Bo-Christer Björk, Wojtek Sylwestrzak, Jakub Szprot

Analysis of Economic Issues Related to Open Access to Scientific Publications

PLATFORMA OTWARTEJ NAUKI

UNIWERSYTET WARSZAWSKI
Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego
www.icm.edu.pl

Bo-Christer Björk, Wojtek Sylwestrzak, Jakub Szprot

Analysis of Economic Issues Related to Open Access to Scientific Publications

Wydawnictwa ICM
Warszawa 2014

Editing and proofreading
Maciej Chojnowski

Cover and title page design, graphics, typesetting and text makeup
Jakub Rakusa-Suszczewski

ISBN 978-83-63490-09-6

Published by
Wydawnictwa ICM
ul. Pawińskiego 5a
02-106 Warszawa

Table of Contents

List of Tables

List of Figures

History of scientific publishing from the economic point of view

Scientific publishing over the past centuries has undergone a number of important paradigm shifts, triggered by new technologies. The first was the invention of the printing press. Before that books were extremely expensive and rare commodities and the scholarly record was preserved in only a limited number of places housing hand-made manuscripts, such as the famous Library of Alexandria or medieval monasteries. After Gutenberg's invention of the printed press monographs became the primary vehicle for spreading scientific ideas, as exemplified by Newton's *Philosophiæ Naturalis Principia Mathematica* or Darwin's *On the Origin of Species.*

Printing also enabled another innovation, the periodical scientific journal, which allowed scholarly societies to spread ideas far beyond the privileged few who could attend meetings and listen to lectures in person. The first two journals of this kind were published 350 years ago – *Philosophical Transactions of the Royal Society* and the French *Journal des Sçavans.* Particularly during the 20th century the scholarly journal became the main channel through which scientists were communicating new research findings. At the same time the peer review system, which is an essential aspect of journals, has evolved and scientific publishing, once almost exclusively within the domain of scholarly societies, has also become a lucrative form of business.

Over the last couple of centuries the number of scholarly journal articles has steadily risen by 3% per year.[1] Since scientific societies were often unwilling to start new journals there has been a growing demand for new titles, particularly after the World War II, when the number of scientists rapidly rose. As a result a number of commercial publishers were able to capture a large share of the market, and these companies have further been consolidated via mergers and acquisitions. Although they price their subscriptions much higher than society publishers, they have for instance offered authors the advantage of not imposing the page charges, something society publishers quite commonly did as a secondary revenue stream a few decades ago.

The Internet has revolutionized scientific publishing, like so many other publishing industries. It essentially reduces the marginal cost of producing and distributing copies of articles to almost zero. It has also spawned two new types of sales vehicles, the "big deal" electronic license and the sale of access to individual articles. Of these two the electronic bundled licenses have been a great success, and today constitute the bulk of the revenue for leading publishers.

Universities have benefited since their scholars and students have gained access to as much as ten times more journals for a cost which is only slightly higher than what they used to pay for the journals subscribed in a paper form. They also save in handling and archiving costs. Publishers have benefited in particular via the high degree of lock-in they gain with their customers. It is almost impossible for a bigger university to discontinue subscribing to an e-license with one of the leading half-dozen publishers. The major reason for this is that discontinuing a license would lead to dramatically worse service to its staff and students and the ensuing strong pressure these would exert on the library. Each of the publishers has acquired a sort of monopoly position for their journal portfolio. It is symptomatic that the details of most of these multi-year contracts are kept secret via non-disclosure clauses. One of the reasons for this is that pricing of the contracts depends not only on the costs to the publisher or even usage levels at a particular university, but is primarily related to the ability of the customer to pay. For example the national e-licenses covering the major commercial and society publishers for countries like Finland (USD 26 million[2]) and Serbia (less than USD 2 million[3]) are on totally different levels, depending on the level of GDP per capita and in particular on the total amount of university budget funding in each country.

University librarians for the past three decades have been complaining about the subscription prices constantly rising faster than the general inflation, a phenomenon termed "the serials crises".[4] The price rises have continued also after the shift to predominantly electronic license access. The reason is simply the extremely strong oligopolistic bargaining power of the major commercial publishers in price-setting. The profit margins of leading commercial scientific journal publishers are reported to be in the range of 30–40%.[5] These are figures which are almost impossible to achieve in any competitive business.

Another downside to the dominance of the big deals is that universities end up spending most of the money allocated for subscriptions on a few licenses and that they have to cancel many individual subscriptions to less popular journals, particularly from small society publishers and other universities. Another factor is also the transaction costs of negotiating subscriptions, adding journals to the electronic intranets, which works in favour of the bigger bundles. Next factor influencing the spread of research results published in small society and university published journals is a fact that the practice of exchanging journal copies for free between universities is no longer of much relevance.

The other new type of sales channel, the sale of individual articles for a price of typically USD 30, has never become very popular, mainly because individual academic readers would have to cover the costs themselves or from their departmental budgets, in contrast to the subscribed content which is usually paid centrally by the university and thus experienced as a free good by researchers.

In the table below some key figures for the publishing of scientific peer reviewed articles are given.

Table 1. Some key facts about the global scientific publishing market*

Number of journals indexed in Web of Science	10,675
Number of scholarly peer-reviewed journals	28,100
Yearly total number of articles	1.8–1.9 million
Revenue of major English Language STM publishers	USD 9.4 billion
Commercial publisher's share of WoS articles	64%
Society publisher's share of WoS articles	30%

* Ware, Mark and Michael Mabe, *The STM Report: An Overview of Scientific and Scholarly Journal Publishing*, op. cit.

