditions specified in the RoMEO publisher policy summary against the background context of the number of publishers offering paid open access options such as Gold or hybrid journals. Just as there is an upward trend among these 107 publishers in the number of publishers allowing some form of self-archiving (Figure 4), there is also a year-on-year rise in the number of restrictions and conditions constraining the right to self-archive and the offer of paid open access options. See Appendix 2 for a full list of coded restrictions and conditions.

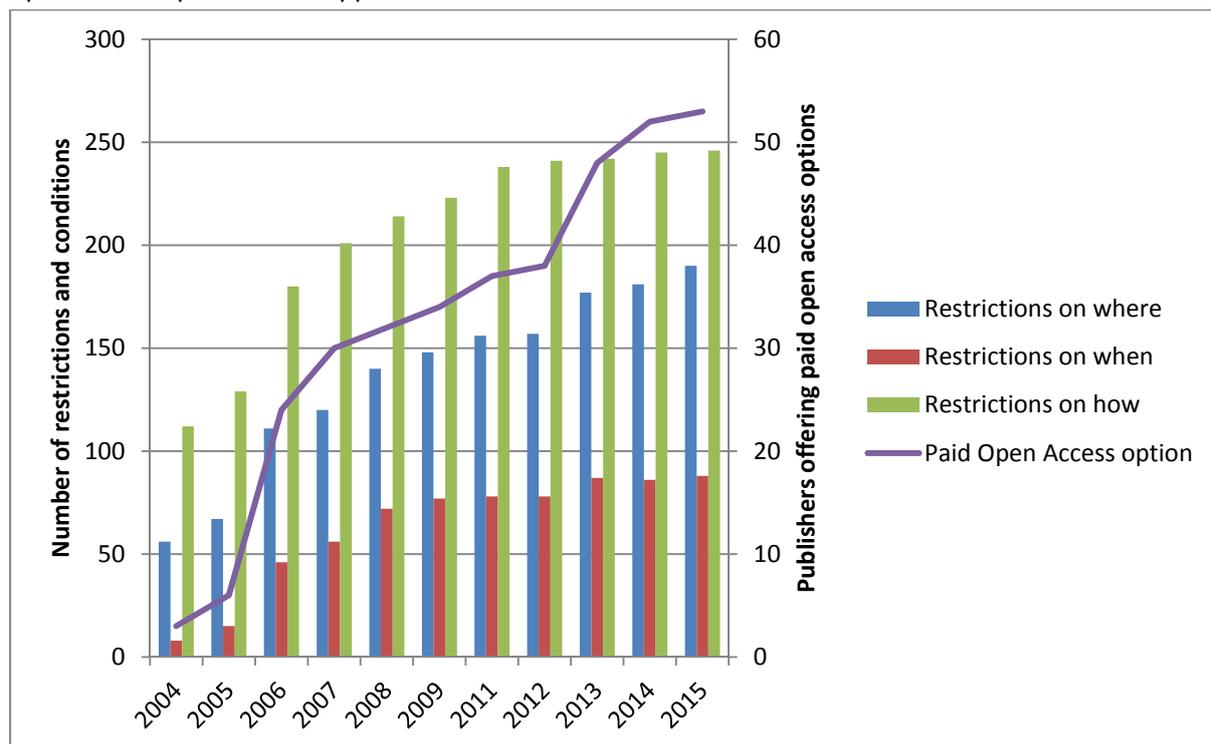


Figure 5 Introduction of restrictions and conditions over time\* mapped against the introduction of a paid open access option

\*No data was available for 2010.

In Figure 5 the bars marked 'Restrictions on when' illustrate the growing trend of publishers imposing embargo periods and other forms of delay such as a requirement for self-archiving to take

place “post-publication”. Restrictions around when a paper may be self-archived grew from eight incidences in 2004, with a massive leap in 2007, to a total of 88 separate instructions in 2015 – a growth of 1000%. By the end of the study period, many publishers had a range of circumstance-dependent or ‘linked’ embargoes. Circumstances might include whether a paper was funded by a particular funding agency, or where the paper was to be archived. Embargo periods are explored further below.

Restrictions relating to *where* a paper may be archived were even more prevalent and followed a similar growth pattern, again because of linked restrictions, e.g., an embargo on when an open access copy is allowed in an institutional repository. In 2004 there were 56 policy items relating to where a paper could be self-archived. After a leap in 2007, 190 policy items were recorded in 2015 – a growth of 190%. The types of permitted locations developed over time, and interpreting the meaning of the policy description was not always straightforward.

By far the tallest bars on the chart are those relating to the additional restrictions and conditions authors must adhere to in order to comply with a publisher’s open access policy. These grew from 112 in 2004 to 246 in 2015 – a growth of 119%. Nineteen different restrictions and conditions were coded overall, but the linking of these restrictions with the *where* and *when* of self-archiving meant that rarely did any two publishers adopt the same approach – a situation no doubt confusing to authors and likely to discourage them from self-archiving or from caring about compliance with publisher policy.

Figure 5 also shows the changing context in which the publishers’ self-archiving policies were operating. In 2004 just three publishers offered a paid open access option, by 2015 this rose to 53 – a growth of 1,670%. Note that there is a strong and positive correlation between the number of publishers offering paid open access with the number placing restrictions on *when* authors may self-archive (Pearson  $r=0.9648$ ); *where* authors may self-archive ( $r=0.9914$ ), and *how* ( $r=0.9603$ ). All correlations are significant at the  $p=0.01$  level. This demonstrates a clear link between the introduction of Gold open access and the increasing restrictions around Green open access.

