



## REVIEW

# The academic, economic and societal impacts of Open Access: an evidence-based review [version 1; referees: awaiting peer review]

Jonathan P. Tennant<sup>1</sup>, François Waldner<sup>2</sup>, Damien C. Jacques<sup>2</sup>, Paola Masuzzo<sup>3,4</sup>, Lauren B. Collister<sup>5</sup>, Chris. H. J. Hartgerink<sup>6</sup>

<sup>1</sup>Department of Earth Science and Engineering, Imperial College London, London, UK

<sup>2</sup>Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

<sup>3</sup>Medical Biotechnology Center, VIB, Ghent, Belgium

<sup>4</sup>Department of Biochemistry, Ghent University, Ghent, Belgium

<sup>5</sup>University Library System, University of Pittsburgh, Pittsburgh, PA, USA

<sup>6</sup>Department of Methodology and Statistics, Tilburg University, Tilburg, Netherlands

**v1** First published: 11 Apr 2016, 5:632 (doi: [10.12688/f1000research.8460.1](https://doi.org/10.12688/f1000research.8460.1))

Latest published: 11 Apr 2016, 5:632 (doi: [10.12688/f1000research.8460.1](https://doi.org/10.12688/f1000research.8460.1))

## Abstract

Ongoing debates surrounding Open Access to the scholarly literature are multifaceted and complicated by disparate and often polarised viewpoints from engaged stakeholders. At the current stage, Open Access has become such a global issue that it is critical for all involved in scholarly publishing, including policymakers, publishers, research funders, governments, learned societies, librarians, and academic communities, to be well-informed on the history, benefits, and pitfalls of Open Access. In spite of this, there is a general lack of consensus regarding the advantages or disadvantages of Open Access at multiple levels. This review aims to be a resource for current knowledge on the impacts of Open Access by synthesizing important research in three major areas of impact: academic, economic and societal. While there is clearly much scope for additional research, several key trends are identified, including a broad citation advantage for researchers who publish openly, as well as additional benefits to the non-academic dissemination of their work. The economic case for Open Access is less well-understood, although it is clear that access to the research literature is key for innovative enterprises, and a range of governmental and non-governmental services. Furthermore, Open Access has the potential to save publishers and research funders considerable amounts of financial resources. The social case for Open Access is strong, in particular for advancing citizen science initiatives, and leveling the playing field for researchers in developing countries. Open Access supersedes all potential alternative modes of access to the scholarly literature through enabling unrestricted re-use, and long-term stability independent of financial constraints of traditional publishers that impede knowledge sharing. Open Access remains only one of the multiple challenges that the scholarly publishing system is currently facing. Yet, it provides one foundation for increasing engagement with researchers regarding ethical standards of publishing. We recommend that Open Access supporters focus their efforts on working to establish viable new models and systems of scholarly communication, rather than trying to undermine the existing ones as part of the natural evolution of the scholarly

## Open Peer Review

**Referee Status:** AWAITING PEER

REVIEW

**Discuss this article**

Comments (1)



ecosystem. Based on this, future research should investigate the wider impacts of an ecosystem-wide transformation to a system of Open Research.



This article is included in the **The Future of Scholarly Publishing** channel.

**Corresponding author:** Jonathan P. Tennant ([jon.tennant.2@gmail.com](mailto:jon.tennant.2@gmail.com))

**How to cite this article:** Tennant JP, Waldner F, Jacques DC *et al.* **The academic, economic and societal impacts of Open Access: an evidence-based review [version 1; referees: awaiting peer review]** *F1000Research* 2016, 5:632 (doi: [10.12688/f1000research.8460.1](https://doi.org/10.12688/f1000research.8460.1))

**Copyright:** © 2016 Tennant JP *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Grant information:** This research was partly funded by the Belgian National Fund for Scientific Research through a FRIA grant. PM acknowledges support from the European Commission Horizon 2020 Programme under Grant Agreement 634107 (PHC32-2014) 'MULTIMOT'. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Competing interests:** JPT currently blogs for the PLOS Paleo Community, and works for ScienceOpen. CHJH is a Center for Open Science ambassador. DCJ and FW are members of the Open Access Working Group of EURODOC. PM is a Research Data Alliance member. LBC works for the University of Pittsburgh, which has an Open Access library publishing department. All views presented here are strictly personal.

**First published:** 11 Apr 2016, 5:632 (doi: [10.12688/f1000research.8460.1](https://doi.org/10.12688/f1000research.8460.1))



## Introduction

Open Access (OA) refers to the removal of major obstacles to accessing, sharing and re-using the outputs of scholarly research. The rationale is that the research process is facilitated by ensuring rapid and widespread access to research findings such that all communities have the opportunity to build upon them. Reflecting this ambition, there are currently over 700 OA policies and mandates recorded worldwide from a range of research institutes and funding bodies ([roarmap.eprints.org](http://roarmap.eprints.org)). OA pertains only to documents made available via two main pathways: the ‘Gold’ route and the ‘Green’ route ([Harnad et al., 2008](#)). The Gold route refers to freely accessible research articles at the point of publication. This route is often, but not always, accompanied by article processing charges (APC). The Green route refers to author self-archiving, in which a version of the peer-reviewed article is posted online to a repository or website. This route is dependent on journal or publisher policies on self-archiving ([sherpa.ac.uk/romeo](http://sherpa.ac.uk/romeo)). Some publishers require an embargo period before deposit in public repositories is allowed. These embargoes are applied in order to avoid putative reductions in subscription income due to such self-archiving. Through these dual pathways, almost 25% of all scholarly documents archived on the Web are now obtainable via OA somewhere on the Internet ([Khabsa & Giles, 2014](#)).

A core issue remains: universal or even marginal access to ~75% of articles is not directly possible unless one is in a privileged position to work at an institute which has subscription access to a portion of these articles. A subscription to all peer-reviewed journals is not affordable for any single individual, research institute or university ([Odlyzko, 2006](#)). Consequently, the potential impact of research articles is never fully realized, impeding scientific progress by a lack of use, while simultaneously negatively affecting the recognition of individual researchers ([Hitchcock, 2013](#)) and the funders who support their work.

Based on these problems, free and unrestricted access to primary research literature has become a global goal of the OA movement. The steady increase in OA over the past two decades has required negotiations with a range of stakeholders (e.g., librarians, funders, academics). Much of the driving force behind this global change has been through a combination of direct, grassroots advocacy initiatives in conjunction with policy changes from funders and governments. The debates regarding the benefits of OA over subscription-based access often hinge on the increased value to academics. However, increased access has broader benefits to research through enhanced visibility, facilitating innovation by businesses, decreasing financial pressure on academic/research libraries (known more broadly as the ‘serials crisis’ ([Russel, 2008](#))), and fostering a culture of greater scientific literacy. OA also includes a moral aspect, where access to scientific knowledge and information is regarded as a fundamental aspect of global human equality. For example, Article 27 of the United Nations Declaration of Human Rights states that “*Everyone has the right to freely participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.*” ([United Nations, 1948](#)).

This review paper aims to provide information on the various benefits and drawbacks of Open Access to scholarly research. We

consider the case for OA from the academic, economic, and societal perspectives. In addition, we shortly consider the broader implications of OA on Open Data, a closely related issue, united under a general theme of ‘Open Research’. By aggregating evidence from primary sources, this review should be useful to those more broadly interested in the impact of scholarly research, as well as policymakers and others interested in implementing OA policies and strategies.

## A brief history of Open Access

The Open Access movement is intrinsically tied to the development of the Internet and how it redefined communication and publishing ([Laakso et al., 2011](#)). With increased availability of Internet bandwidth, print articles have become virtually redundant, and the costs per article potentially decrease as a result of not investing material resources in print publications. As a result, widespread dissatisfaction with the expensive traditional publishing model increased, resulting in the OA movement and concomitant innovations in scholarly publishing.

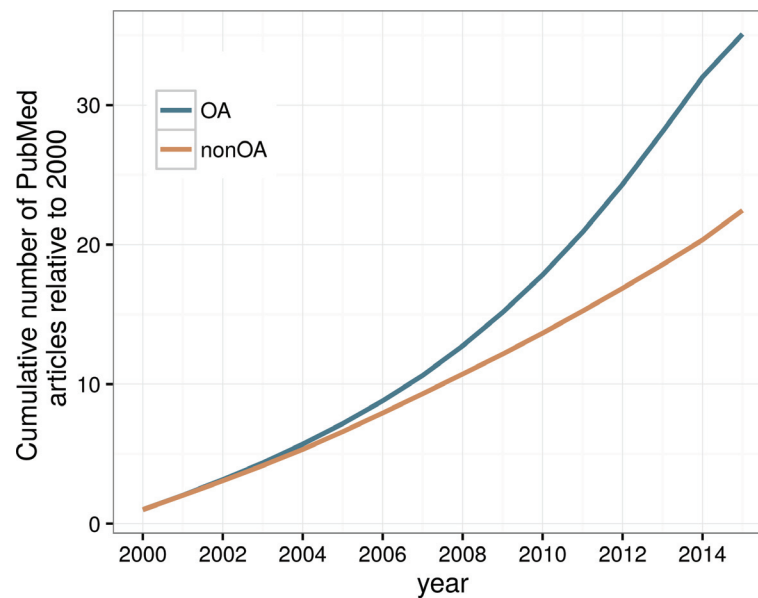
Interest in using the Internet for opening access to scientific research coalesced throughout the 1990s, culminating with the 2001 conference on “Free Online Scholarship” by the Open Society Institute in Budapest. The result of this conference was the release of the Budapest Open Access Initiative (BOAI), which is recognized as one of the defining points of the OA movement. The BOAI was the first initiative to use the term “Open Access” and articulated the following public definition:

*By “open access” to [peer-reviewed research literature], we mean its free availability on the public internet, permitting any 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, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.*

One result of the growing OA movement is the rise of OA-only publishers, who publish exclusively digital content and have demonstrated that such a business model is financially sustainable. For example, pioneer OA publishers BioMed Central and the Public Library of Science (PLOS) were founded in the early 2000s and remain successful businesses to date. More recently, OA publishing has gained increasing momentum among researchers, funders, and governments. This has led to a proliferation of innovative approaches to publishing (e.g., *PeerJ* ([peerj.com](http://peerj.com)), *F1000Research* ([f1000research.com](http://f1000research.com)), *Open Library of Humanities* ([openlibhums.org](http://openlibhums.org))) and a range of different policies from research funders and institutes mandating OA. All of these different policies and new business models, combined with traditional publishers innovating by launching their own OA titles and programs, have made the overall OA ecosystem quite complex.

Even with this growing prevalence of publishers that facilitate Open Access to the scholarly literature, Open Access is hardly





**Figure 1.** Ratio of the cumulative sum of articles in PubMed Central relative to 2000.

ubiquitous as of yet. Bjork *et al.* (2009) estimated that in 2006 the total number of published articles that year was approximately 1,350,000. Of these, 4.6% became immediately accessible and an additional 3.5% after an embargo period of, typically, one year. Furthermore, usable copies of 11.3% could be found in repositories or on the home pages of the authors. Since the National Institutes of Health mandated archival of articles in the public PubMed Central repository in 2008, the cumulative number of OA articles in PMC has increased more than the number of non-OA articles (see Figure 1). In 2013, the total percentage of OA articles available was estimated at 24% of English-language scholarly documents accessible on the Web (Khabsa & Giles, 2014).

Table 1 shows a non-exhaustive summary of the developments in the advancement of scholarly publishing and the Open Access movement. Included are the founding of major institutions in the movement as well as policy and legal developments. Several controversial moments are included, because they have spurred action or generated awareness for the movement. An example is the suicide of Aaron Swartz, who was arrested for downloading JSTOR articles on the grounds that he allegedly intended to make these publicly available. Another ongoing controversy is scholarly piracy; this includes the Sci-Hub and LibGen projects, which have created an online repository of pirated scholarly papers. Both gained increased attention after becoming the target of a lawsuit by the publisher Elsevier. There have been mixed responses to these kinds of activities, polarising the view that illegal acts regress or weaken the case for OA, while some hail the development as the 'Napster moment' for the OA movement which will force the established industry to change.