Overview of Open Access and its different forms

In addition to making possible the electronic dissemination via digital subscriptions and pay-per-view for individual articles, the Internet is also an enabler of more fundamental changes in the revenue and business models for scholarly publishing. Similar disruptive innovations have for instance occurred in telephony (Skype), the dictionary business (Wikipedia) and the production and sales of software (Linux, MySQL, Firefox).

The Open Access paradigm emerged in the early 1990s as individuals and groups of academics realized the potential of the Web. The first Open Access journals had small or simply no budgets and were run by volunteers on homemade software platforms mostly hosted on university servers of the editors. At a time when subscription journals still appeared only in print versions such journals became the focal points for OA enthusiasts. In parallel, scientists in some areas with long traditions of disseminating working papers and preprints of manuscripts undergoing peer review founded repositories where the authors could upload manuscript copies, which anybody with Web access could read. Such repositories were particularly popular in physics (arXiv, 1991) and economics (RePec, 1997), which had already developed a habit of disseminating preprints and working papers.

There is a number of Open Access (OA) definitions but the basic idea behind it is very simple: OA is about making scientific publications freely available on the Web for anybody with Internet access to view, download and print out. In relation to data and other scientific output, however, different aspects of OA are significant. These include the right to process and reuse the data. The recently developing trends to textmine large bodies of publications imply that these aspects of OA are increasingly important. Contrast that with the costly and time-consuming ways of getting access to peer-reviewed articles in the era before the big electronic bulk licenses, involving the staff of university libraries and individual researchers themselves in retrieving the issues stored in the university archives. Not to mention the time and effort of retrieving items not held by the local library, for instance via interlibrary loans. Academic readership was fairly concentrated to the leading journals in the scholars own field.

The idea of OA is somehow similar to the idea behind Wikipedia, with the main difference in the way material is produced and reviewed. It is this simplicity (like the simplicity of the Web itself), which seduced many of

the pioneers who launched OA journals in the 1990s. Imagine a future where all the scientific literature would be freely available on the Web and a reader could instantly follow up any reference in a given publication by just clicking on the hyperlink!

OA as a concept applies equally well to all sorts of scientific material: monographs, conference papers, theses, and not least importantly data. This text is however concerned primarily with the scientific peer reviewed journal literature and with the specific circumstances surrounding it.

In the debate about Open Access a distinction is often made between "gratis access" and "libre access" and some OA proponents would not even accept free access alone as a sufficient condition for Open Access. Gratis OA means that the texts are technically open on the Web for anyone to read, download and print. Libre OA means that the licensing statements attached to the text give readers specific rights to do other things with the texts. The most widely quoted definition of Open Access (The Budapest Open Access Initiative declaration) states for instance that:

By „open access" to this 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.[6]

While the majority of today's readers may still be content just to be able to read and cite articles, the additional rights given in libre OA are increasingly relevant to certain users (including automated tools). Therefore the current gratis OA materials may need to be relicensed to libre OA in the future in order to satisfy these requirements. It is also important to note that green manuscript copies usually provide only gratis access, not libre access, since the copyright mostly resides with the publisher of the official versions.

Hence in the following when we speak about Open Access we mean it in its general sense, including both gratis and libre access. There are also other aspects to open accessibility, in particular the time delay between the original publication and the OA availability. An important group of around 500 high quality subscription journals make the electronic version available within a delay of one year and many repository copies are made available with a delay either due to authors' decisions or embargo periods imposed by publishers. Thus while delayed OA in not fully as valuable as immediate OA, it is clearly much better than restricted access behind pay walls.

The main distinction in the literature is made between gold OA, that is articles published as OA by the publisher, and green OA, by which manuscript copies of published articles are meant, which have been self-archived by the authors or third parties.

Table 2. A classification of variants of different statuses of OA publications

Availability	Prior to publication	Immediately after publication	Up to 12 months after publication	In excess of 12 months after publication
Version	Original publication	Personal version	Preprint version	
Usage rights	Libre	Gratis		
Provider of OA copy	Publisher	Funder or employer	Author	Third-party
Setting	In fully free journal	Individually opened		
Location	Publisher's site	Subject-based repository	Institutional repository	Homepage or other
Permanency	Permanently OA	Temporarily OA		
Legal status	Legal copy	Illegal copy		