### **When – an analysis of embargo periods**

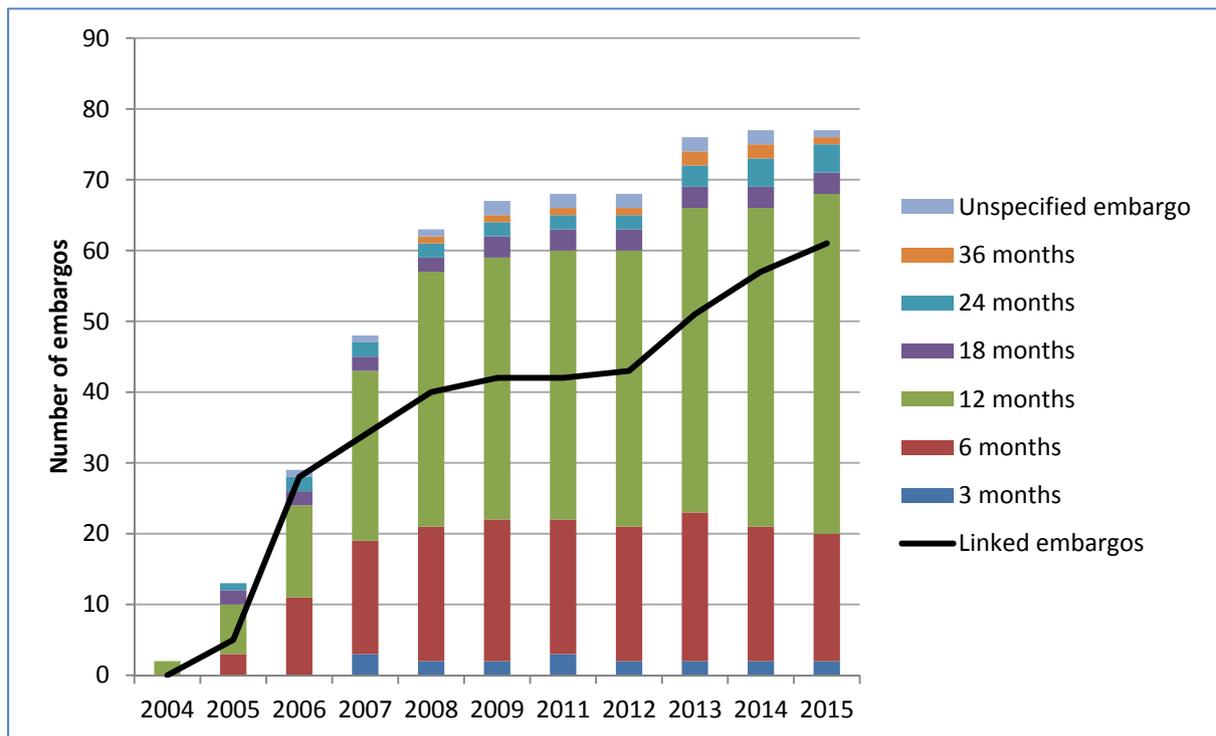


Figure 6 Growth of embargo periods over time

Figure 6 illustrates the growth in the number of publisher imposed embargo periods over the twelve year period. The numbers do not coincide with the *when* data in Figure 5, as that data includes other restrictions around when a paper can be self-archived, such as ‘post-publication’. Once again, some publishers have more than one embargo depending on the circumstances, so these data indicate incidences of embargoes, rather than the number of publishers with embargos. The black line shows the rise in circumstance-linked embargos, which accounts for much of the growth.

The growth in the number of imposed embargos was rapid until 2008, and thereafter held rather steady until 2012, when it began to accelerate again. The apparent plateau between 2011 and 2012 may be due to the missing data points on these dates as described above. After a fairly even split between 6 month and 12 month embargoes in 2005, by 2015, 62% of embargos were for 12 months, which seems to be becoming the industry norm in conformance with funder mandates. In 2015, 31 of the 61 linked embargoes were the result of funder mandates. Medical and biological science funders such as the National Institutes for Health (NIH), Wellcome Trust, the Biotechnology and Biological Sciences Research Council (BBSRC) and the Medical Research Council (MRC) were often mentioned by name in publisher open access policies, as was PubMed, their repository of choice. A further 23 of the 61 linked embargos related to where the paper is self-archived, of which 12 related to deposit in an institutional or subject repository, three in a designated funder repository, and three in any open access repository.

#### Where – an analysis of permitted self-archiving locations

To better understand publisher requirements regarding where open access copies may be made available, an analysis of deposit location terminology was performed. See Figure 7. Not all policy summaries used the same terms, so we manually categorised and coded for this analysis. Only locations in which the publisher permitted some form of deposit were coded.