### The academic case for Open Access

The academic case for OA is two-faceted: (i) it is associated with a higher documented impact of scholarly articles through

availability and re-use; and (ii) it non-restrictively allows researchers to use automated tools to mine the scholarly literature. For the former, major arguments in favor of OA include the evidence that work that is openly available generates more academic citations, but also has more societal impact. The latter major argument involves non-restrictive access to the scholarly literature through appropriate licensing, making it possible to use automated tools to collect and analyze the entire body of scholarly literature in a legally sound framework and irrespective of copyright laws. The following sections cover these effects of OA on both the impact of scholarly research and through the mining of scholarly literature.

### The impact advantage

**Academic impact.** Academic impact is frequently measured through citation counts, and these remain fundamental as the 'currency units' for researchers, research groups, institutes and universities. Lawrence (2001) was the first to propose that OA would have a citation advantage. The utility and consistency of the citation advantage across different research fields has been intensively debated because its magnitude substantially varies depending on the discipline (Table 2). However, the general tendency from studies to date indicates that there is at least some association between OA publishing and increased citation counts across most disciplines (Antelman, 2004; Hajjem *et al.*, 2006) (Figure 2 and Table 2).

Estimates for the open citation advantage range from +36% (Biology) to +600% (Agricultural Sciences) (Swan, 2010). A longitudinal study Eysenbach (2006) compared the bibliometric impact of a cohort of articles from a multi-disciplinary journal (*Proceedings of the National Academy of Sciences*) that offers both OA and non-OA publishing options. After adjusting for potentially confounding variables, the results indicated that non-OA papers were twice as likely to remain uncited six months after

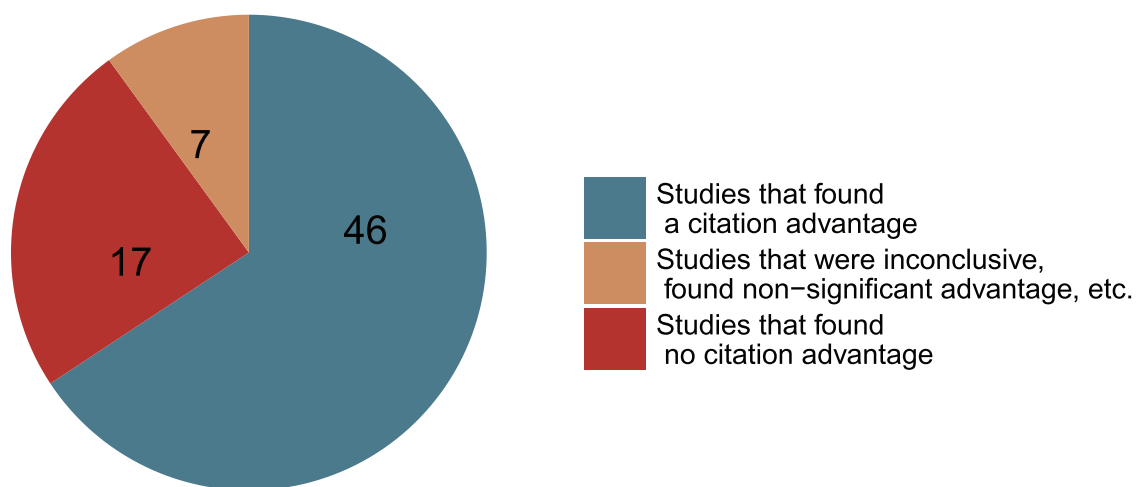


**Table 1. Major historical milestones in the progress of open research publishing.**

YEAR	MILESTONE
1454	Invention of <b>printing</b>
1665	January 5: First issue of The <b>Journal des sçavans</b> (later spelled <i>Journal des savants</i> ), the earliest academic journal published in Europe and established by Denis de Sallo.
1807	25-year-old <b>Charles Wiley</b> opens a small printing shop at 6 Reade Street in lower Manhattan.
1842	May 10: <b>Julius Springer</b> founded what is now Springer Science+Business Media in Berlin.
1848	<b>John Wiley</b> (son of Charles Wiley) gradually started shifting his focus away from literature toward scientific, technical, medical, and other types of nonfiction publishing.
1880	Foundation of <b>Elsevier</b> .
1936	First scientific book published by <b>Elsevier</b> .
1990	First <b>web page</b> .
1991	An online repository of electronic preprints, known as e-prints, of scientific papers is founded in Los Alamos by the American physician Paul Ginsparg. It was renamed to <b>ArXiv.org</b> in 1999. The total number of submissions by March 21st, 2016 (after 24.6 years) is 1,127,232 ( <a href="http://arxiv.org/stats/monthly_submissions">arxiv.org/stats/monthly_submissions</a> ).
1993	Creation of the <b>Open Society Institute</b> (renamed to the Open Society Foundations [OSF] since 2001) by the progressive liberal business magnate George Soros. The OSF financially supports civil society groups around the world, with a stated aim of advancing justice, education, public health and independent media.
1997	Launch of <b>SciELO</b> in Brazil. There are currently 14 countries in the SciELO network and its journal collections: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Mexico, Peru, Portugal, South Africa, Spain, Uruguay, and Venezuela.
1998	<b>Public Knowledge Project</b> (PKP) is founded by John Willinsky in the Faculty of Education at UBC, with Pacific Press Professorship endowment, dedicated to improving the scholarly and public quality of research.  PKP has created the <b>Open Conference Systems</b> (2000), <b>Open Journals Systems</b> (2001), <b>Open Harvester Systems</b> (2002) and the <b>Open Monograph Press</b> (2013).
2000	<b>BioMed Central</b> , the self-described first and largest OA science publisher and <b>PubMed Central</b> , a free digital repository for biomedical and life sciences journal, is founded. In 2008, Springer announces the acquisition of BioMed Central, making it, in effect, the world's largest open access publisher.
2001	An online petition calling for all scientists to pledge that from September 2001 they would discontinue submission of papers to journals which did not make the full-text of their papers available to all, free and unfettered, either immediately or after a delay of several months is released. The petition collected 34,000 signatures but publishers took no strong response to the demands. Shortly thereafter, the <b>Public Library of Science</b> (PLOS) was founded as an alternative to traditional publishing. <i>PLOS ONE</i> is currently the world's largest journal by number of papers published (about 30,000 a year in 2015).  December 1–2: <b>Conference convened in Budapest</b> by the Open Society Institute to promote open access – at the time also known as Free Online Scholarship. Where the Budapest Open Access Initiative (BOAI) was born.
2002	January 14th: Release of the <b>Budapest Open Access Initiative</b> (BOAI), a public statement of principles relating to OA to the research literature. This small gathering of individuals is recognised as one of the major defining events of the OA movement. On the occasion of the 10th anniversary of the initiative, it was reaffirmed in 2012 and supplemented with a set of concrete recommendations for achieving “the new goal that within the next ten years, Open Access will become the default method for distributing new peer-reviewed research in every field and country.”  Start of the Research in Health - <b>HINARI</b> programme of the World Health Organization and major publishers to enable developing countries to access collections of biomedical and health literature online at reduced subscription costs. Together with Research in Agriculture - <b>AGORA</b> , Research in the Environment - <b>OARE</b> and Research for Development and Innovation - <b>ARDI</b> programmes, it currently forms <b>Research4Life</b> that provides developing countries with free or low cost access to academic and professional peer-reviewed content online.
2008	The <b>National Institutes of Health</b> (NIH) Public Access Policy, an OA mandate requiring that research papers resulting from NIH funding must be freely and publicly available through PubMed Central within 12 months of publication, is officially recorded.



YEAR	MILESTONE
2009	The <b>Fair Copyright in Research Works Act</b> (Bill H.R 801 IH, also known as the "Conyers Bill") is submitted as a direct response to the National Institutes of Health (NIH) Public Access Policy; intending to reverse it. The bill's alternate name relates it to U.S Representative John Conyers (D-MI), who introduced it at the 111th United States Congress on February 3, 2009.
2011	Arrest of <b>Aaron Swartz</b> after he systematically downloaded articles from JSTOR, for alleged copyright infringement. In reaction to the high cost of research papers behind paywalls, <b>Sci-Hub</b> , the first known website to provide automatic and free, but illegal, access to paywalled academic papers on a massive scale, is founded by Alexandra Elbakyan from Kazakhstan.
2012	Start of the <b>Academic Spring</b> , a trend wherein academics and researchers began to oppose restrictive copyright in traditional academic journals and to promote free online access to scholarly articles. Start of the <b>Cost of Knowledge</b> campaign which specifically targeted Elsevier. It was initiated by a group of prominent mathematicians who each made a commitment to not participate in publishing in Elsevier's journals, and currently has over 15,933 co-signatories. Start of the United States-based campaign <b>Access2Research</b> in which open access advocates (Michael W. Carroll, Heather Joseph, Mike Rossner, and John Wilbanks) appealed to the United States government to require that taxpayer-funded research be made available to the public under open licensing. This campaign was widely successful, and the directive and FASTR (the Fair Access to Science and Technology Research Act) have become defining pieces in the progress of OA in the USA at the federal level. Launch of <b>PeerJ</b> , an OA journal that charges publication fees through researcher memberships, not on a per-article basis, resulting in what has been called "a flat fee for 'all you can publish'".
2013	January: The suicide of <b>Aaron Swartz</b> gains international attention for the Open Access movement. November: <b>Berlin 11 Satellite Conference</b> for students and early career researchers, which brought together more than 70 participants from 35 countries to engage on Open Access to scientific and scholarly research.
2014	First <b>OpenCon</b> in Washington DC, an annual conference for students and early career researchers on Open Access, open data, and open educational resources. Open Access is embedded the European Commission's <b>Horizon 2020</b> Research and Innovation programme.
2015	Academic publisher <b>Elsevier</b> makes a complaint in New York City for copyright infringement by <b>Sci-Hub</b> . Sci-Hub is found guilty and ordered to shut down. The website re-emerges under a different domain name as a consequence. A second hearing in March 2016 is delayed due to failure of the defendant to appear in court, and to gather more evidence for the prosecution.



**Figure 2. Studies that investigated the citation advantage grouped by their conclusion.** The majority concluded that there is a significant citation advantage for Open Access articles. Source: Data from The Open Access Citation Advantage Service, SPARC Europe, accessed March 2016.