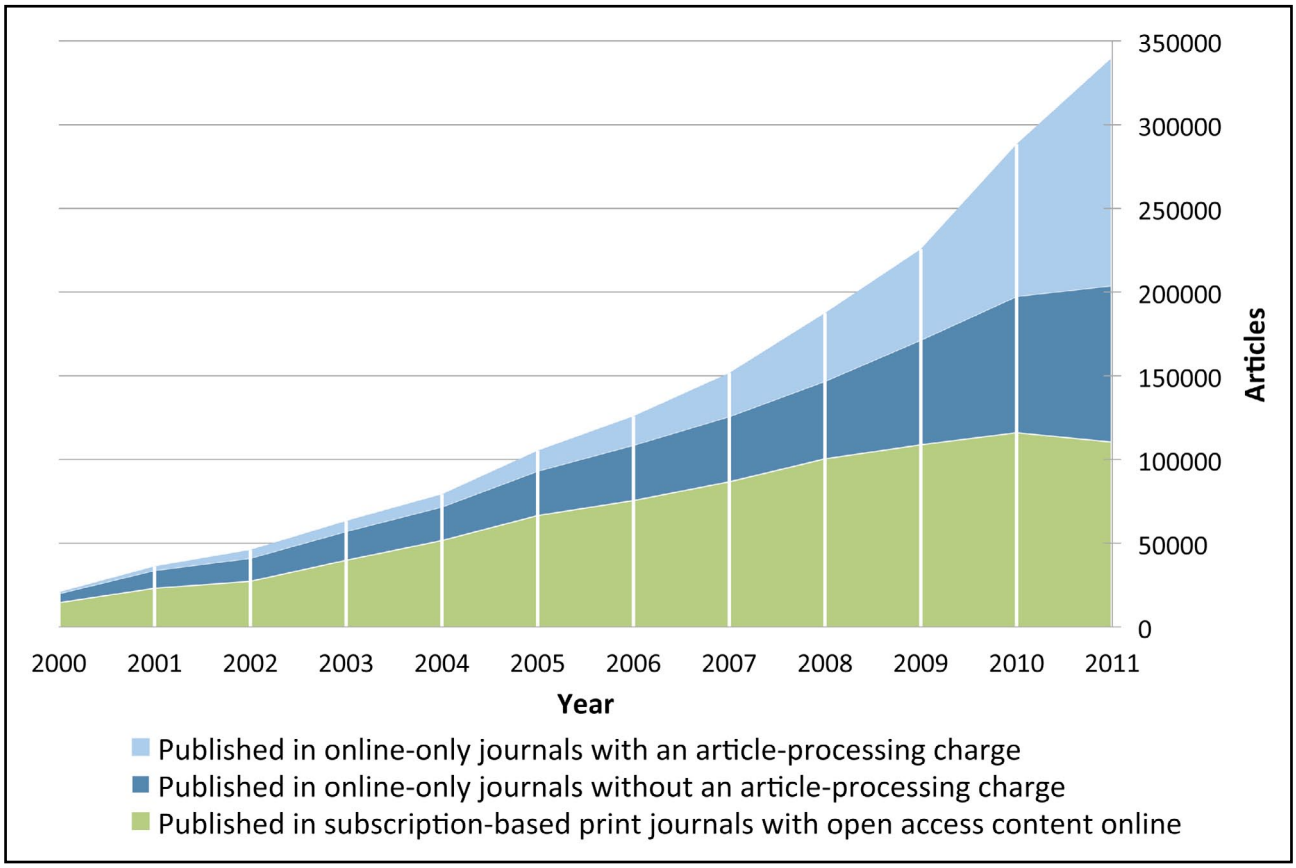
A special variation of gold OA is when subscription publishers make individual articles OA against the payment of fee (hybrid OA), usually around USD 3,000.[7]

Among full Open Access journals three distinct groups can be seen. Firstly there are journals which charge neither authors nor readers. These journals are electronic only and are often found in the social sciences and humanities. The majority of them have been founded by OA enthusiasts.

The second category predominantly consists of society or university published journals which often have a long history of paper publishing and which also, when started to publish a parallel electronic version, decided to make it OA. Such journals are very common outside the main Anglo-Saxon countries, often subsidized by research funders and supported by nationally funded portals like Scielo in Latin America. They are also very popular in Poland with several hundreds journals already following this model and making their content openly available through the Library of Science platform, funded by Polish Ministry for Science and Higher Education.

The third category consists of professionally OA journals published by commercial or in some cases noncom--mercial (i.e. PLoS) or society publishers which earn their revenues from so-called article processing charges (APCs). Figure 1 shows the evolution of the number of articles published in OA journals, broken down into these three categories.

Figure 1. The evolution of the number of OA articles, broken down into three main groups[8]



There have been several recent studies on the prevalence of OA. Gold OA is more straightforward to study, because it seems possible to count full OA journals indexed in DOAJ[9] whereas hybrid and green OA are more difficult to measure. It has to be noted, though, that while DOAJ is fairly complete with respect to English language journals, it lacks a significant fraction of titles in other languages. Only less than a third of the Polish research OA journals are listed there. A recent study by the Science-Metrix[10] commissioned by the EU found high prevalence rates approaching 50% for the OA availability of articles published in 2008–2011. While there is a number of methodological question marks concerning the global figures, the country-specific results indicate that 13% of articles by Polish authors could be found in full gold OA journals and 29% as green or hybrid copies, to be compared with EU averages of 8% and 37% respectively. It is also interesting to note that Croatia with its national OA journal portal Hrčak had the highest gold OA rate of 24% and the Netherlands with an extensive national infrastructure of institutional repositories the highest green rate of 52%.

A more conservative estimate of how common OA is would be that globally around 35% of recently published scholarly articles indexed in Scopus are available openly.[11] Of these close to 15% are in full OA journals, around 5% in delayed OA journals, 1% in hybrid journals and 15% as self-archived manuscript copies.

The Open Access citation advantage

Intuitively it would seem logical that articles which are available openly would be downloaded, read and hence cited more often than comparable articles which are not. If such an advantage could be empirically proved, it would provide a strong sales argument to promote OA to authors, in particular as an incentive for uploading article copies to repositories. Two recent bibliographies[12] provide good overviews of the studies and evidence up to date. An overwhelming majority of the studies have shown the existence of a considerable citation advantage. Swan only lists 4 out of 31 studies which do not show the advantage and Wagner only 3 out of 46. In addition to a statistically significant effect some studies have gone further and tried to examine other factors behind the effect. One possible factor is that manuscripts which are made available prior to peer review and formal journal publication have a longer time to accumulate citations (early advantage according to Swan). Another is that authors select their better articles for uploading to repositories (selection bias).

If all articles could be freely found, either because the journals had converted to OA, or because authors would consistently upload manuscripts to repositories, there would no longer be any citation advantage. What the studies indirectly show is, nevertheless, that OA increases readership and thus the potential impact on other researchers as well as on industrial development.