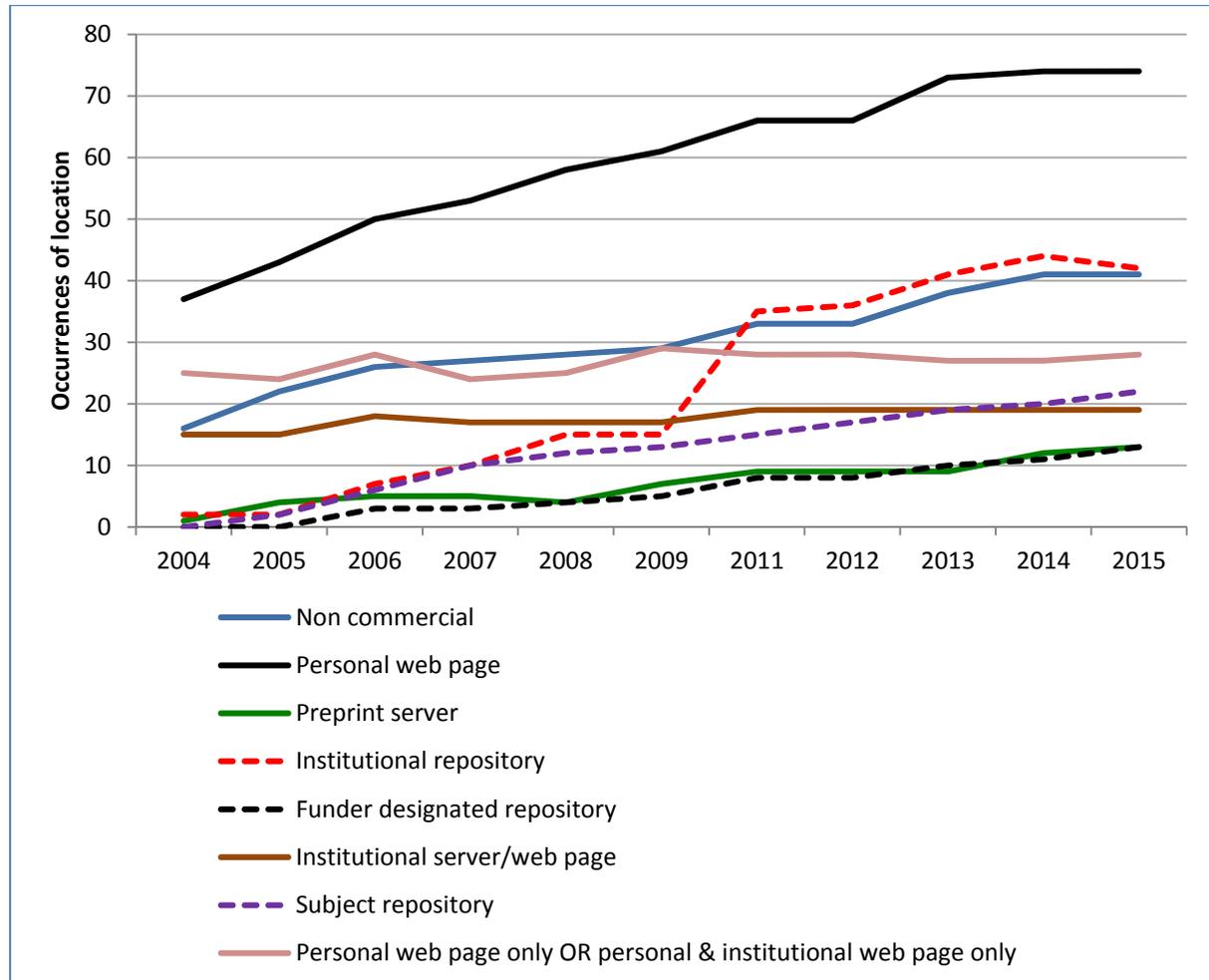


Figure 7 Occurrences of location terms for acceptable deposits of open access copies

The prevalence of terms specifying acceptable locations for open access copies has increased over time, again mainly because of the linking of embargo periods with particular locations. The most common term, and the one showing the greatest growth, was that of the personal web page. In many cases a publisher would move away from a more restrictive position (e.g., from RoMEO 'white' to another colour code) simply by allowing authors to self-archive a version of their paper immediately on a personal web page alone. Approximately 10% of publishers each year “earned” their colour by allowing authors to self-archive on a personal web page alone; about 25% only allowed self-archiving on a personal or institutional web page, but not in any form of repository. Note in Figure 7 that the locations preferable to authors or their funders and institutions – institutional, subject or funder-designated repositories (Fry et al, 2011; Harnad, 2015; HEFCE, 2015) – are mentioned far less frequently than personal web pages as permitted locations.

Mentions of institutional repositories grew significantly between 2009 (15 occurrences) and 2011 (35 occurrences). Some of this was background growth as a result of the increasing number of

institutional repositories. However in many cases, this was again the result of publishers placing an embargo on deposit in institutional repositories. Indeed, by 2015, 28 of the 42 occurrences of the term 'institutional repository' (two-thirds) were linked to some form of restriction on their use. In 2004 there were no mentions of funder designated repositories or subject repositories. However, by the end of the twelve years, there were 13 references to the former and 22 references to the latter.

Interpreting the meaning of the location terms used in RoMEO policy summaries was challenging. For example, in the first six years (2004-2009), the term "institutional server" was more prevalent than "institutional web page". In the latter five years (2011-2015), the prevalence of each term was reversed. Whether publishers simply replaced the term "server" with "web page", believing them to be synonymous, or whether they perceived the terms to mean two different locations is unclear. The growing number of references to institutional repositories in 2011 suggests that publishers not using this term probably did so deliberately, believing an institutional web page or server to be something significantly different from a repository. Indeed, some publishers placed an embargo on repository deposits but permitted immediate deposit on an institutional web page. However, this is by no means certain, and no advice on interpreting the publishers' terminology is provided by the RoMEO database.