**Table 2. Main scientific papers that have investigated and quantified the citation advantage as well as its origin.**

REFERENCE	DISCIPLINE	CITATION ADVANTAGE	ORIGIN
Antelman (2004)	Mathematics, Electrical Engineering, Political Science, Philosophy	+91%, +51%, +86%, +45% per discipline respectively	NA
Atchison & Bull (2015)	Political Science	Statistically significant citation advantage	NA
Cheng & Ren (2008)	Medicine, Biology, Agricultural Sciences, Chemistry and University Journals	+200%	NA
Davis & Fromerth (2007)	Mathematics	+35%	Quality advantage, no evidence of early advantage
Davis <i>et al.</i> (2008)	Physiology	-5%	NA
Davis (2011)	Sciences, Social Sciences, and Humanities	+1% but statistically indistinguishable	No evidence of an early advantage
Evans & Reimer (2009)	All	+8% for newly published articles; +16% for citations coming from developing countries	NA
Eysenbach (2006)	Natural Sciences	+210 up to +290%	NA
Frandsen (2009)	Biology, Mathematics, Pharmacy and Pharmacology	No clear tendency towards an increase in impact	NA
Gargouri <i>et al.</i> (2010)	Engineering, Biology, Biomedicine, Chemistry, Psychology, Mathematics, Clinical Medicine, Health, Physics, Social Science, Earth Sciences	+?% to ?% depending on the discipline	Quality advantage is confirmed no evidence for selection bias
Gaule & Maystre (2011)	Biology	No evidence of citation advantage	NA
Gentil-Beccot <i>et al.</i> (2010)	High Energy Physics	+200%	Early advantage confirmed
Hajjem <i>et al.</i> (2006)	Biology, Psychology, Sociology, Health, Political Science, Economics, Education, Law, Business, Management	+36% to 172%	NA
Harnad & Brody (2004)	Physics	+250% to 580%	NA
Henneken <i>et al.</i> (2006)	Astronomy and Physics	+200%	NA
Kousha & Abdoli (2010)	Agricultural Science	+621% but not to every journal	NA
Kurtz <i>et al.</i> (2005)	Astronomy	None	Selection bias and early advantage
Kurtz & Henneken (2007)	Astronomy	+200%	Early advantage confirmed
Lansingh & Carter (2009)	Ophthalmology	No	NA
Lawrence (2001)	Computer Science	+157% up to +284% for top publication	NA
McCabe & Snyder (2014)	Ecology, Botany, Multidisciplinary Science and Biology	+8%	NA
McVeigh (2004)	Natural Sciences	Approximately 0-+50% in 2003 depending on field, negative citation advantage in 2000	NA
Metcalfe (2005)	Astronomy	+200%	NA
Metcalfe (2006)	Solar Physics	+170% and +260% depending on the online repository	No evidence for selection bias



REFERENCE	DISCIPLINE	CITATION ADVANTAGE	ORIGIN
Moed (2006)	Condensed Matter Physics	NA	Confirm early access advantage and selection bias but no general OA effect
Norris <i>et al.</i> (2008)	Ecology, Applied Mathematics, Sociology and Economics	+157%	NA
Sahu <i>et al.</i> (2005)	Medicine	+300% up to +450%	NA
Schwarz & Kennicutt Jr (2004)	Astronomy	+200%	Early advantage
Vanclay (2013)	Environmental Science	Not significant	NA
Wang <i>et al.</i> (2015)	All	+111% up to 152%	NA
Wohlrabe & Birkmeier (2014)	Economics	+35% up to 64% depending on the database used	NA
Xu <i>et al.</i> (2011)	Oxford Open Journals	+138.87%	NA
Zhang (2006)	Communication Studies	+200%	NA

publication when compared to OA articles. Additionally, the average number of citations for OA articles was more than double that of the non-OA articles. The study also differentiated the type of OA article, namely the self-archived (i.e., Green OA) and the publisher version of record (VOR) that is freely available (i.e., Gold OA). Gold OA was found to have a higher overall academic impact than Green OA.

Despite strong evidence for a citation advantage, the magnitude of this advantage remains variable. The substantial heterogeneity in observed citation advantages can be due to different academic cultures or could simply be spurious. For example, self-archiving prior to publication is a community standard in fields such as high energy physics or mathematics, but has yet to be widely adopted among the life sciences. Such ‘pre-prints’ have also been associated with an overall increase in the average number of citations, the total number of citations, and the speed of citation accumulation (Aman, 2014; Gentil-Beccot *et al.*, 2010). Other studies could only replicate immense citation advantages (+600%) if relevant predictors were omitted (McCabe & Snyder, 2014), which indicates a potential spurious effect. When taking into account these relevant predictors, the citation advantage became much smaller (i.e., +8%).

One alternative explanation for the citation advantage could be that researchers choose to publish OA when a finding is more impactful, but empirical evidence contradicts this selection effect. Gargouri *et al.* (2010) compared citation counts within a cohort of OA articles that had either been self-selected as OA or mandated as OA (e.g., by funders). The study concluded that both were cited significantly more than non-OA articles. As such, these findings rule out a selection bias from authors as the cause for the citation advantage (Gargouri *et al.*, 2010).

In sum, evidence indicates that OA is broadly related to increased academic impact as assessed through citations (Figure 2; see also

McKiernan *et al.* (2016), but given the large variability in results, further research should aim to synthesize these results in a meta-analysis and try to explain the cause of this variability.

**Societal impact.** Scholarly articles also have a societal impact, such as when they are covered in news media or are discussed in social media outlets; alternative metrics, or altmetrics, can be used as a guide to measure this impact. Article-level information such as social media usage, Mendeley readership, and media attention (Piwowar, 2013) can be tracked by altmetrics providers. As such, when an article generates discussion outside of the academic literature, altmetrics track this usage. Despite limitations such as academics discussing their own research on Twitter, altmetrics provide a general view of the wider societal impact of research articles. Considering the increased pressure on researchers and research institutes to communicate research findings to the public, altmetrics can provide additional insight into which research enjoys public interest. A working group established by NISO is investigating the future role of altmetrics in research communication and assessment ([www.niso.org/topics/tl/altmetrics\\_initiative/](http://www.niso.org/topics/tl/altmetrics_initiative/)).

Open Access articles would be expected to have an altmetrics advantage compared to the non-OA literature; if an article has fewer restrictions to articles for journalists, citizens, businesses, and policy-makers, it seems logical that this would enable the research to be publicly re-used. Furthermore, those parties may be more likely to ‘push’ articles which are publicly accessible into different communication channels. In other words, increased access removes barriers to widespread societal engagement, whereas a relative lack of article access discourages engagement by making it monetarily difficult to access the research.

There is research showing evidence for an altmetrics advantage for Open Access articles, but this does not reflect itself in the most impactful articles. Wang *et al.* (2015) found evidence that OA



articles receive more attention through social media. The authors compared social media attention (Twitter and Facebook) between OA and non-OA articles at *Nature Communications* and found that OA articles get 1.2–1.48 times as much social media attention as compared to non-OA articles [see also [Adie, 2014](#)]. Nonetheless, of the top 100 articles of 2015 as presented by Altmeteric.com, only 42 articles were OA ([www.altmetric.com/top100/2015/](http://www.altmetric.com/top100/2015/)). This 42% is larger than the overall proportion of OA articles in the literature, which indicates that OA contributes relatively more impact per paper. However, it also indicates that the open impact advantage can be overshadowed by the intrinsic nature of the research published.

[Allen et al. \(2013\)](#) found that a social media announcement of the release of a research article increases the number of users who view or download that article, but does not translate to increases in the citation count in the field of clinical pain research. [Costas et al. \(2015\)](#) found a relatively weak correlation between social media activity and citation counts for the articles in their sample (over 1.5 million article records), while [Mohammadi et al. \(2015\)](#) found that the number Mendeley readers with a status of graduate student or faculty correlated with citation counts. When OA to the articles is factored into an analysis, there is a potential recursive relation between citation counts and altmetrics due to OA. [Eysenbach \(2011\)](#) indicated that there is a moderate correlation (0.42–0.72) between the tweets and citations of articles from an OA journal (*Journal of Medical Internet Research*). Highly tweeted articles were eleven times more likely to be highly cited than less-tweeted articles, or vice versa (75% of highly tweeted articles were highly cited; 7% of less-tweeted articles were highly cited). However, it is difficult to assess causality in these cases: do research papers that have more academic impact make their way more frequently into societal discussions, or does increasing online discourse increase their potential citation rates. Overall, this evidence implies that there is a general media advantage with OA (see also [McKiernan et al. \(2016\)](#)), which can be used as a proxy or pathway to indicate greater societal impact.

Altmetrics themselves should not be conflated with citations when it comes to assessing impact, even though some providers such as Altmeteric.com provide a single score that can be used to rank an article in a similar way to a journal's Impact Factor. Each measure of altmetrics tells a different story about the impact of research, and a careful understanding of the altmetrics landscape can lead to a clearer picture of article-level impact.

### Text- and data mining

Traditionally, researchers hand over their copyright via a Copyright Transfer Agreement in order to publish a paper. Copyright transfer as the default has far-reaching consequences on the ability of both the authors and others to re-use that research, and many authors are not aware of the impact of these transfers on their ownership of the work. Academics frequently give the copyright to the publishers in exchange for the 'prestige' of publishing in one of their venues (e.g., [Mueller-Langer & Watt, 2010](#)). Essentially, copyright is a tool wielded by traditional publishers for financial gain rather than fostering creativity, innovation, or protecting authors ([Okerson, 1991](#); [Willinsky, 2002](#)). However, in the digital age

copying is essential to perform necessary research tasks. These activities range from viewing the article (i.e., downloading requires copying) to re-using figures from an article in a book. The interaction of OA and copyright is complex and deserves multiple papers in itself (e.g., [Scheufen \(2015\)](#)). We will highlight how OA views copyright and relate this to its effects on text- and data-mining (TDM).

The majority of 'born Open Access' journals and publishers do not request or receive copyright from authors. Instead, publishers are granted non-exclusive rights to publish, and copyright is retained by authors through a Creative Commons license (typically CC-BY, as befitting the definition of OA from the BOAI). Importantly, this represents a power shift from publisher-owned to author-owned rights to research. This model of author-retained copyright appears to be favoured by the majority (71%) of the research community ([Hoorn & van der Graaf, 2006](#)). Shifting copyright to stay with the author allows for wider re-use, including TDM, and forms the basis for a robust and developing public domain.

As such, copyright in OA publications is non-restrictive and also allows machines to freely access it. In traditional publishing, human reading and computer reading are seen as two separate things which require different agreements, whereas OA publishing views them both in the same, non-restrictive manner. In other words, in order to mine OA journals, one only needs the technical skills to do so. In order to mine traditional closed access journals, one needs to sign or negotiate access conditions, even if legitimate access to the articles has already been bought ([Bloudoff-Indelicato, 2015](#)).

Automated extraction of information from scholarly research via TDM is a methodology that can be applied to investigate the scholarly literature at an enormous scale, creating new knowledge by combining individual findings. This has already proven to be useful for a large variety of applications ([Glenisson et al., 2005](#)). Moreover, OA publishers facilitate TDM on a massive scale by allowing multiple options for collecting the literature needed. For example, PLOS is non-restrictive and allows users to scrape articles directly from the website or using its API. As a result, scraping tools can be used or an application called RPlos, an R package developed to search and download full-text scholarly papers ([Chamberlain et al., 2015](#)).

TDM is not only a knowledge-generation tool; it also allows for automated screening for errors and automated literature searches that renew scientific discovery. With TDM it becomes possible to easily compare one's results with those of the published literature, identify convergence of evidence and enable knowledge discovery ([Natarajan et al., 2006](#)) or discover frequent tentative hypotheses that can be used for new research ([Malhotra et al., 2013](#)). It has already been used to make major advances in fields such as biomedicine ([Gonzalez et al., 2016](#)). TDM also allows for computer applications that can download all scholarly literature given certain search terms (e.g., ContentMine's 'getpapers' tool ([github.com/ContentMine/getpapers](https://github.com/ContentMine/getpapers))), simplifying and shortening the tedious literature search. TDM can also serve a screening purpose similar to plagiarism scanners, helping to detect statistical errors in the scholarly literature (e.g., [Nuijten et al. \(2015\)](#)). TDM can be used in various innovative ways and is an emerging and rapidly



advancing field; non-restrictive licensing through OA promotes its wider application.

Given the exponential increase in the number of scholarly publications, (semi-) automated methods to synthesize results have become increasingly important. TDM decreases the time dedicated to the search for relevant information in scholarly literature by categorizing information (Leitner & Valencia, 2008), highlighting and annotating relevant results to specific users (Shatkay *et al.*, 2008), and profiling research (Porter *et al.*, 2002). Furthermore, TDM also prevents researchers and readers from wasting time on reinventing the wheel simply because one can no longer keep up with the published literature (Harmston *et al.*, 2010).

Because of traditional copyright transfers, TDM has often been stymied by traditional, closed access publishers who frequently see it as a copyright infringement. Researchers using software that harvests data from online publications have been (threatened to be) cut off from accessing the articles. These researchers found themselves trapped in negotiations to resume their research — even though their universities had paid subscription fees for access (e.g., Bloudoff-Indelicato, 2015; Van Noorden, 2012). Standard subscriptions do not permit systematic downloads because publishers fear that their content might be stolen and revenue lost (Van Noorden, 2012). In 2014, Elsevier opened its papers for TDM via a proprietary API (Van Noorden, 2014), but places restrictions on the researcher using the API; however, researchers are not legally required to comply with these restrictions in some countries [e.g., U.K., U.S.A., Handke *et al.*, 2015].