Analysis and comparison of the main forms of Open Access (green and gold) from the economic point of view

A number of studies have tried to numerically model the effects of the different forms of OA scenarios on the global overall costs of scientific publishing and getting access to the articles. By necessity such modelling requires making a lot of different assumptions, which makes the results unreliable and difficult to compare from one study to the other.

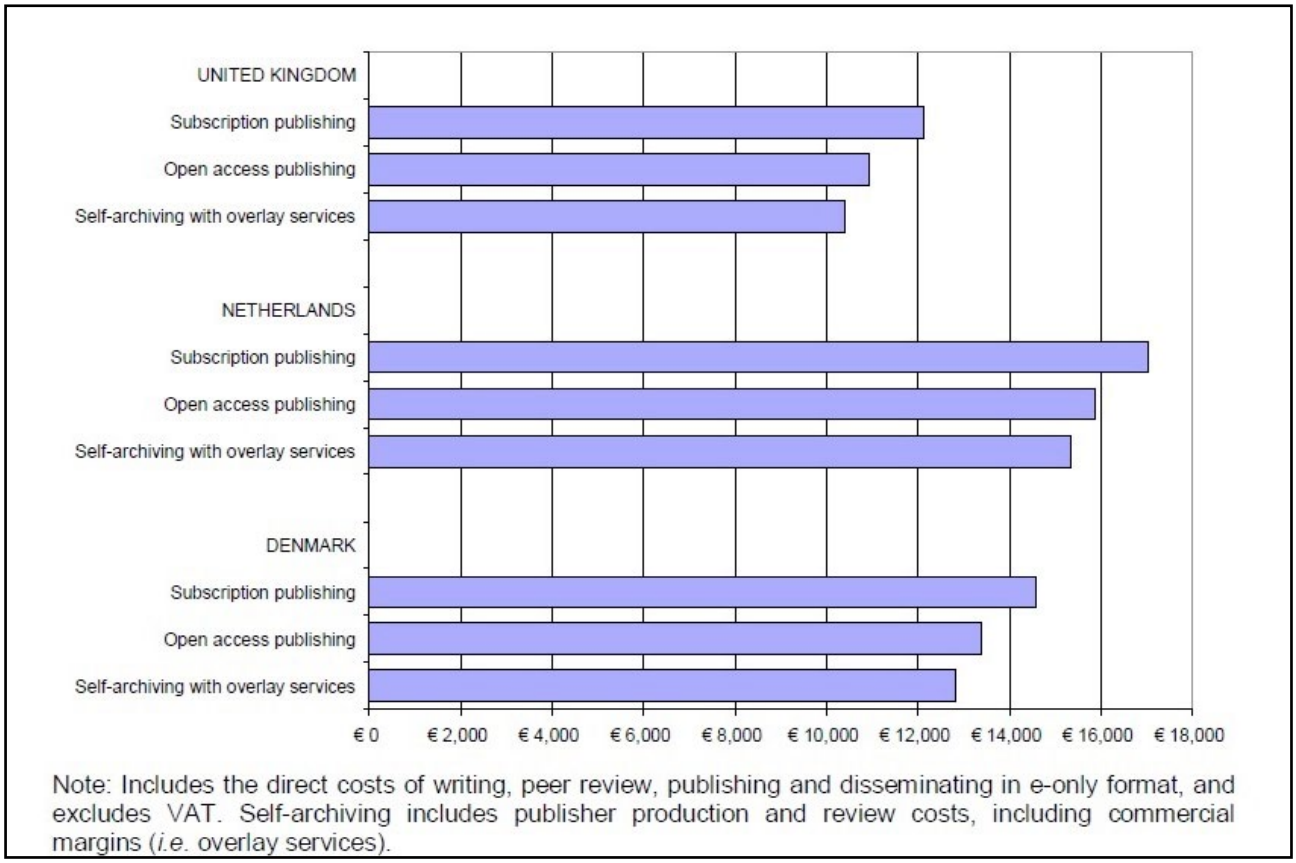
The effects of OA on direct publishing costs are the easiest to study. One of the most evident factors is that full OA publishing takes place only electronically and hence there are considerable savings in terms of printing and distribution costs. This advantage might, however, disappear if subscription based journals go over to electronic only distribution, which they so far usually have not. Sales and marketing costs may also decrease since marketing is no longer done to attract subscribers but only to attract authors.

A more comprehensive analysis also includes the costs to the readers and their organizations for retrieving the articles. The publishing costs are part of these costs via the subscription prices paid, but there are important additional costs in the form of the labour, IT-infrastructure and archive space costs (incurred for instance by university libraries) and the costs in the form of time spent by researchers in getting hold of the copies. The shift to electronic delivery has significantly reduced the manual handling costs of libraries as well as the retrieval costs of authors but has had little effect on the total sum of subscription prices paid by libraries.

In a study published by the UK Research Information Network[13] the current global costs of journal publishing were studied and the costs of OA scenarios were modelled. The researchers estimated that the cost effects of transition to electronic only publication would reduce the overall cost for publishing, dissemination and local library access provision by 13% and that a further transition to OA publishing financed with author-side payments would reduce it by additional 7%. In another study by Houghton et al.[14] the costs of a subscription-based but electronic only base scenario were compared with OA via journals (10% cost reduction) and OA via repositories (14% cost reduction). It is worth noting that a big share of the cost that is left unaffected are the reviewing costs, which in any case are not paid by the publishers but which are incurred by the universities of the editors and reviewers.