#### **How – an analysis of the additional requirements of publishers**

Other studies have examined the *when* and *where* of self-archiving from different perspectives (Laakso, 2014; Bjork et al, 2014), but there is very little in the literature about the plethora of additional demands that publishers place on *how* authors are to self-archive their work (Troll Covey, 2013). These demands might include acknowledging the publisher, perhaps with a set statement, linking to the publisher version, removing the pre-print on publication of the post-print, and so on. (A full list of the requirements coded in the study is provided in Appendix 2.) As Figure 5 shows, the number of these requirements outstrips those around the *when* and *where*. Too often, these demands fall to an institutional repository manager to meet, rather than the authors themselves, and there is evidence that this can be particularly onerous (Boyes, 2015). Indeed the Research Consulting study (2014) found that "several institutions participating in [their] project raised the question of whether libraries' current role in rigorously policing individual deposits to ensure compliance with publisher policies is sustainable as volumes increase". Figure 8 provides a breakdown of the most prevalent additional requirements.

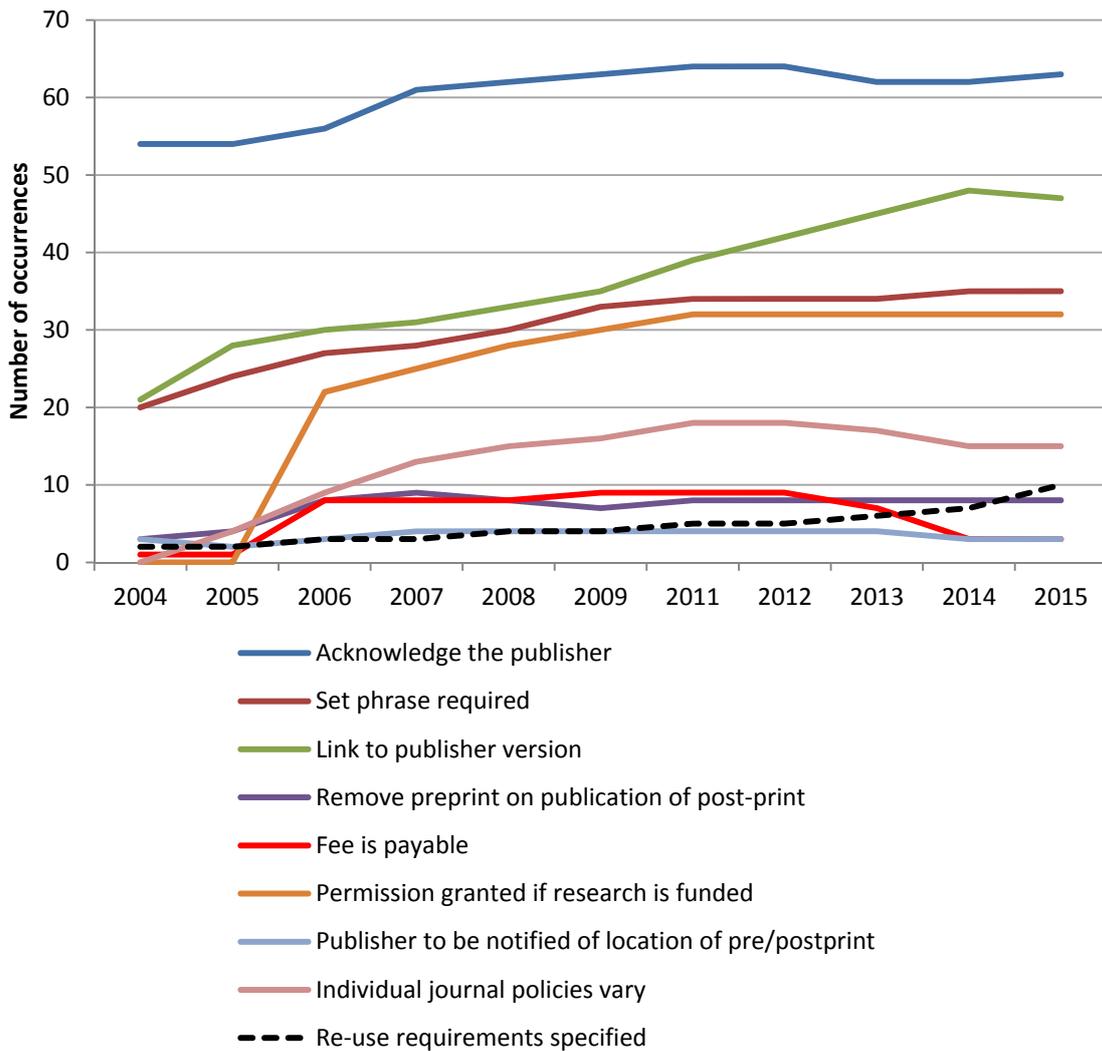


Figure 8 Occurrences of additional publisher restrictions and conditions in RoMEO publisher policies

Unlike the *when* and *where* analyses, it was less likely for a publisher to link a requirement as to *how a paper was to be self-archived* to another aspect of their policy. The main exception to this was where permission was only granted if research was subject to a funder mandate, in which case an embargo and funder designated repository were often specified. As a result, we see some publisher requirements reaching a plateau over time, or even diminishing. Acknowledging the publisher was the most common requirement. By 2015, 63 publishers requested this (59%) and for 35 of these (33% overall) a set phrase was required. This is significant because a repository manager could create a generic coversheet for institutional repository papers to acknowledge the publisher, but the requirement of a set phrase makes standardizing this task impossible. The requirement to link to the published version also grew over the period and seems to have plateaued in 2014-15 at 44%.