Retrospectively making the enormous corpus of closed access papers publicly available might be possible through legal action at an institutional or governmental level. The position of these institutes regarding copyright transfer remains generally unclear. While academics themselves may have little power in debates regarding copyright, institutes could claim ownership of the work they likely already own by invoking their rights under the work made-for-hire doctrine (Denicola, 2006). This is further supported by Shavell (2010) and Eger & Scheufen (2012) who ascertained that transition towards an OA model could not be smooth without first undertaking the necessary legislative steps. In light of the potential copyright problems for closed access articles, TDM will be easier and legally safer for OA journals. As a consequence, TDM is likely to be more readily applied to OA literature when compared to closed access literature.

## The economic case for Open Access

### The case for publishers

Any publisher has to cover operating costs, which are through (i) article processing; (ii) management and investment; and (iii) other costs. Article processing includes editing, proofreading and typesetting, among other things. Management and investment are the marginal costs to establish and keep the journal running. Other costs include PR, hosting and infrastructural services, conference sponsoring, and other services that are extrinsic to research articles themselves. The average production cost for one paper is estimated to be around \$3500–\$4000 (Van Noorden, 2013) but are highly variable depending on the publisher. For example, Philip

Campbell (Editor-in-Chief of *Nature*) stated that his journal's internal costs were at \$20,000–\$30,000 per paper (Van Noorden, 2013). One possible reason for such variation is that it is unclear whether proposed costs relate to those directly involved in article processing or those required in order for a publisher to 'break even' if they receive zero subscription income for an article made OA.

In order to cover those costs and make a profit or surplus, closed access publishers charge for access via subscriptions, whereas many OA publishers or journals charge to publish. Due to increased subscription costs, closed access publishing is becoming an increasingly unsustainable business model (Odlyzko, 2013) with prices estimated to have increased at 250% of that for inflation ([www.eff.org/issues/open-access](http://www.eff.org/issues/open-access)), which will slowly but surely diminish the scope of access to the scholarly literature as fewer organisations are able to pay such high costs. Only recently has any transparency into the detailed costs of subscriptions been gained by using Freedom of Information Requests to bypass non-disclosure agreements between libraries and publishers (Lawson & Meghreblian, 2015), and provide the basis for understanding the economics of scholarly communication. Two potential ways to prevent future retention of an unsustainable model is through decreasing the subscription prices, thereby lowering publishers' profit margins and the financial burden on subscribers, or switching to new OA-oriented business models and creating new value.

OA publishing is most prevalent in the form of 'pay-to-publish', which can be viewed as a fundamental conflict of interest for researchers. However, this payment model has proven itself to function properly when editorial decisions are separated from the business-side of the publisher (i.e., editorial independence), and many publishers (e.g., PLOS, PeerJ, and many learned societies) operate fee waiver schemes for researchers unable to obtain funds for publication fees.

Subscription-based publishers frequently still produce print versions of journals, which increases their costs and is potentially done in order to justify charging for readership, or satisfy an increasingly small demographic who prefer this mode of reading. After all, subscriptions to print journals make sense and, if large-scale printing is still in place, simply transferring this idea to the digital versions creates continuity. Print versions are accompanied by logistical costs to print and ship each issue, but these are partially offset with reprint orders, additional charges for colour figures, and print-based advertising. For some of the largest subscription-oriented publishers the annual net profit on investment reaches up to 40 percent, which makes academic journal publishing highly lucrative for investors (Satyanarayana, 2013), further increases investment to sustain this type of publishing model, and allows maintenance of a system which has been described as an 'oligopoly' (Larivière *et al.*, 2015).

Open Access publishers only publish digitally and have opened up avenues for innovation. For example, PeerJ has introduced a wholly different OA business model, where readers pay nothing to access articles, but authors pay a membership fee once to publish for a lifetime. The Open Library of Humanities (OLH) is another innovative business model in which libraries pay a small fee to support OLH and scholars are able to publish for free; this support also enables



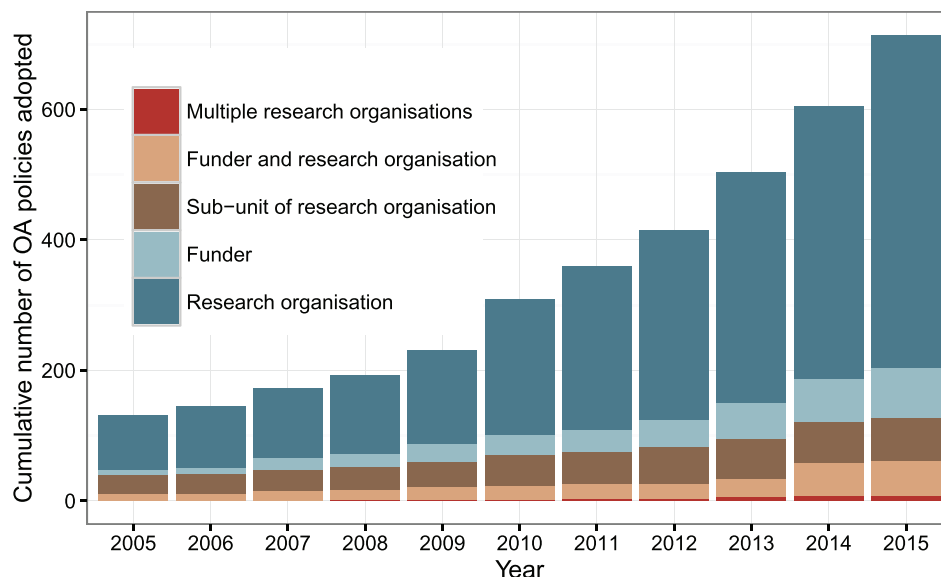
the OLH to help journals ‘flip’ from a subscription model to OA; for example, the recent case of *Lingua* (<https://www.timeshighereducation.com/research-intelligence/open-library-humanities-aims-flip-journals-open-access>). Library publishing has also developed in response to the OA movement; in this model, academic libraries begin publishing operations in the interest of providing added value to their patrons and contributing to the growth of knowledge ([librarypublishing.org](http://librarypublishing.org)). In terms of innovating in the publishing platform itself, *eLife* have introduced the Lens as a novel way of viewing research articles online ([lens.elifesciences.org/about/](http://lens.elifesciences.org/about/)), and *F1000Research* has introduced so-called ‘living figures’ to enable researchers to interact with data underlying research findings [e.g., Colomb & Brembs, 2015]. Such innovations add increased value to the research communication process, contrary to services such as paying to print colour figures, and represent just several cases of a recent explosion in innovation across the publishing ecosystem. One can imagine that publishing costs in OA journals become dependent on the value added on a per-article basis, which can help reshape and improve scholarly communication. As such, making publication costs dependent on the value added aligns the interests of publishers with those of scholars where improving the quality of the process of scholarly communication is the end-goal. The motivation behind this could come from the currently available data that suggests that hybrid publishing options offered by traditional publishers, while being of higher cost, provide a much lower overall quality publishing process ([blog.wellcome.ac.uk/2016/03/23/wellcome-trust-and-coaf-openaccess-spend-2014-15/](http://blog.wellcome.ac.uk/2016/03/23/wellcome-trust-and-coaf-openaccess-spend-2014-15/)).

### The case for non-publishers

The implementation of OA models has implications beyond the publishing industry in terms of economics. Research funding comes from multiple sources, including, national funding agencies

and industries, as well as private funders. Much primary research actually takes place outside of academia inside research & development departments. A report from 2004 by Arzberger and colleagues into the scientific, social and economic development of access to research results concluded that access should be promoted to the largest extent possible. According to this report, access to research results can only be responsibly restricted in the case of national security, privacy, or those involving IP rights of the authors (Arzberger *et al.*, 2004). A major principle underlying this is the ownership of research results: publicly funded research and data are public goods and because they have been produced in the public interest they should be considered and maintained as such. Indeed, such a principle has become one of the focal rallying points of the global OA movement. Appropriate licensing and accessibility can influence re-use through commercialization, and can empower citizens and industry to recognize great economic benefits. This apparently resonates with many organisations, as indicated by the increased numbers of OA policies on a global basis (see Figure 3).

With access to scholarly articles, entrepreneurs and small businesses can accelerate innovation and discovery, which is advantageous for advancing the ‘entrepreneurial state’ (Mazzucato, 2011). Access to research results has clear advantages for a range of industries and can help stimulate regional and global economies. Increased access to research results has been associated with considerable increases of return on financial investment (Beagrie & Houghton, 2014). Furthermore, OA facilitates collaborations between publishers and industrial partners to leverage the potential of structured information networks for advanced data mining projects, such as that recently announced between IBM Watson and PLOS (Denker, 2016). One of the major driving forces behind the development of OA in the UK on a national level, the ‘Finch Report’, also concluded that



**Figure 3.** Cumulative number of Open Access policies adopted by multiple research organisations, funder and research organisation, sub-unit of research organisation, funder, and research organisation. Figures are given at the beginning of each year. Source: ROARMAP, accessed March 2016.



OA was an essential source for information and innovation to the civil service, commercial sectors, small- and medium-sized enterprises (SMEs), and the general public ([www.researchinfonet.org/publish/finch/](http://www.researchinfonet.org/publish/finch/)).

Taking cancer research as one high impact case study, there is substantial evidence for the economic benefit of OA. In 2011–12 prices, the total expenditure on research relating to cancer in the period of 1970–2009 was £15 billion (Glover *et al.*, 2014). 5.9 million quality adjusted life years were gained from the prioritized interventions in 1991–2010, of which the net-monetary benefit was an estimated £124 billion (i.e., eight-fold return on investment). However, only 17% of the annual net-monetary was estimated to be attributable to research performed in the UK (Glover *et al.*, 2014), suggesting that 83% of the economic return on cancer research is drawn from research from non-UK sources. As such, opening up research for global access rather than localized and restricted use has the potential to increase the economic return, as demonstrated with the case on cancer research.

### The price of Open Access

The question of the current publication cost is difficult and confounded by estimates of the total global publishing costs and revenue. Data provided by Outsell, a consultant in Burlingame, California, suggest that the science-publishing industry generated \$9.4 billion in revenue in 2011 and published around 1.8 million English-language articles. This equates to an approximate average revenue per article of \$5,000. A white paper produced by the Max Planck Society estimated costs at €3,800–€5,000 per paper through subscription spending, based on a total global spending of €7.6 billion across 1.5–2 million articles per year in total (Schimmer *et al.*, 2014). Other estimates suggest that the total spend on publishing, distribution and access to research is around £25 billion per year, with an additional £34 billion spent on reading those outputs, a sum which equates to around one third of the total annual global spend on research (£175 billion; Research Information Network (2008)).

Such high costs are at odds with estimates of the cost of OA publishing. For example, the Scientific Electronic Library Online (SciELO) is a pan-Latin American bibliographic database, digital library, and co-operative electronic publishing model of Open Access journals. It is estimated that their costs are between \$70 and \$600 per OA article depending on the services provided (Brembs, 2015). OA now dominates the Latin American publishing landscape, with an estimated 72–85% of articles now with full text OA articles publicly available ([www.sparc.arl.org/news/open-access-latin-americaembraced-key-visibility-research-outputs](http://www.sparc.arl.org/news/open-access-latin-americaembraced-key-visibility-research-outputs)). Even more extreme estimates of the cost of OA come from Standard Analytics, who suggested the absolute minimum per-article costs of publishing could fall to between \$1.36 and \$1.61 with sufficient cloud-based infrastructure (Bogich *et al.*, 2016). However, it is likely that this estimate under-emphasizes marginal costs that are beyond a per-article cost basis. What is clear from these analyses is that OA has the opportunity to become a cost-reducing mechanism

for scholarly publishing. Open Journals System (OJS), an open source software available to anyone to use and download without charge, is another example of this. Additionally, researcher-led initiatives such as the recently launched *Discrete Analysis* have costs that average around \$30 per article, with no cost to authors or readers, and utilise the infrastructure offered by the arXiv to keep costs low ([discreteanalysisjournal.com](http://discreteanalysisjournal.com)).