Of particular importance is subsequent Houghton study on the costs effects in three countries: the UK, the Netherlands and Denmark. The methodology was essentially the same as in the global study. Figure 2 shows the comparison of the overall costs per article for three scenarios (traditional subscription publishing, OA publishing of all articles financed by APCs and green parallel self-archiving).[15]

Figure 2. Scholarly communication costs in different scenarios for three countries



Source: John Houghton, *Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark*.

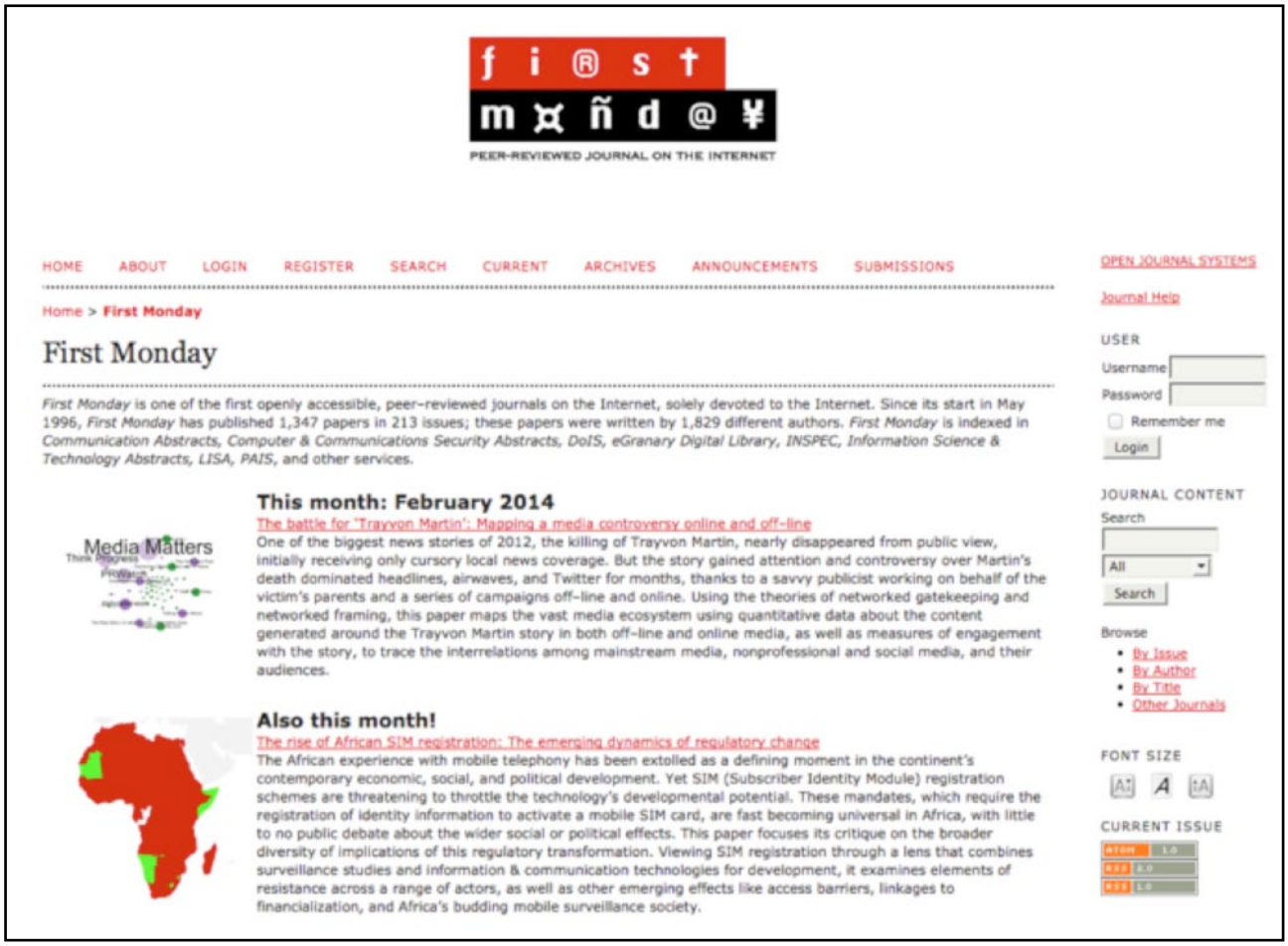
The overall conclusions of the study were compared to current costs. The subscription model gold OA would entail total cost savings of EUR 70 million per year in Denmark, EUR 133 million in the Netherlands and EUR 480 million in the UK whereas green self-archiving without subscription cancellations 30, 50 and EUR 125 million respectively. Obviously arriving at these results entails a lot of assumption and they should be treated with caution. One aspect is that massive self-archiving, if practiced in a majority of countries, would be likely to lead changes in subscription pricing, thus affecting the static assumptions of the model.

A country such as Poland faces quite different challenges compared to the three states mentioned above. In particular the costs of providing green OA to a large extent follow the country's general wage whereas the cost of OA in APC-funded full OA or hybrid journals tend to be the same globally (in contrast to subscription costs to bundle e-licenses, as discussed above). Many OA journals (but not hybrid) allow authors from developing countries to publish without paying the APC, but Poland would not be a country to which this was applicable.

Case studies of Open Access journals

There are several markedly different types of OA journals, the only common denominator being that the articles are freely available. The range is from very established subscription journals, which make the electronic version freely available since they get enough revenue from other sources, to small journals funded by academics and run using free software and voluntary labor. An example of the former, *British Medical Journal*, has been free since 1998. *First Monday* **(figure 3)** in turn is a journal specializing in Internet research and has been in existence since 1996.