Other requirements were less prevalent and show no signs of growth, but still create additional work for a repository manager because they do not apply to all papers. These include a requirement for a fee to be paid if an author wants to make their paper available earlier than the embargo period or wants to use the publisher PDF; a requirement to remove the pre-print once the post-print is available; and a requirement to notify the publisher of the location of the pre-print or post-print. In

addition to the generic restrictions and conditions listed on a publisher's policy summary, some summaries indicated that the publisher policy varied according to the journal. As Waajters et al (2008) have commented, "there are more policies than publishers as some publishers even apply a different policy to the different journals they publish". Among the 107 publishers studied, this peaked in 2011-12 at 18 and dropped to 15 publishers (14%) in 2015.

Interestingly, the number of publishers specifying particular re-use requirements has grown over the period but remains low. In 2004, only two publishers restricted re-use of the self-archived papers and the restrictions were not specified in the RoMEO policy summary. By 2015, ten publishers specified re-use requirements, mainly in response to funder mandates requiring that use of papers arising from their funding be licensed under specific Creative Commons (CC [n.d.]) licences. However, some of the licences specified would not meet funder requirements. Indeed, four different CC licences were listed in 2015, including two publishers opting for the most restrictive CC-BY-NC-ND licence, which requires author attribution, allows only non-commercial re-use, and prohibits the creation of derivatives – restrictions that do not comply with the Bethesda definition of open access.

### **Discussion - What *should* green mean?**

It has been demonstrated that the Bethesda definition of open access includes elements not taken into consideration by the RoMEO colour codes. These include the need for immediate (embargo-free) deposit in an open-access repository under liberal re-use conditions without having to seek permission. Neither do the RoMEO colour codes address the strong preference of academic authors and other users to read and cite the version of record (Fry et al, 2011), i.e., the publisher PDF. The evidence presented in this paper demonstrates how publishers have in theory become 'greener' over time by allowing some sort of self-archiving while simultaneously increasing restrictions around the *when*, *where* and *how* of self-archiving. The complexity and variety of publisher self-archiving policies make compliance ever more difficult for academic authors or their representatives, at best discouraging the practice they purport to support; at worst engendering disrespect for those policies.

The growth of embargo periods since the inception of the RoMEO database has been considerable, and, as Jenkins et al (2008) have noted, goes against the "spirit and meaning of open access". Set against a backdrop of ever-increasing paid Gold open access options, it is not surprising that publishers would simultaneously seek to withdraw a free product (immediate open access) that they are also trying to sell through hybrid or gold OA journals. As Troll Covey (2013) has noted, "Publishers impose embargoes on self-archiving purportedly to avoid cancelled subscriptions and recoup costs. However, the convergence of embargoes and paid open access options suggests not cost recovery, but a new revenue stream."

Remarkably, the growth in embargo periods has not caused a reduction in publishers being allocated a RoMEO colour code - quite the opposite. The way many publishers have achieved this is to link embargo periods with deposit in a repository (a preferred location stipulated in the Bethesda definition of OA), while still allowing authors to either self-archive immediately on a personal web page only (10% of publishers) or on both a personal and institutional web page (25% of publishers).

Indeed the number of mentions of personal web pages was the greatest and fastest growing location of deposit over the period studied. Of course, self-archiving on web pages reduces the discoverability and permanence of the paper. The growth in the number of publishers allowing the deposit of the publisher PDF was encouraging, but again there were considerable restrictions around this.

To assess publisher compliance with the Bethesda definition of open access and alignment with authors' preference for providing open access to the post-print version of their papers, if not the publisher PDF, we redefined what it means to be green and conducted an analysis comparing 'redefined green' and the original 'RoMEO green'. The conditions for 'redefined green' were that a publisher offered immediate (embargo -free) deposit of the post-print in an institutional repository without any requirement to notify the publisher they have done so. The results are shown in Figure 9. As many authors have a strong preference for the version of record, i.e., the publisher PDF, the chart also shows where a publisher allows the self-archiving of the publisher PDF under 'redefined green' conditions

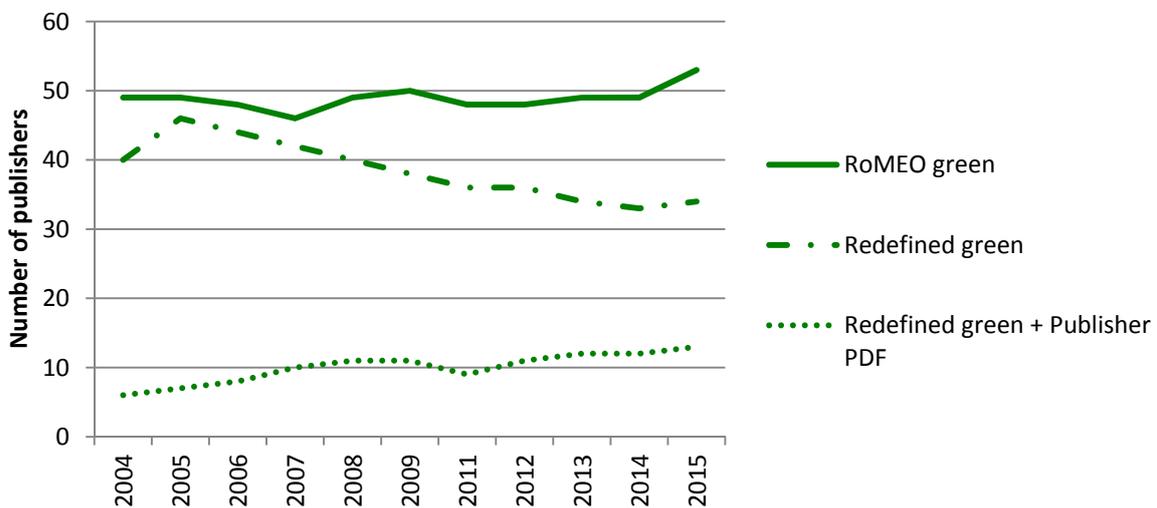


Figure 9. The trajectory of 'RoMEO green' and 'redefined green' over time.