In his article, Sutton (2011) argued that current scholarly journals are digital products and that as such they are driven by very different economic principles and social forces than their print ancestors. Based on Anderson (2013), the author made the case that changes in both the delivery of scientific content and in publishers' business models was inevitable when journals moved online. Sutton (2011) considered that scientific literature is no different from other digital products with respect to distribution costs and as such it is no exception to the 'zero is inevitable' rule of pricing.

### The societal case for Open Access

OA to the scholarly literature does not just benefit academics, but also has wider impacts on other domains in society. It makes research available to anyone with an Internet connection who has the ability to search and read the material. Therefore, it transcends academic affiliation and supports sustainable lifelong learning. Examples of groups who might benefit most from OA include citizen scientists, medical patients and their supporting networks, health advocates, NGOs, and those who work in linguistics and translation. In theory, OA affects anyone who uses information and opens up possibilities for knowledge to be used in unexpected, creative and innovative ways beyond mainstream professional research.

Access to knowledge has been called a human rights issue, considering it is included in Article 27 of the United Nations Declaration of Human Rights. Willinsky (2006) has argued that "*Access to knowledge is a human right that is closely associated with the ability to defend, as well as to advocate for, other rights*". This is not only true for access to knowledge from research that could save human lives, but also, as argued by Jacques Derrida, to the right of access to philosophy and the humanities disciplines that stem from it. Derrida writes about the field of Philosophy, "*No one can forbid access to it. The moment one has the desire or will for it, one has the right to it. The right is inscribed in philosophy itself*" (Derrida, 2002).

Society's ability to make research publicly accessible supports the long-term interest and investment in research. Citizens support research through taxes and therefore one could argue that efforts to support public access should be a fundamental part of the research process. While OA is not a solution to all aspects of research accessibility (e.g., language barriers and disability access remain continuing issues to be addressed), it most certainly increases accessibility greatly and simultaneously allows innovations to remove other barriers (e.g., OA articles can be freely translated to address language barriers and can be changed to different formats to accommodate screen readers). Anecdotal evidence suggests that public access to research is required from a range of public spheres



([whoneedsaccess.org/](http://whoneedsaccess.org/)). Nonetheless, the fact that access to knowledge is actively prohibited in fields like public health should be of major concern to all stakeholders engaged in academic publishing.

In addition to professional research by, for example, academics, there is the dimension of citizen science. In citizen science, the broader public participates in the research process itself. Numerous projects such as Galaxy Zoo, Zooniverse, Old Weather, Fold It, Whale FM, and Bat Detective are all different initiatives in which citizens publicly and openly engage with research. These initiatives introduce new ways of knowledge creation and these groups also require thorough access to actually be able to do non-redundant research. Citizen science forms part of the societal case for Open Access, because it indicates anyone can be engaged with research, not only professional researchers.

Some traditional publishers, and some academics, have argued that public access to research is not required because research papers cannot be understood by non-specialists. However, citizen science initiatives already indicate the general public *is* interested and understands the research. Whereas this is hyper-variable, and strongly dependent on a range of factors, it is the fact that any public interest in science that is of importance. These publishers and academics argue that specialization is sufficient reason for confining access to professional research bodies through subscriptions. Such statements conflate a lack of desire or need for access with the denial of opportunity to access research. Isolated incidents such as the crashing of servers of Physical Reviews Letters upon the ‘Gravitational Waves’ announcement and OA publication (Feb, 2016; [Abbott et al. \(2016\)](#)) indicate that there are cases of extreme public interest in science that closed access would only impede. Furthermore, claims that only experts can and should read research articles does little to break down the ‘ivory tower’ perception that still pervades academia, and undermines the enormous amounts of resources invested in science communication and public engagement activities. Such perceptions run counter to the idea of access to knowledge as a right, retaining it as a privilege based on financial or academic status.

### Open Access in developing countries

The arguments outlined above form the basis for democratic and equal access to research, which come to light even stronger in the developing world. For low- and middle-income countries (LMIC), OA publishing breaks traditional financial barriers and allows unrestricted, equal access to scholarly information to people all over the globe. Due to the high prices of journal subscriptions, developing countries struggle with access just as in developed countries, but to a greater extent and consequently with greater negative repercussions. For example, a research paper from 1982 that indicated Liberia should be included in the Ebola endemic zone was unknown to Liberian officials in the 2014 Ebola outbreak ([Knobloch et al., 1982](#)), and paywalled. In general, lack of access can have major deleterious consequences for students and researchers in that they do not have sufficient material to conduct their own primary research or education.

OA provides a mechanism to level the playing field between developed and developing countries by providing equal access. This

increases fair competition and increases the scientific potential of the developing world ([Chan et al., 2005](#)). This is linked to the wider issue of open licensing, which is essential for effective marketing of medicines and medical research in developing countries ([Flynn et al., 2009](#)), and justifies the necessity of Open Access in the wider context of social welfare. Developing countries clearly acknowledge the need for access and as such have launched many repositories to increase access with self-archiving of research articles. In 2014, over 100 institutions in Africa launched a network of over 25 fully-operational OA repositories in Kenya, Tanzania and Uganda ([www.ubuntunet.net/april2014#researchrelevant](http://www.ubuntunet.net/april2014#researchrelevant)). Such developments suggest that African nations are leaning more towards a ‘green’ model of OA adoption.

The shift from a ‘reader pays’ to an ‘author pays’ model with OA potentially limits its adoption in developing countries. The pay-to-publish system is a potentially greater burden for authors in developed countries, considering that they are not used to paying publication costs, and funding systems for OA are not as well-established as those in the Western world. Publication fees present an even greater relative burden ([Matheka et al., 2014](#)) given that they can often exceed a monthly salary. This has been at least partially mitigated with waiver fees for authors from developing countries and additional provisions in research grants. In November 2015, Research4Life ([www.research4life.org](http://www.research4life.org)) and DOAJ announced a working partnership that will help to ensure that the Research4Life users will have access to the largest possible array of OA journals from publishers with a certain quality standard. While Research4Life does not directly cover OA publication costs, a lot of publishers propose full or partial waivers if they are based in countries eligible by Research4Life. However, determining which countries qualify for access to scientific journals through these programs, and which journals they are provided access to, is a fairly closed process. They are also not entirely stable, as publishers can opt out of the initiative, or be selective about which countries they choose to serve. In 2011, publishers withdrew free access to 2500 health and biomedicine journals for Bangladesh ([Kmietowicz, 2011](#)) through the HINARI programme. While access was subsequently reinstated, this demonstrates that such initiatives are not an adequate replacement for full OA ([Chatterjee et al., 2013](#)). Despite these programs purporting to provide essential articles to researchers in poor nations, they exclude some developing countries (e.g., India) and limit access to researchers who work in registered institutions.

Initiatives such as the Journals Online Project developed by INASP (International Network for the Availability of Scientific Publications; [www.inasp.info/en/](http://www.inasp.info/en/)) has helped to develop a number of online OA platforms in the Global South. These were launched in 1998 with the African Journals Online (AJOL) platform, a project currently managed in South Africa. More recently, INASP have set up Latin American Journals Online (LAMJOL) which hosts journals in El Salvador, Honduras, and Nicaragua. In Asia, Bangladesh Journals Online (BanglaJOL), Nepal Journals Online (NepJOL), and Sri Lankan Journals Online (SLJOL), all facilitated through INASP, continue to develop and now around 95% of their articles are full-text Open Access. As mentioned previously, improved



access should not be limited to professional researchers only, considering that there is also global interest from the broader public, including health professionals.

### Predatory publishers

One negative effect on OA comes from entities that attempt to profit by exploiting the pay-to-publish system OA publishers use. These publishers operate a sub-category of OA journals known as vanity presses, predatory publishers (Beall, 2012) or pseudo-journals (McGlynn, 2013). These journals seem to be in the scholarly publishing business primarily to collect publication fees (i.e., APCs) in exchange for rapid publication without formal peer-review. Beall (2015) has defined a list of criteria for identifying predatory journals and an index of publishers and individual journals that meet these criteria is continuously updated (scholarlyoa.com).

While not all scholars and advocates agree with the criteria proposed by Jeffrey Beall, there are several factors that many agree on to identify a predatory publisher. Predatory publishers tend to charge low publication fees (Xia, 2015), most below \$100 and few charge more than \$200. On the contrary, the average publication fee of journals indexed in the Directory of Open Access Journals (DOAJ) is around \$900–\$1,000 (Solomon & Björk, 2012) and leading universities in the UK and Germany pay on average \$1,200–\$1,300 per article (Schimmer *et al.*, 2015). The editorial and peer-review aspects of predatory publishers are either non-existent or suspect; they also falsely claim to have ratings such as a Journal Impact Factor and to be indexed in major databases such as Scopus (Djuric, 2015). Editors from these journals solicit articles that have no relation to the topic of their journal and do not send the documents out to be properly peer-reviewed (Bowman, 2014).

The problem of predatory OA seems to highly affect countries where the academic evaluation strongly favors international publication without further quality checks (Shen & Björk, 2015). Xia *et al.* (2015) collected and analyzed the publication record, citation count, and geographic location of authors from the various groups of journals. Statistical analyses verified that predatory and non-predatory journals have distinct author populations: authors who publish in predatory journals tend to be inexperienced early-career researchers from developing countries. The spatial distribution of both the predatory publishers and those authors who submit in pseudo-journals is highly skewed: Asia and Africa contributes three quarters of authors (Xia *et al.*, 2015) and Indian journals form the overwhelming proportion of predatory publishers (Xia, 2015). An interesting finding is the very low involvement of South America, both among predatory publishers (0.5 %) and corresponding authors in predatory journals (2.2%). The OA infrastructure in Latin America is different compared to other developing countries, which reveals a possible reason for this asymmetric situation. Latin American journals and universities are engaged in OA publication models at a higher degree than other regions (Alperin *et al.*, 2011). As a result, scholars from this region are not only more aware of OA issues, but they have more options for publishing OA than those from other regions (Alperin *et al.*, 2011). Moreover, SciELO (Packer, 2009) and the creation of Latin American databases (Octavio Alonso-Gamboa & Russell, 2012) have played

a tremendous part in this process by bringing recognition and a good reputation to publishing outlets in Latin America.

Considerable attention is given to the subject of predatory publishers, who have become conflated with the OA movement in general to the detriment of genuine OA publishers. For example, a ‘sting’ operation that outed bad peer-review instead got misinterpreted as bad peer-review in OA journals (Bohannon, 2013), but was probably more indicative of issues to do with the traditional closed and over-burdened system of peer review ([www.scilogs.com/communication\\_breakdown/jon-tennant-oa/](http://www.scilogs.com/communication_breakdown/jon-tennant-oa/)). Overall, the predatory publisher phenomenon is one major negative aspect that spawns many misconceptions and misgivings about publishing OA. Recently launched industry-led initiatives such as “Think, Check, Submit” ([thinkchecksubmit.org](http://thinkchecksubmit.org)) provide a checklist to help researchers identify trustworthy journals, and will likely be a pivotal tool in combating predatory publishers.

### Open Access and ‘open research’

OA exists in a constantly evolving scholarly research ecosystem. As such, it is important to note how it is interconnected to other facets of the scholarly communication system. Here, we discuss the implications that the transition to OA has on developments in the broader context of ‘Open Data’ and ‘Open Science’, or ‘Open Research’.