Figure 3. Screenshot of *First Monday*, a free for the authors OA journal, which uses the open source Open Journals System software as a technical platform



In the following, a number of cases of OA journals are presented, highlighting the different options available.

The majority of journals that started publishing OA during the mid and late 1990s were new electronic-only journals (often with "electronic" or "on-line" as part of the name), founded by individual academics or groups of academics. The setting up of a new electronic only OA journal was simple and required little infrastructure or capital, in particular since there was no need for marketing to get subscribers. The central asset was the personal network of the editor, needed to recruit a credible editorial board and to solicit the first submissions. The journals were usually hosted on the website of the editor-in-chief's university with home crafted simple static Web pages. The managing and peer review process was usually done on a voluntary basis and the way journals were operated was close in spirit to the way many open source software development projects worked. The manuscript volumes handled were usually rather low.

Medical Education Online is a good example of a pioneer volunteer-based OA journal. From the start (1996) the journal was envisaged as a portal for experts interested in medical education, also containing material other than just peer-reviewed articles (i.e. short discussion items, book reviews and a resource section where academics could upload material), but over the years the journal material has been more and more concentrated on articles. Accepted articles are published once they are prepared, not in regular issues, which speeds up publication. The look and feel of the articles is nevertheless exactly the same as in traditional scholarly paper journals.

MEO was originally launched with a number of invited articles and for the first five years the number of submissions and published articles was low. But after having survived the first critical years the numbers have increased (currently around 20 published articles per year) and the journal has established itself within the research community. For the first decade *MEO* was published using a Web platform programmed by the editor-in-chief. Over the years the platform was improved to include for instance the possibility for electronic submission of manuscripts. In 2008, due to increase in the workload, the journal adopted article processing charges (APCs) in order to generate a modest revenue. From the start of 2010 the journal is published by a company specialized in OA publishing (Co-Action Publishing) and uses Open Journal Systems (OJS) software, a widely used open source solution for publishing scholarly journals and handling the review process. The level of the APCs has been gradually raised to the current EUR 850, in order to cover the costs of professional copyediting and publishing.

Elore is the oldest OA journal from Finland and a good example of how scientific publishing in languages other than English can benefit from OA. It is published by the Finnish Folklore Society and operates with a minimal budget using mainly volunteer labor. It publishes articles in both the national languages (Finnish and Swedish), but also in English, and includes other items than peer-reviewed articles, too.

In some countries national or regional portals for OA journals provide free (or very cheap) infrastructure for making the electronic versions of journals available. The Latin American portals Scielo and Redalyc, African Journals OnLine, Japanese J-Stage, and the Polish Library of Science[16] provide excellent examples. The Croatian portal Hrčak currently hosts 170 OA journals. In addition to offering very cost-effective IT platforms these

portals can also reach indexing agreement on a broad scale with organizations like the Web of Science (see Scielo-Thomson agreement as an example[17]). A large proportion of journals that use such portals are old established journals published by scientific societies, universities or university departments. A good example here is *Boletim do Museu Paraense Emílio Goeldi: Ciências Humanas*. This journal has its roots in one of the oldest scholarly journals in Brazil (established in 1894) and publishes articles in the social sciences and humanities with topics related to the Amazonas region. The language of the articles in the journal is Portuguese, but all articles have also abstracts in English. 15 years ago just finding out about articles in the journal would have been very difficult unless the reader belonged to a very select group of people who either had a personal subscription or whose university happened to subscribe to and archive the paper journal. Now it is only a Google search away.

The third wave of OA journals are those which have been launched by specialized professional OA publishers. Typical for these is that they fund their operations with so-called article processing charges (APCs). A very small minority also have submission charges, but the norm is a flat uniform standard fee. Many biomedical journals offer waivers from these fees to authors from developing countries. While almost all such publishers are commercial a notable exception is Public Library of Science, which is a non-profit operation initiated by a USD 7 million grant. BioMedCentral was started by venture capital and became successful enough to be purchased by Springer in 2008. Hindawi has been able to leverage the much lower staff costs in the home country of the company, Egypt, and has turned into a highly profitable company. From a Polish perspective an interesting publisher is DeGruyter, which in 2012 acquired the Polish publisher Versita.[18]

Since around 2010 the leading subscription publishers have realized the potential of OA publishing and increasingly launched OA journals of their own, which clearly benefit from the publisher's brand names in attracting submissions and probably get accepted for indexing in the Web of Science (and hence get impact factors) more rapidly than journals of less known OA publishers.