For the 107 publishers studied, the trajectory of RoMEO 'green' is rather stable or slightly upward over time with an overall increase of 8%. The trajectory for 'redefined green' is strikingly different, far lower in comparison and declines by 35% over time. Interestingly, the trajectory for the number of 'redefined green' publishers allowing the self-archiving of the publisher PDF is increasing. This may indicate that publishers allowing self-archiving in repositories have recognised that this is not a threat to their business model. As Fry et al (2011) found, "Open Access Repositories are perceived by researchers as complementary to, rather than replacing, current forums for disseminating and publishing research."

### Conclusions and recommendations

The data strongly suggest that achieving RoMEO green status has become a target for publishers rather than a commitment to open access *per se*. This is not a surprise. For most publishers and learned societies, publishing is primarily a commercial or income-generating activity, and the threat

of unfettered open access to their commercial interests has been much discussed. However, the original purpose of the RoMEO colour codes was to categorise a publisher's commitment to Green open access. By failing to address some of the key elements and developments in the Bethesda definition of open access, the system enables publishers to be perceived as meeting the highest standard, 'green', without allowing authors to achieve the highest levels of openness, discoverability, permanence and re-use. In answer to our original research question, it is apparent that in many cases changes to publisher open access policies over time enable the illusion of full support (sufficient to secure RoMEO code green), but in reality discourage self-archiving in a manner consistent with author preferences for discovery, preservation, reuse and increased impact.

Colour-coding open access implies that there is a hierarchy: the Gold route of open access publishing being the pinnacle of achievement and the Green route of self-archiving, particularly RoMEO green, being the next best thing and RoMEO white the worst. This is not borne out in practice. Many Gold open access publishers charge unaffordable Author Processing Charges that are not welcomed by the scholarly community (Pinfield et al, 2015). Some RoMEO green publishers are fairly restrictive about the self-archiving they permit. Some white publishers actually allow self-archiving. The vast majority of publishers in this study are introducing an ever-increasing range of restrictions and conditions making compliance with their policies by authors or repository managers almost unmanageable. The four RoMEO colour codes imply that there are four key open access positions. In fact, the current situation is far more complex. Indeed, one of the critical elements of open access – liberal re-use – is not systematically included in the RoMEO policy summaries at all.

We recommend that open access advocates, institutions, funders and authors revisit together the open access future they would like to see, and create a new grading system to define and guide that future. The result might not be a series of four colour codes, but a spectrum, as with the Open Access Spectrum Evaluation Tool (SPARC, 2016). The response of publishers to funder mandates and their apparent gaming of the RoMEO colour codes demonstrate that their policies are not so much a statement of principle, but a pragmatic approach to a changing open access environment. If scholars, institutions and funders work together to devise and implement a new grading system for publisher open access policies and practices, they could create the conditions that enable open access to achieve its full potential to increase the impact of scholarly work and accelerate advances in the disciplines.

## References

Anon, 2015. Internet archive: Wayback Machine. Available at: <http://archive.org/web/> [Accessed February 6, 2015].

Archambault, E. et al., 2014. Proportion of Open Access Papers Published in Peer-Reviewed Journals at the European and World Levels- 1996 – 2013. Available at: <http://science-metrix.com/en/publications/reports/proportion-of-open-access-papers-published-in-peer-reviewed-journals-at-the>

- Bailey, B.C.W., 2001. What Is Open Access? “ Open Access ” Defined. Available at: <http://digital-scholarship.org/cwb/WhatIsOA.pdf>.
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. 2003. Available at: <http://openaccess.mpg.de/Berlin-Declaration>.
- Bjork, B., Laakso, M. & Welling, P., 2014. Anatomy of green open access. *Journal of the Association for Information Science and Technology*, 65(2), pp.237–250.
- Boyes, P., 2015. A Day in the Life of an Open Access Research Adviser. *Unlocking research. Blog of the University of Cambridge Office for Scholarly Communication*. Available at: <https://unlockingresearch.blog.lib.cam.ac.uk/?p=340>.
- Budapest Open Archive Initiative. 2002. Available at: <http://www.budapestopenaccessinitiative.org/read>.
- Creative Commons, [n.d.] *Creative Commons Licences*. Available at: <https://creativecommons.org/licenses/>
- Fry, J. et al., 2011. *PEER behavioural research : authors and users vis-a-vis journals and repositories. D4.2 final report*. Available at: <https://dspace.lboro.ac.uk/2134/12967>.
- Gadd, E., Oppenheim, C. & Proberts, S., 2003. RoMEO studies 4: an analysis of journal publishers' copyright agreements. *Learned publishing*. Available at: <http://www.ingentaconnect.com/content/alpsp/lp/2003/00000016/00000004/art00009>
- Harnad, S. et al., 2004. The green and the gold roads to Open Access. *Nature Focus*, 17 May. Available at: <http://www.nature.com/nature/focus/accessdebate/21.html>.
- Harnad, S. et al., 2008. The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. *Serials Review*, 34(1), pp.36–40
- Harnad, S., 2011. Open Access to Research: Changing Researcher Behaviour Through University and Funder Mandates. *JeDEM: eJournal of eDemocracy & Open Government*, 3(1), pp.33–41. Available at: [www.jedem.org](http://www.jedem.org).
- HEFCE, 2014. Research Excellence Framework. Available at: <http://www.ref.ac.uk/>.
- HEFCE, 2015. Policy for open access in the post-2014 Research Excellence Framework. , 07(July 2015), pp.1–11. Available at: <http://www.hefce.ac.uk/pubs/year/2014/201407/#d.en.86771> [http://www.hefce.ac.uk/media/hefce/content/pubs/2014/201407/HEFCE2014\\_07.pdf](http://www.hefce.ac.uk/media/hefce/content/pubs/2014/201407/HEFCE2014_07.pdf).
- Hubbard, B., [n.d.] *Green, Blue, Yellow, White & Gold*. Available at: <http://www.sherpa.ac.uk/documents/sherpaplusdocs/Nottingham-colour-guide.pdf>.