### Open Access and the case for Open Data

The overall movement of OA has become conjoined with the push for Open Data. Data sharing is fundamental to scientific progress, because data lead to the knowledge generated in research articles. Furthermore, data sharing has recently become a common requirement, together with OA, for both research funding and publication. The data sharing policy from PLOS illustrates the high degree of overlap between OA and Open Data; authors of articles published in PLOS are required to share the data except if they have valid reasons not to (i.e., an opt-out system; [journals.plos.org/plos-one/s/data-availability](http://journals.plos.org/plos-one/s/data-availability)).

The benefits of Open Data are diverse, including a citation advantage. Combined with the citation advantage for OA articles, providing data alongside publications can increase citations on average by 30% (Piwowar & Vision, 2013) and up to 69% (Piwowar *et al.*, 2007), but this evidence is entirely field-dependent [e.g., Dorch *et al.*, 2015]. Below we cover six additional benefits of Open Data.

First, data sharing enhances reproducibility, a crucial aspect in a time where some scientific domains appear to have problems with reproducibility [e.g., Open Science Collaboration, 2015]. Several factors could form the basis for this ‘crisis’, such as an overemphasis on novelty instead of rigour, selective reporting of results, an overemphasis on statistical significance, and insufficient documentation of the research methods. Publicly sharing data, code, and materials can alleviate issues with reproducibility. This is especially pertinent in the modern sciences, where a substantial proportion of published results draw on quantitative experiments and computer simulations. As such, it is largely impossible to reproduce these experiments as they become more complex and associated datasets increase in complexity. When full access to the data, metadata, and the code



used to produce the ultimate results are provided alongside publication, this greatly improves reproducibility.

Second, publicly available data can be used to stimulate innovations, such as new analytical methods. An excellent example of this is provided by the neuroimaging OpenfMRI project, where shared data have been used to examine the effects of different processing pipelines on analysis outcomes (Carp, 2012) and test new methods to characterize different cognitive tasks (Turner & Laird, 2012). Another good example is the Protein Data Bank (PDB) (Berman *et al.*, 2000), a project which has enabled the re-use of the primary structural data and opened up new avenues of research, despite the latter not being expected.

Third, data sharing enables new research questions that can only be answered by combining datasets, which now remain separated. Analyzing vast volumes of data can yield novel and perhaps surprising findings. This allows for integrated research hypotheses on the underlying processes behind the original data and observations. Exploratory approaches to large datasets can be seen as hypothesis generating tools, which later fuel hypothesis testing to confirm or disprove these hypotheses (Wagenmakers *et al.*, 2012).

Fourth, the realization that data will ultimately be shared and visible to the community provides a strong incentive for researchers to ensure they engage in better data documentation and, therefore, research methods. For example, the willingness to publicly share data has been associated with fewer statistical errors in the final research article (Wicherts *et al.*, 2011).

Fifth, public data sharing provides a digital backup for datasets, protecting valuable scientific resources. Moreover, a considerable amount of data produced every day does not ultimately lead to publication and often remain hidden. Such data might remain in a hidden file-drawer despite being valid, creating a systematic bias in the information available. Public data sharing opens this file-drawer and, consequently, allows independent assessments of whether the data are valid or not.

Finally, and tightly connected with the fifth point, sharing data can certainly reduce the cost of performing research. A file-drawer has been indicated to greatly reduce the efficiency of research in detecting effects (van Assen *et al.*, 2014). Open Data, as such, discourages redundant data collection (i.e., data that have been already collected but never made publicly accessible) and simultaneously allows researchers to better approximate what is happening in their fields. This will have a large effect on research costs, resulting in savings that can be then be used for more productive research goals.

### Open Access and Open Science

Beyond OA and Open Data lies a more integrated approach to research, referred to more broadly as Open Science (i.e., Science 2.0, Open Scholarship). According to the European Commission's Horizon 2020 programme, Open Science is defined as "*The transformation, opening up and democratisation of science and research through ICT, with the objectives of making science more efficient, transparent and interdisciplinary, of changing the interaction between science and society, and of enabling broader societal impact and*

*innovation*". Consequently, we see OA as only one of the multiple challenges currently facing the 'open transformation' of the scholarly publishing system (Watson, 2015), and should be considered in the wider contexts and complimentary domains of research transparency and open source.

As Kriesekorte *et al.* (2012) pointed out, OA is now widely accepted as desirable and becoming a reality in many academic spheres. However, the second essential complementary element to research – evaluation – has received less attention despite the large amount of research that has been done to document its current limitations (Benos *et al.*, 2007; Birukou *et al.*, 2011; Ioannidis, 2005; Ioannidis, 2012a; Ioannidis, 2012b; John *et al.*, 2012; Nosek & Bar-Anan, 2012; Simmons *et al.*, 2011). Open evaluation, an ongoing post-publication process of transparent peer review and rating of papers, promises to address the problems of the current assessment systems Kriesekorte *et al.* (2012). Future research regarding better ways to improve scholarly communication will be instrumental in providing evidence to support the transformation of the publishing system and design new alternatives (Buttliere, 2014; Ghosh *et al.*, 2012; Kriesekorte *et al.*, 2012; Pöschl, 2012), which will draw heavily upon an open publishing framework driven by developments and newly emerging models in OA.

### Conclusions

This review presents published evidence of the impact of Open Access on the academy, economy and society. Overall, the evidence points to a favorable impact of OA on the scholarly literature through increased dissemination and reuse. OA has the potential to be a sustainable business venture for new and established publishers, and can provide substantial benefits to research- and development-intensive businesses, including health organisations, volunteer sectors, and technology. OA is a global issue, highlighted by inequalities beset at all levels between developing and developed nations, and largely fueled by financial inequality. Current levels of access in the developing world are insufficient and unstable, and only OA has the potential to foster the development of stable research ecosystems. While predatory publishing remains an ongoing issue, particularly in the developing world, increasing public engagement, development of OA policies, and discussion of sustainable and ethical publishing practices can remove this potential threat to OA.

For libraries, universities, governments, and research institutions, one major benefit of lowering the cost of knowledge is a budget that allows them to spend their resources more wisely. For researchers themselves, the goal is to increase their audience and impact by delivering wider and easier access for readers. For publishers, promoting OA is a reaction to the wants and needs of their community. Furthermore, subscription-based publishers have (partly) answered the call of the an increasing global demand for OA by giving their green light to author self-archiving (Harnad *et al.*, 2008), as well as through establishing numerous 'hybrid' OA options. In an author survey, Swan & Brown (2004) reported that the vast majority of their sample indicated that they would self-archive willingly if their employer (or funding body) required them to do so. Similarly, in a study by Swan & Brown (2005) the vast majority of researchers (81%) indicated that they would comply with mandates that made OA a condition of funding or employment. Fortunately, it seems that funders and research organisations are moving in that



direction. Since 2005, the number policies supporting OA publishing increased steadily. Consequently, it is now the responsibility of researchers to ensure OA to their publications either by choosing the green or the gold road.

The fact that OA impacts upon such a diverse range of stakeholders, often with highly polarised and emotional viewpoints, highlights the ongoing need for evidence-informed discussion and engagement at all levels. As Peter Suber, a leading voice in the OA movement, stated: “*As long as they do not have the power to stop Open Access, the toll-access publishers are not the enemy*”. Therefore, Open Access supporters should focus their efforts on working for new models and systems rather than trying to undermine or punish the existing ones. OA remains only one of the multiple challenges that the scholarly publishing system is currently facing. As highlighted in this review, the empirical evidence for OA is overwhelmingly positive. Further research needs to move from investigating the effects of OA to researching the broader effects of Open Science.

### Author contributions

All authors contributed equally to the writing of this manuscript using the Overleaf collaborative writing platform.

### Competing interests

JPT currently blogs for the PLOS Paleo Community, and works for ScienceOpen. CHJH is a Center for Open Science ambassador.

DCJ and FW are members of the Open Access Working Group of EURODOC. PM is a Research Data Alliance member. LBC works for the University of Pittsburgh, which has an Open Access library publishing department. All views presented here are strictly personal.

### Grant information

This research was partly funded by the Belgian National Fund for Scientific Research through a FRIA grant. PM acknowledges support from the European Commission Horizon 2020 Programme under Grant Agreement 634107 (PHC32-2014) ‘MULTIMOT’.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### Acknowledgments

We would like to collectively acknowledge the OpenCon community for inspiring this paper, and for providing continuous discussion about the various aspects of Open Access. In particular, we are grateful to Brett Buttlere, Audrey Risser, Sarah Barkla, and April Clyburne-Sherin for contributing resources to the development of this paper, and Tracey Depellegrin Connelly, Matt Menzenski, and Joseph McArthur for helpful comments on an earlier draft. We also thank Neil Saunders who provided the base code to extract data from PubMed Central. We would also like to thank Andy Nobes for drawing our attention to the work of INASP. PM would like to thank Lennart Martens for insightful discussions on Open Science.

## References

- Abbott BP, Abbott R, Abbott TD, *et al.*: **Observation of Gravitational Waves from a Binary Black Hole Merger.** *Phys Rev Lett.* 2016; **116**(6): 061102.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Adie E: **Attention! a study of open access vs non-open access articles.** *Figshare.* 2014.  
[Publisher Full Text](#)
- Allen HG, Stanton TR, Di Pietro F, *et al.*: **Social media release increases dissemination of original articles in the clinical pain sciences.** *PLoS One.* 2013; **8**(7): e68914.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Alonso-Gamboa JO, Russell JM: **Latin American scholarly journal databases: a look back to the way forward.** In *Aslib Proceedings*. Emerald Group Publishing Limited, 2012; **64**(1): 32–45.  
[Publisher Full Text](#)
- Alperin JP, Fischman GE, Willinsky J: **Scholarly communication strategies in Latin America's research-intensive universities.** *Educación superior y sociedad.* 2011; **16**(2).  
[Reference Source](#)
- Aman V: **Is there any measurable benefit in publishing preprints in the arxiv section quantitative biology?** *CoRR.* abs/1411.1955. 2014.  
[Reference Source](#)
- Anderson C: **Free: How today's smartest businesses profit by giving something for nothing.** Random House, 2013.  
[Reference Source](#)
- Antelman K: **Do open-access articles have a greater research impact?** *Coll Res Libr.* 2004; **65**(5): 372–382.  
[Publisher Full Text](#)
- Arzberger P, Schroeder P, Beaulieu A, *et al.*: **Promoting access to public research data for scientific, economic, and social development.** *Data Sci J.* 2004; **3**: 135–152.  
[Publisher Full Text](#)
- Atchison A, Bull J: **Will open access get me cited? an analysis of the efficacy of open access publishing in political science.** *PS Polit Sci Polit.* 2015; **48**(01): 129–137.  
[Publisher Full Text](#)
- Beagrie N, Houghton JW: **The value and impact of data sharing and curation: A synthesis of three recent studies of UK research data centres.** 2014.  
[Reference Source](#)
- Beall J: **Criteria for determining predatory open-access publishers.** *Scholarly Open Access.* 2015.  
[Reference Source](#)
- Beall J: **Predatory publishers are corrupting open access.** *Nature.* 2012; **489**(7415): 179.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Benos DJ, Bashari E, Chaves JM, *et al.*: **The ups and downs of peer review.** *Adv Physiol Educ.* 2007; **31**(2): 145–152.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Berman HM, Westbrook J, Feng Z, *et al.*: **The protein data bank.** *Nucleic Acids Res.* 2000; **28**(1): 235–242.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Birukou A, Wakeling JR, Bartolini C, *et al.*: **Alternatives to peer review: novel approaches for research evaluation.** *Front Comput Neurosci.* 2011; **5**: 56.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bjork B, Roos A, Lauri M: **Scientific journal publishing: yearly volume and open access availability.** *Inform Res.* 2009; **14**(1).  
[Reference Source](#)
- Bloudoff-Indelicato M: **Text-mining block prompts online response.** *Nature News.* 2015; **527**(7579): 413.  
[Publisher Full Text](#)
- Bogich T, Ballesteros S, Berjon R, *et al.*: **On the marginal cost of scholarly communication.** 2016.  
[Reference Source](#)