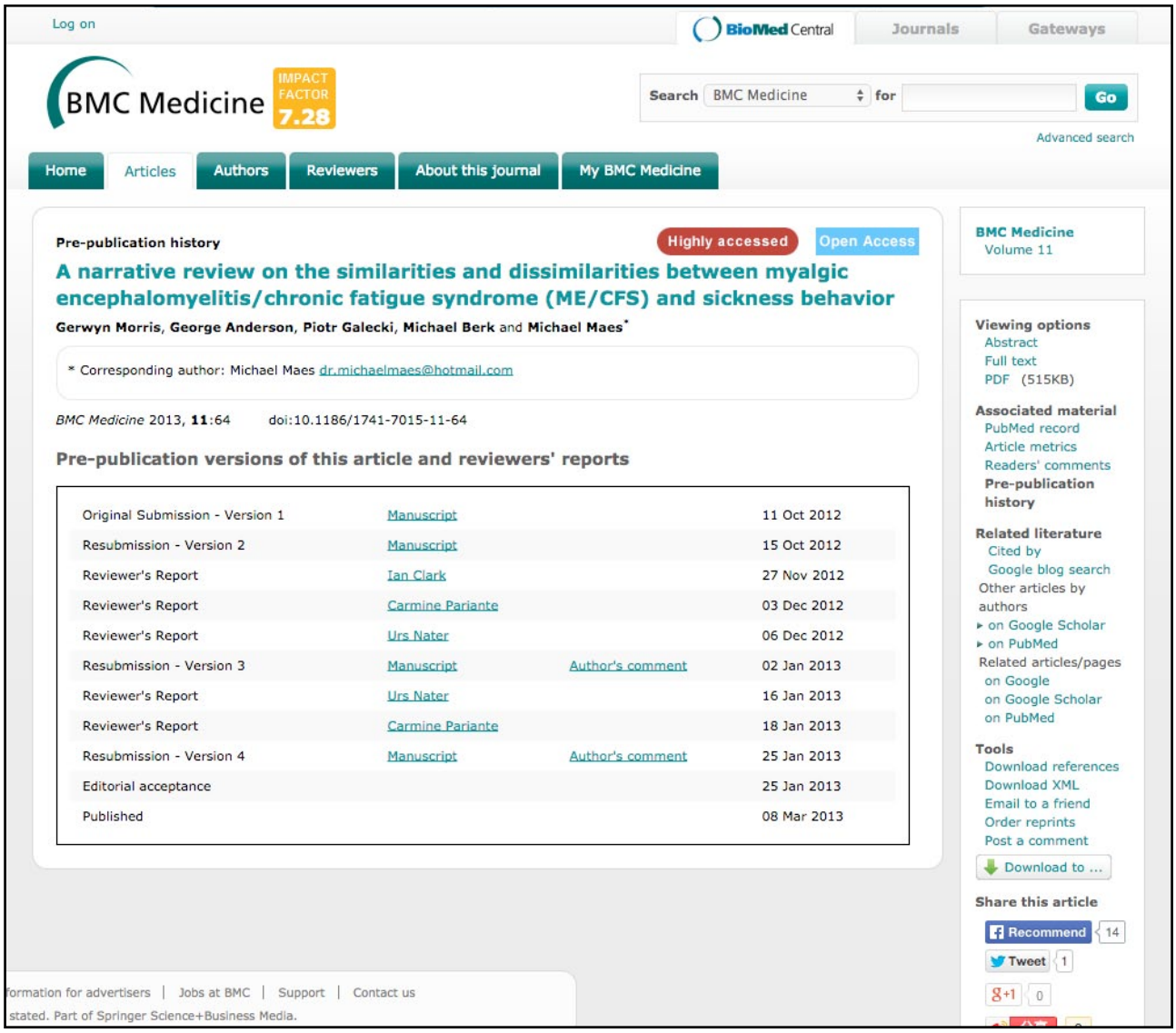
A study of the APCs in OA journals showed a global average charge level of USD 906 with higher levels in biomedicine than in other disciplines.[19] Recently when major subscription publishers have entered the OA market they have tended to set APCs higher, with an average close to USD 2,000.[20]

An unfortunate consequence of the low cost of establishing new OA journals is the emergence of hundreds of journals from publishers whose main interest is in collecting publishing fees. Such publishers are commonly nick-named "predatory publishers"[21] and they are highly visible to most academics due to the amount of spam email (asking to submit manuscripts, join editorial boards etc.) they generate. Many of these journals tend to accept articles without a proper peer review, thus giving OA a bad name, despite the relatively low number of articles they actually publish. Nevertheless their actual impact is low, since they usually publish low numbers of articles per journals and these get rather few citations.

At the other extreme there are very serious OA publishers such as PLoS and BioMedCentral who use the potential of the net for innovative features such as real-time download and citation statistics (including mentions in

social media like Facebook and Twitter) as well as publishing all the manuscript versions and review reports as appendices to the article.

Figure 4. Example of BMC Medicine alternative metrics and transparent peer review



Transitional forms of access to scientific publications (hybrid Open Access articles, delayed Open Access)

A couple of journals began to experiment with what is now called hybrid OA as early as 1999, but in earnest hybrid OA was started by Springer in 2004 with its *Open choice* programme, which encompassed the majority of its subscription portfolio of over 1,000 journals. Springer apparently saw hybrid OA as a possible transition route towards full OA funded by APC, and it set USD 3,000 charge – if all journals would convert to OA founded by APCs at such level, the revenue to the publisher would remain the same as before via subscriptions.

Subsequently other publishers also started experimenting with hybrid OA and in the last three years the number of journals offering this option has quadrupled, reaching over 8,000 today.[22] The pricing has usually been uniform for all journals of a particular publisher, around USD 3,000, although recently Elsevier has started differentiating its prices according to the subject area and scientific prestige of the journals. Uptake levels have remained very low (1–2%) for most journals, with a few high-prestige journals in biomedicine achieving higher levels.

As major society journals started to make available parallel electronic versions around the millennium shift some highly impactful journals decided to make the e-edition free after an embargo period of typically one year. This has long been an overlooked form of OA, partly because many OA proponents do not accept this as being OA at all (but on the other hand they accept self-archived green OA as a valid form of OA, which also often takes place after a delay of up to one year or more). Another reason for overlooking this phenomenon is that there is no central index for delayed OA journal, like there is for immediate OA journals (DOAJ).

A recent study has shown that there are around 500 journals with delayed OA within maximum one year, and that these publish around 5% of all peer-reviewed journal articles.[23] Furthermore many of these journals tend to be high-quality society journals, with an average citation impact three times that of full OA journals and two times that of subscription journals in general.

The green route to Open Access

Until now there have been three major locations for scholars self-archiving their manuscripts: author or departmental home pages, subject repositories or institutional repositories.