Jenkins, C. et al., 2008. RoMEO studies 7: creation of a controlled vocabulary to analyse copyright transfer agreements. *Journal of Information Science*, 34(3), pp.290–307. Available at: <http://jis.sagepub.com/content/34/3/290> \n<http://jis.sagepub.com/content/34/3/290.abstract> \n<http://jis.sagepub.com/content/34/3/290.full.pdf>.

Jenkins, C. et al., 2007. RoMEO Studies 8: self-archiving: the logic behind the colour-coding used in the Copyright Knowledge Bank. *Program*, 41(2), pp.124–133. Available at: <http://www.emeraldinsight.com/10.1108/00330330710742908>.

Kramer, B. & Bosnen, J., 2016. 101 Innovations in Scholarly Communication [Results: selecting a journal to submit to]. *101 Innovations in Scholarly Communication*. Available at: [http://dashboard101innovations.silk.co/page/Select-journal-to-submit-to -](http://dashboard101innovations.silk.co/page/Select-journal-to-submit-to-).

Laakso, M., 2014. Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed. *Scientometrics*, 99, pp.475–494.

Loughborough University (2016), Research Data Archive. Available at: <http://lboro.figshare.com>

Mabe, M., 2009. *Scholarly Publishing*, Available at: [http://www.journals.cambridge.org/abstract\\_S1062798709000532](http://www.journals.cambridge.org/abstract_S1062798709000532).

Miguel, S., 2011. Open Access and Scopus: A New Approach to Scientific Visibility From the Standpoint of Access. *Journal of the American Society for Information Science & Technology*, 62(6), pp.1130–1145.

Nicholas, D., Huntington, P. & Rowlands, I., 2005. Open access journal publishing: the views of some of the world's senior authors. *Journal of Documentation*, 61(4), pp.497–519. Available at: <http://www.emeraldinsight.com/doi/pdfplus/10.1108/00220410510607499>

Pinfield, S., Salter, J. & Bath, P.A., 2015. The “Total Cost of Publication” in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination With Subscriptions. *JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY*, .

Proudman, V., [n.d.] Analysis of funder open access policies around the world. Available at: <http://sparceurope.org/analysis-of-funder-open-access-policies-around-the-world/>.

Research Consulting, 2014. *Counting the Costs of Open Access*, Available at: <http://www.ariessys.com/wp-content/uploads/Research-Consulting-Counting-the-Costs-of-OA-Final.pdf>

ROARMAP, 2016. Registry of Open Access Repository Mandates and Policies. Available at: <http://roarmap.eprints.org/>.

SHERPA, 2016a, RoMEO database. Available at: <http://www.sherpa.ac.uk/romeo>

SHERPA, 2016b. JULIET database. Available at: <http://www.sherpa.ac.uk/JULIET>.

Solomon, D.J., Laakso, M. & Björk, B.C., 2013. A longitudinal comparison of citation rates and growth among open access journals. *Journal of Informetrics*, 7(3), pp.642–650. Available at: <http://dx.doi.org/10.1016/j.joi.2013.03.008>.

SPARC, 2016. Open Access Spectrum Evaluation Tool. Available at: <http://oaspectrum.org/>.

SPARC, 2014. How open is it? Version 2.0. Available at: [https://www.plos.org/files/HowOpenIsIt\\_English.pdf](https://www.plos.org/files/HowOpenIsIt_English.pdf)

Spezi, V. et al., 2013. Researchers' green open access practice: a cross-disciplinary analysis. *Journal of Documentation*, 69, pp.334–359. Available at: <http://www.emeraldinsight.com/journals.htm?issn=0022-0418&volume=69&issue=3&articleid=17088385&show=html>.

Suber, P. et al., 2003. Bethesda statement on open access publishing. Available at: <http://dash.harvard.edu/handle/1/4725199>.

Suber, P., 2004. Praising progress, preserving precision. *SPARC Open Access Newsletter*, 77.

Swan, A. & Brown, S., 2004. Authors and open access publishing. *Cogprints*, pp.219–224. Available at: <http://www.trans.uma.es/numeros.html>.

Troll Covey, Denise., 2009. Self-Archiving Journal Articles: A Case Study of Faculty Practice and Missed Opportunity. *portal: Libraries and the Academy*, 9(April), pp.223–251.

Troll Covey, D., 2013. Publisher Support for Self-Archiving : Laudatory or Predatory ? *Proceedings of the International Conference on Digital Libraries*, (January). Available at: [http://works.bepress.com/denise\\_troll\\_covey/79/](http://works.bepress.com/denise_troll_covey/79/).