- Bohannon J: **Who's afraid of peer review?** *Science*. 2013; **342**(6154): 60–5.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bowman JD: **Predatory publishing, questionable peer review, and fraudulent conferences.** *Am J Pharm Educ*. 2014; **78**(10): 176.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Brembs B: **What goes into making a scientific manuscript public?** 2015; Accessed: 2016-3-24.  
[Reference Source](#)
- Buttliere BT: **Using science and psychology to improve the dissemination and evaluation of scientific work.** *Front Comput Neurosci*. 2014; **8**: 82.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Carp J: **On the plurality of (methodological) worlds: estimating the analytic flexibility of fmri experiments.** *Front Neurosci*. 2012; **6**: 149.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chamberlain S, Boettiger C, Ram K: **rplos: Interface to the Search 'API' for 'PLOS' Journals.** 2015.  
[Reference Source](#)
- Chan L, Kirsop B, Arunachalam S: **Open access archiving: the fast track to building research capacity in developing countries.** 2005.  
[Reference Source](#)
- Chatterjee P, Biswas T, Mishra V: **Open access: the changing face of scientific publishing.** *J Family Med Prim Care*. 2013; **2**(2): 128–30.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cheng W, Ren S: **Evolution of open access publishing in Chinese scientific journals.** *Learn Publ*. 2008; **21**(2): 140–152.  
[Publisher Full Text](#)
- Colomb J, Brembs B: **Sub-strains of *Drosophila* Canton-S differ markedly in their locomotor behavior.** *F1000Res*. 2015; **3**: 176.  
[Publisher Full Text](#)
- Costas R, Zahedi Z, Wouters P: **Do "altmetrics" correlate with citations? extensive comparison of altmetric indicators with citations from a multidisciplinary perspective.** *J Assoc Inf Sci Technol*. 2015; **66**: 2003–2019.  
[Publisher Full Text](#)
- Davis PM, Fromerth MJ: **Does the arxiv lead to higher citations and reduced publisher downloads for mathematics articles?** *Scientometrics*. 2007; **71**(2): 203–215.  
[Publisher Full Text](#)
- Davis PM, Lewenstein BV, Simon DH, *et al.*: **Open access publishing, article downloads, and citations: randomised controlled trial.** *BMJ*. 2008; **337**: a568.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davis PM: **Open access, readership, citations: a randomized controlled trial of scientific journal publishing.** *FASEB J*. 2011; **25**(7): 2129–2134.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Denicola R: **Copyright and open access: reconsidering university ownership of faculty research.** *Nebraska Law Review*. 2006; **85**(2).  
[Reference Source](#)
- Denker SP: **Collaboration with IBM Watson supports the value add of open access.** 2016.  
[Reference Source](#)
- Derrida J: **Who's afraid of philosophy?: Right to philosophy 1.** Stanford University Press, 2002; **1**.  
[Reference Source](#)
- Djuric D: **Penetrating the omerta of predatory publishing: the Romanian connection.** *Sci Eng Ethics*. 2015; **21**(1): 183–202.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Dorch SBF, Drachen TM, Ellegaard O: **The data sharing advantage in Astrophysics.** *arXiv*. 2015.  
[Reference Source](#)
- Eger T, Scheufen M: **The past and the future of copyright law: Technological change and beyond.** *Liber Amicorum Boudewijn Bouckaert*, forthcoming. 2012; 37–64.  
[Reference Source](#)
- Evans JA, Reimer J: **Open access and global participation in science.** *Science*. 2009; **323**(5917): 1025.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Eysenbach G: **Citation advantage of open access articles.** *PLoS Biol*. 2006; **4**(5): e157.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Eysenbach G: **Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact.** *J Med Internet Res*. 2011; **13**(4): e123.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Flynn S, Hollis A, Palmedo M: **An economic justification for open access to essential medicine patents in developing countries.** *J Law Med Ethics*. 2009; **37**(2): 184–208.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Frandsen TF: **The integration of open access journals in the scholarly communication system: Three science fields.** *Inf Process Manag*. 2009; **45**(1): 131–141.  
[Publisher Full Text](#)
- Gargouri Y, Hajjem C, Larivière V, *et al.*: **Self-selected or mandated, open access increases citation impact for higher quality research.** *PLoS One*. 2010; **5**(10): e13636.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gaule P, Maystre N: **Getting cited: does open access help?** *Research Policy*. 2011; **40**(10): 1332–1338.  
[Publisher Full Text](#)
- Gentil-Beccot A, Mele S, Brooks T: **Citing and reading behaviours in high-energy physics.** *Scientometrics*. 2010; **84**(2): 345–355.  
[Publisher Full Text](#)
- Ghosh SS, Klein A, Avants B, *et al.*: **Learning from open source software projects to improve scientific review.** *Front Comput Neurosci*. 2012; **6**: 18.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Glenisson P, Glänzel W, Janssens F, *et al.*: **Combining full text and bibliometric information in mapping scientific disciplines.** *Inf Process Manag*. 2005; **41**(6): 1548–1572.  
[Publisher Full Text](#)
- Glover M, Buxton M, Guthrie S, *et al.*: **Estimating the returns to UK publicly funded cancer-related research in terms of the net value of improved health outcomes.** *BMC Med*. 2014; **12**(1): 99.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gonzalez GH, Tahsin T, Goodale BC, *et al.*: **Recent Advances and Emerging Applications in Text and Data Mining for Biomedical Discovery.** *Brief Bioinform*. 2016; **17**(1): 33–42.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hajjem C, Harnad S, Gingras Y: **Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact.** 2006; arXiv preprint cs/0606079.  
[Reference Source](#)
- Handke C, Guibault L, Vallbé JJ: **Is Europe falling behind in data mining? Copyright's impact on data mining in academic research.** 2015.  
[Publisher Full Text](#)
- Harmston N, Filsell W, Stumpf MP: **What the papers say: text mining for genomics and systems biology.** *Hum Genomics*. 2010; **5**(1): 17–29.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Harnad S, Brody T: **Comparing the impact of open access (OA) vs. non-OA articles in the same journals.** *D-lib Magazine*. 2004; **10**(6).  
[Publisher Full Text](#)
- Harnad S, Brody T, Vallieres F, *et al.*: **The access/impact problem and the green and gold roads to open access: An update.** *Serials Review*. 2008; **34**(1): 36–40.  
[Publisher Full Text](#)
- Henneken EA, Kurtz MJ, Eichhorn G, *et al.*: **Effect of e-printing on citation rates in astronomy and physics.** *arXiv*. 2006.  
[Reference Source](#)
- Hitchcock S: **The effect of open access and downloads ('hits') on citation impact: a bibliography of studies.** 2013.  
[Reference Source](#)
- Hoorn E, van der Graaf M: **Copyright issues in open access research journals: The authors perspective.** *D-Lib Magazine*. 2006; **12**(2): 6.  
[Publisher Full Text](#)
- Ioannidis JP: **Why most published research findings are false.** *PLoS Med*. 2005; **2**(8): e124.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ioannidis JP: **Why Science Is Not Necessarily Self-Correcting.** *Perspect Psychol Sci*. 2012a; **7**(6): 645–654.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ioannidis JP: **Scientific communication is down at the moment, please check again later.** *Psychol Inq*. 2012b; **23**(3): 267–270.  
[Publisher Full Text](#)
- John LK, Loewenstein G, Prelec D: **Measuring the prevalence of questionable research practices with incentives for truth telling.** *Psychol Sci*. 2012; **23**(5): 524–532.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Khabisa M, Giles CL: **The number of scholarly documents on the public web.** *PLoS One*. 2014; **9**(5): e93949.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kmietowicz Z: **Publishers withdraw 2500 journals from free access scheme in Bangladesh.** *BMJ*. 2011; **342**: ISSN 0959-8138.  
[Reference Source](#)
- Knobloch J, Albiez EJ, Schmitz H: **A serological survey on viral haemorrhagic fevers in Liberia.** In *Annales de l'Institut Pasteur/Virologie*. Elsevier. 1982; **133**: 125–128.  
[Publisher Full Text](#)
- Kousha K, Abdoli M: **The citation impact of open access agricultural research: A comparison between OA and non-OA publications.** *Online Information Review*. 2010; **34**(5): 772–785.  
[Publisher Full Text](#)
- Kriegeskorte N, Walther A, Deca D: **An emerging consensus for open evaluation: 18 visions for the future of scientific publishing.** *Front Comput Neurosci*. 2012; **6**: 94.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kurtz MJ, Henneken EA: **Open access does not increase citations for research**



articles from the astrophysical journal. *arXiv*. 2007.

#### Reference Source

Kurtz MJ, Eichhorn G, Accomazzi A, *et al.*: **The effect of use and access on citations.** *Inf Process Manag*. 2005; **41**(6): 1395–1402.

#### Publisher Full Text

Laakso M, Welling P, Bukvova H, *et al.*: **The development of open access journal publishing from 1993 to 2009.** *PLoS One*. 2011; **6**(6): e20961.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Lansingh VC, Carter MJ: **Does open access in ophthalmology affect how articles are subsequently cited in research?** *Ophthalmology*. 2009; **116**(8): 1425–1431.

#### PubMed Abstract | Publisher Full Text

Larivière V, Haustein S, Mongeon P: **The Oligopoly of Academic Publishers in the Digital Era.** *PLoS One*. 2015; **10**(6): e0127502.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Lawrence S: **Online or invisible?** *Nature*. 2001; **411**(6837): 521.

#### Reference Source

Lawson S, Meghreblian B: **Journal subscription expenditure of UK higher education institutions [version 3; referees: 4 approved].** *F1000Res*. 2015; **3**(274).

#### PubMed Abstract | Publisher Full Text

Leitner F, Valencia A: **A text-mining perspective on the requirements for electronically annotated abstracts.** *FEBS Lett*. 2008; **582**(8): 1178–1181.

#### PubMed Abstract | Publisher Full Text

Malhotra A, Younesi E, Gurulingappa H, *et al.*: **'HypothesisFinder': a strategy for the detection of speculative statements in scientific text.** *PLoS Comput Biol*. 2013; **9**(7): e1003117.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Matheka DM, Nderitu J, Mutonga D, *et al.*: **Open access: academic publishing and its implications for knowledge equity in Kenya.** *Global Health*. 2014; **10**(1): 26.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Mazzucato M: **The entrepreneurial state.** *Soundings*. 2011; **49**(49): 131–142.

#### Publisher Full Text

McCabe MJ, Snyder CM: **Identifying the effect of open access on citations using a panel of science journals.** *Econ Inq*. 2014; **52**(4): 1284–1300.

#### Publisher Full Text

McGlynn T: **The evolution of pseudojournals.** Small Pond Science. Dominguez Hills, CA, 2013.

#### Reference Source

McKiernan E, Bourne PE, Brown CT, *et al.*: **The open research value proposition: How sharing can help researchers succeed.** *Figshare*. 2016.

#### Publisher Full Text

McVeigh ME: **Open access journals in the ISI citation databases: analysis of impact factors and citation patterns: a citation study from Thomson Scientific.** Thomson Scientific. 2004.

#### Reference Source

Metcalfe TS: **The rise and citation impact of astroph in major journals.** *arXiv preprint astro-ph/0503519*, 2005.

#### Reference Source

Metcalfe TS: **The citation impact of digital preprint archives for solar physics papers.** *Solar Phys*. 2006; **239**(1–2): 549–553.

#### Reference Source

Moed H: **The effect of "open access" upon citation impact: an analysis of arxiv's condensed matter section.** 2006.

#### Reference Source

Mohammadi E, Thelwall M, Haustein S, *et al.*: **Who reads research articles? an altmetrics analysis of Mendeley user categories.** *J Assoc Inf Sci Technol*. 2015; **66**(9): 1832–1846.

#### Publisher Full Text

Müller-Langer F, Watt R: **Copyright and open access for academic works.** *Review of Economic Research on Copyright Issues*. 2010; **7**(1): 45–65.

#### Reference Source

Natarajan J, Berran D, Dubitzky W, *et al.*: **Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line.** *BMC Bioinformatics*. 2006; **7**(1): 373.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Norris M, Oppenheim C, Rowland F: **Open access citation rates and developing countries.** In *ELPUB*. 2008; 335–342.

#### Reference Source

Nosek BA, Bar-Anan Y: **Scientific utopia: I. opening scientific communication.** *Psychological Inquiry*. 2012; **23**(3): 217–243.

#### Publisher Full Text

Nuijten MB, Hartgerink CH, van Assen MA, *et al.*: **The prevalence of statistical reporting errors in psychology (1985-2013).** *Behav Res Methods*. 2015; 1–22. ISSN 1554-351X, 1554-3528.

#### PubMed Abstract | Publisher Full Text

Odlyzko A: **Economic costs of toll access.** *Open Access: Key Strategic, Technical and Economic Aspects*. 2006; **4**: 39–43.

#### Publisher Full Text

Odlyzko AM: **Open access, library and publisher competition, and the evolution of general commerce.** *CoRR*. 2013.

#### Reference Source

Okerson A: **With feathers: Effects of copyright and ownership on scholarly publishing.** *College and Research Libraries*. 1991; **52**(5): 425–38.

#### Publisher Full Text

Open Science Collaboration: **PSYCHOLOGY. Estimating the reproducibility of psychological science.** *Science*. 2015; **349**(6251): aac4716.

#### PubMed Abstract | Publisher Full Text

Packer AL: **The SciELO open access: a gold way from the south.** *The Canadian Journal of Higher Education*. 2009; **39**(3): 111–126.

#### Reference Source

Piwowar H: **Altmetrics: Value all research products.** *Nature*. 2013; **493**(7431): 159.

#### PubMed Abstract | Publisher Full Text

Piwowar HA, Vision TJ: **Data reuse and the open data citation advantage.** *PeerJ*. 2013; **1**: e175. ISSN 2167-8359.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Piwowar HA, Day RS, Fridsma DB: **Sharing detailed research data is associated with increased citation rate.** *PLoS One*. 2007; **2**(3): e308.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Porter AL, Kongthon A, Lu JC: **Research profiling: Improving the literature review.** *Scientometrics*. 2002; **53**(3): 351–370.

#### Publisher Full Text

Pöschl U: **Multi-stage open peer review: scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation.** *Front Comput Neurosci*. 2012; **6**: 33.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Research Information Network: **Activities, costs and funding flows in the scholarly communications system in the UK.** Technical report, 2008.

#### Reference Source

Russel RD: **The business of academic publishing: A strategic analysis of the academic journal publishing industry and its impact on the future of scholarly publishing.** *Electron Journal of Academic and Special Librarianship*. 2008; **9**(3).

Sahu DK, Gogtay NJ, Bavdekar SB: **Effect of open access on citation rates for a small biomedical journal.** 2005.

#### Reference Source

Satyanarayana K: **Journal publishing: the changing landscape.** *Indian J Med Res*. 2013; **138**(1): 4–7.

#### PubMed Abstract | Free Full Text

Scheufen M: **Copyright Versus Open Access: On the Organisation and International Political Economy of Access to Scientific Knowledge.** Springer, 2015.

#### Publisher Full Text

Schimmer R, Geschuhn KK, Vogler A: **Disrupting the subscription journals' business model for the necessary large-scale transformation to open access.** 2014.

#### Reference Source

Schimmer R, Geschuhn KK, Vogler A: **Disrupting the subscription journals' business model for the necessary large-scale transformation to open access.** 2015.

#### Publisher Full Text

Schwarz GJ, Kennicutt RC Jr: **Demographic and citation trends in astrophysical journal papers and preprints.** 2004; **36**(5). *arXiv preprint astro-ph/0411275*.

#### Reference Source

Shatkay H, Pan F, Rzhetsky A, *et al.*: **Multi-dimensional classification of biomedical text: toward automated, practical provision of high-utility text to diverse users.** *Bioinformatics*. 2008; **24**(18): 2086–2093.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Shavell S: **Should copyright of academic works be abolished?** *Journal of Legal Analysis*. 2010; **2**(1): 301–358.

#### Publisher Full Text

Shen C, Björk BC: **'predatory' open access: a longitudinal study of article volumes and market characteristics.** *BMC Med*. 2015; **13**(1): 230.

#### PubMed Abstract | Publisher Full Text | Free Full Text

Simmons JP, Nelson LD, Simonsohn U: **False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant.** *Psychol Sci*. 2011; **22**(11): 1359–66.

#### PubMed Abstract | Publisher Full Text

Solomon DJ, Björk BC: **A study of open access journals using article processing charges.** *J Am Soc Inf Sci Technol*. 2012; **63**(8): 1485–1495.

#### Publisher Full Text

Sutton C: **Is free inevitable in scholarly communication? the economics of open access.** *College & Research Libraries News*. 2011; **72**(11): 642–645.

#### Reference Source

Swan A: **The open access citation advantage: Studies and results to date.** 2010.

#### Reference Source

Swan A, Brown S: **Authors and open access publishing.** *Learn Publ*. 2004; **17**(3): 219–224.

#### Reference Source

Swan A, Brown S: **Open access self-archiving: An author study.** 2005.

#### Reference Source

Turner JA, Laird AR: **The cognitive paradigm ontology: design and application.** *Neuroinformatics*. 2012; **10**(1): 57–66.

#### PubMed Abstract | Publisher Full Text | Free Full Text

United Nations: **Universal declaration of human rights.** 1948.

#### Reference Source

van Assen MA, van Aert RC, Nuijten MB, *et al.*: **Why publishing everything is**



**more effective than selective publishing of statistically significant results.**

*PLoS One*. 2014; **9**(1): e84896. ISSN 1932-6203.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Van Noorden R: **Trouble at the text mine.** *Nature*. 2012; **483**(7388): 134–135.

[PubMed Abstract](#) | [Publisher Full Text](#)

Van Noorden R: **Open access: The true cost of science publishing.** *Nature*. 2013; **495**(7442): 426–429.

[PubMed Abstract](#) | [Publisher Full Text](#)

Van Noorden R: **Elsevier opens its papers to text-mining.** *Nature*. 2014; **506**(7486): 17.

[PubMed Abstract](#) | [Publisher Full Text](#)

Vanclay JK: **Factors affecting citation rates in environmental science.** *J Informetr*. 2013; **7**(2): 265–271.

[Publisher Full Text](#)

Wagenmakers EJ, Wetzels R, Borsboom D, *et al.*: **An agenda for purely confirmatory research.** *Perspect Psychol Sci*. 2012; **7**(6): 632–638.

[PubMed Abstract](#) | [Publisher Full Text](#)

Wang X, Liu C, Mao W, *et al.*: **The open access advantage considering citation, article usage and social media attention.** *Scientometrics*. 2015; **103**(2): 555–564.

[Publisher Full Text](#)

Watson M: **When will 'open science' become simply 'science'?** *Genome Biol*. 2015; **16**(1): 101. ISSN 1465-6906.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Wicherts JM, Bakker M, Molenaar D: **Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results.**

*PLoS One*. 2011; **6**(11): e26828.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Willinsky J: **Copyright contradictions in scholarly publishing.** First Monday. 2002; **7**(11).

[Reference Source](#)

Willinsky J: **The Access Principle: The Case for Open Access to Research and Scholarship.** MIT Press, Cambridge, Mass. 2006.

[Reference Source](#)

Wohlrabe K, Birkmeier D: **Do open access articles in economics have a citation advantage?** 2014.

[Reference Source](#)

Xia J: **Predatory journals and their article publishing charges.** *Learn Publ*. 2015; **28**(1): 69–74.

[Publisher Full Text](#)

Xia J, Harmon JL, Connolly KG, *et al.*: **Who publishes in “predatory” journals?** *J Assoc Inf Sci Technol*. 2015; **66**(7): 1406–1417.

[Publisher Full Text](#)

Xu L, Liu J, Fang Q: **Analysis on open access citation advantage: an empirical study based on oxford open journals.** In *Proceedings of the 2011 iConference*. ACM. 2011; 426–432.

[Publisher Full Text](#)

Zhang Y: **The effect of open access on citation impact: a comparison study based on web citation analysis.** *Libri*. 2006; **56**(3): 145–156.

[Publisher Full Text](#)



# Discuss this Article

## Version 1

Reader Comment 12 Apr 2016

**Ross Mounce**, University of Cambridge, UK

This is in no way a review of the entire manuscript, I don't have time for that at the moment.

This is merely an attempt at constructive criticism of a particular section, namely 'The case for publishers'.

Reading the manuscript as it currently is, one might come away from this section thinking that the average published research paper costs \$3500–\$4000 to publish. I do not believe this to be true. For instance, at the Journal of Machine Learning Research (JMLR) it costs between \$6.50 and \$10 per paper - it is a notably efficient journal with technically competent submitting authors that submit manuscripts in formats that are more immediately publishable (e.g. LaTeX), thus Shieber (2012) absolutely must be cited in any discussion of the cost of publishing. Likewise, another estimate is given by Ubiquity Press: 300 is all they need to sustainably publish manuscripts submitted in less convenient formats (such as MS Word documents), so please do cite: <http://www.ubiquitypress.com/site/publish/> too. Scale and 'selectivity' (the arbitrary rejection of significant percentage of papers that some journals like impose) are also very important factors in the cost of a particular journal's cost per paper. The high cost at Nature is largely down to the cost of rejecting all those papers, not the cost of the papers they actually end up accepting!

You may also want to mention the end-to-end XML based publishing system called ARPHA (Mietchen *et al*., 2015) as used by the open access publisher Pensoft and a similar(-ish?) system licensed by River Valley. Adoption of these more efficient, integrated publishing workflows will inevitably bring down the real costs of publishing - it's a madly inefficient workflow at most biology journals at the moment with completely different systems for authoring, peer-review and making the final version / typesetting / proofs.

Mike Taylor also has useful information on the real cost of typesetting here: <https://svpow.com/2015/06/11/how-much-does-typesetting-cost/>

I'm hugely surprised you don't seem to mention the effect of open access journals on decreasing publication delays. That's also a major value-add for most (all?) stakeholders. Daniel Himmelstein's (2015) excellent data & blog post showing that newer open access journals like eLife and PeerJ are decreasing the average time between submission and publication - faster publication. Furthermore, journals like F1000Research and Research Ideas and Outcomes (<http://riojournal.com/>) and preprint servers are eliminating the delay before submitted manuscripts are made public to zero days.

Finally, I'd advise extreme caution in adopting (willingly?) Beall's phrase "predatory publishers". You seem to indicate that if a journal charges a low APC it may be a predatory publisher? Would a solo author article at PeerJ then thus make PeerJ a "predatory publisher"?

If you still do insist on including a section on Beall's views (although why? whyyyyyyyy?), please do point out by citation/quotation that he is a controversial figure who thinks (his own words) that "The OA movement is an anti-corporatist movement that wants to deny the freedom of the press to companies it disagrees with".



Beall 2013. [The Open-Access Movement is Not Really about Open Access](#). Triple C

Shieber 2012. An efficient journal <https://blogs.harvard.edu/pamphlet/2012/03/06/an-efficient-journal/>

Mietchen D, Mounce R, Penev L (2015) [Publishing the research process](#). Research Ideas and Outcomes 1: e7547. doi: 10.3897/rio.1.e7547

**Competing Interests:** I am one of the founding editors of the open access journal Research Ideas and Outcomes (RIO). I am a Software Sustainability Fellow. I am a Panton Fellow for Open Data in Science I believe open access publishing should be the default I am one of nearly 16,000 academics who publicly boycott Elsevier (<http://thecostofknowledge.com/>)

---