Waaijters, L., 2008. Copyright Angst, Lust for Prestige and Cost Control: What Institutions Can Do to Ease Open Acces. *Ariadne: Web Magazine for Information Professionals*. Available at: <http://www.ariadne.ac.uk/issue57/waaijters-et-al>.

Ware, M. & Mabe, M., 2009. *The STM Report*, (November). Available at: [http://www.stm-assoc.org/2009\\_10\\_13\\_MWC\\_STM\\_Report.pdf](http://www.stm-assoc.org/2009_10_13_MWC_STM_Report.pdf).

Xia, J. & et al, 2012. A review of open access self-archiving mandate policies. *portal: Libraries and the Academy*, 12(1), pp.85–102.

## Appendix 1

Academy of Management
American Anthropological Association
American Association for the Advancement of Science
American Chemical Society
American Economics Association
American Geophysical Union
American Institute of Aeronautics and Astronautics
American Institute of Physics
American Mathematical Society
American Medical Association
American Meteorological Society
American Physical Society
American Physiological Society
American Psychological Association
American Public Health Association
American Society for Clinical Investigation
American Society for Microbiology
American Society for Pharmacology and Experimental Therapeutics
American Society of Civil Engineers
American Society of Hematology
American Society of Limnology and Oceanography (ASLO)
American Society of Plant Biologists
American Sociological Association
Annual Reviews
Arnold Publishers (Later an imprint of Sage)
Association for Computing Machinery
Association for the Advancement of Computing in Education
Association of Applied Biologists
Australian Academic Press
Australian Computer Society Inc
Berkeley Electronic Press
BioMed Central
Blackwell Publishing (Later part of Wiley-Blackwell)
BMJ Publishing Group
British Institute of Non-destructive Testing
CAB International Publishing
Cambridge University Press
Clinical Laboratory Science (Later American Society for Clinical Laboratory Science)
Company of Biologists
Ecological Society of America
Electrochemical Society

Elsevier
Emerald
Endocrine Society
Evolutionary Ecology
Geological Society
Geological Society of America
Georgetown University Law Center
Haworth Press
Hindawi Publishing Corporation
Imperial College Press
Institute of Biology
Institute of Electrical and Electronics Engineers (IEEE)
Institute of Electrical, Information and Communication Engineers
Institute of Mathematical Statistics
Institute of Physics
Institution of Chemical Engineers
Institution of Electrical Engineers (IEE) (Later IET)
International Press
International Union of Pure and Applied Chemistry
Internet Journal of Chemistry
IOS Press
John Wiley & Sons, Inc.
Johns Hopkins University Press
Kluwer (now part of Springer)
Lawrence Erlbaum Associates, Inc.
Lippincott, Williams & Wilkins
Marcel Dekker
Mary Ann Liebert
Massachusetts Institute of Technology Press
Massachusetts Medical Society
Materials Research Society
Medknow Publications
Michigan Law Review
National Academy of Science
National Research Council Canada
Nature Publishing Group
Nordic Ecological Society
Oceanography Society
Optical Society of America
Oxford University Press
Physicians Postgraduate Press
Portland Press
Professional Engineering Publishing (Institutional of Mechanical Engineers)

Resilience Alliance
Rockefeller University Press
Royal College of General Practitioners
Royal Meteorological Society
Royal Society
Royal Society of Chemistry
Royal Society of Medicine
SAGE Publications
School of Management, University of Bath
Sheffield Academic Press
Society for Endocrinology
Society for General Microbiology
Society for Industrial and Applied Mathematics
Society for In-Vitro Biology
Society of Dyers and Colourists
Society of Photo-Optical Instrumentation Engineers
Springer Verlag (Germany)
Stanford University Law School
Taylor & Francis
University of California Press
University of Chicago Press
Wiley-VCH Verlag Berlin
Yale Law School

## Appendix 2

Coded restrictions on *what version* may be self-archived

- Publisher version allowed
- Publisher version allowed with restrictions
- Publisher version required
- Publisher version prohibited
- Only pre-print allowed
- Only post-print allowed
- Pre-print and post-print allowed

Coded restrictions on *where* authors may self-archive the designated version(s)

- Personal website
- Institutional repository
- Subject repository
- Elsewhere online
- On a password protected server
- Institutional or employer website
- Non-commercial server
- Funder designated repository
- Pre-print server
- Open access repository
- Social website

Coded restrictions on *when* authors may self-archive the designated version(s)

- 3 month embargo
- 6 month embargo
- 12 month embargo
- 18 month embargo
- 24 month embargo
- 36 month embargo
- Unspecified embargo
- Post-publication

Coded restrictions on *how* authors may self-archive and the OA context

- Acknowledge publisher
- Acknowledge with set phrase
- Link to published version or journal
- Obtain permission
- Pay a fee

- Archiving only allowed by those with specific funders
- Must notify publisher
- Must not update preprint version
- Post-print must match the published version
- Only applies to those who signed Licence to Publish rather than Copyright Transfer Agreement
- Restrictions on re-use apply
- Pre-print must be removed once Post-print available
- Individual journal titles may vary from general policy
- Author needs to complete form for Publisher to submit to PubMed on their behalf
- Author can only self-archive if voluntary and not mandated.
- If archiving is mandated will only be permitted if a separate agreement exists with the publisher and the mandating body.
- Activity only permitted if required by law
- Applies if deposit mandated