A meta-study of five studies on the location of green copies[24] showed that the manuscripts could be found predominantly on home pages (27–74%), then with slightly lower numbers in institutional repositories (19–44%) and finally in subject repositories (3–43%). The numbers are probably increasing for institutional repositories, which also offer more stable storage and longevity compared to home pages. Good subject repositories are only available in certain subject areas, but can be of regional importance. In addition to the leading physics and biomedical repositories there are around 50 other repositories of varying success.[25]

In history research there are for example interesting repositories dealing with European Integration and Latin American development. A very well functioning repository seems to be the Forced Migration Online Digital Library, which aims to collect a multitude of information resources related to refugees and forced migration.[26] The website is of high quality and the repository aims to reach out also to policy makers, the broader public and teachers. There are several highly specialized repositories in the arts and humanities, but none with broader scope. It is important to note that scholars in these fields tend to publish more in monographs or book chapters. Curiously there are two repositories dealing with different aspects of Basque culture, one broader (Hedatuz) and one concentrated on the language (ArtXiker). Both accept inputs in several languages (Basque, French, Spanish, English). The latter, like hprints.org (the free Nordic Arts and Humanities and Social Sciences e-print repository), uses the French national repository infrastructure HAL. Hprints.org was started with Nordic funding but never really took off.

In their report on repository types Armbruster and Romary[27] in fact distinguish national repositories as a specific category of its own. In addition to the French national portal HAL, which is being administered by large CNRS research institute, they also mention national portals in Japan and Spain. Clearly the founding of institutional repositories in any given country can be initiated top-down or bottom-up. In the more common bottom-up approaches most universities nevertheless opt for using one of a handful of available open source platforms (DSpace, EPrints). In some countries (for instance in the Netherlands) there have been attempts to integrate existing independent platforms. The EU is funding projects aiming at greater

standardization and interoperability (DRIVER[28] and later OpenAIRE).

Currently it seems that the heyday of subject repositories has passed, with the exception of the half-dozen repositories attracting large numbers of manuscripts. Institutional repositories on the other hand are becoming very common, and can fulfill many other functions in addition to offering a place for green copies of peer reviewed journal articles (PhD and Masters theses, teaching material). They are also increasingly tightly integrated with the research reporting (CRIS) systems of universities.

A further factor tipping the balance in favour of institutional repositories is the copyright policies of publishers concerning green copies. Leading commercial and society publishers started explicitly defining authors' rights in terms of where and when they could post manuscript versions on the Web around the millennium shift or slightly after and initially many commercial publishers, led by Elsevier, had quite liberal policies permitting authors uploading their accepted versions to repositories without a delay. Surprisingly society publishers were often more restrictive. But right from the start subject repositories such as the powerful PubMedCentral, which was backed by the NIH, were perceived as stronger threats and often not allowed. In the last few years several publishers have tightened their rules by taking into use embargo periods (Springer 12 months, Emerald 24 months). Seen in conjunction with the rapidly increased number of journals offering a paid hybrid OA option, this could be a strategy aimed at steering authors into paying the hybrid APC instead of going for the free self-archiving alternative.

Elsevier, which previously had allowed "accepted versions" in all institutional repositories, took a step back around 2008–2009 by requiring explicit agreements with the university in question if providing OA copy is mandatory for the university's personnel, a policy of disputable legal basis. This reversal of policy might have been triggered by the increasing numbers of such mandates and research such as the study of Gargouri et al.[29], which demonstrated that such mandates increase the self-archiving from a base level of around 15% to 60% or more.

All in all the publishers' policies concerning self-archiving are relatively permissive. A comprehensive study by Laakso[30] shows that immediate self-archiving is allowed for half of the subscription articles and that the share rises to around 80% after 12 months. The situation differs between sciences, with physical science journals mostly allowing immediate posting (80% allow it) compared to the longer embargoes common in biomedicine and social sciences.

While some early OA advocates claimed green OA to be almost free of cost, there has been a growing realization that setting up and operating repositories requires both human and financial resources, even if open source solutions (e.g. DSpace or EPrints) are mainly used as the IT-infrastructure for institutional repositories at most universities. In addition to the adaption, installation, and maintenance of the software and servers, library personnel might also be needed to (among other tasks) verify the copyrights of uploaded copies, correct references, and provide advice to researchers. Quite often institutional repositories are also integrated with the current research reporting systems which record the publications' meta-data for administrative and evaluation purposes.

Rough estimates of the costs of uploading and storing green OA copies of articles have been used in scenarios comparing different major strategies on how OA should be achieved. Houghton et al.[31] estimated the costs of uploading copies of all journal articles in the UK to repositories at around USD 33 per article, assuming that it takes 10 minutes of the author's time. One of the most systematic attempts to measure such costs was made by the EC funded Publishing and the Ecology of European Research project.[32] Average costs for setting up the IT architecture of a full repository were reported to be USD 60,000. The personnel cost per uploaded article were within the wide range of USD 2 to 53 depending on the repository. These results are very ambiguous; in particular since repositories can contain a wide variety of materials, and since the cost of setting up the repository and of managing staff has to be spread over all these document types.

Open Access to research data

Open data refers to datasets made available online on the Internet for anyone to use, distribute further and use as input in spin-off applications.[33] Such data can take many forms: astronomical and climatological measurements, cartographic and traffic data, economic or demographic statistics, patient data from medical trials, qualitative interviews conducted for an academic dissertation etc. Increasingly open data stemming from publicly funded research and different branches of government is recognized as a valuable source of innovation and growth, and policies are being formulated in many countries to increase the availability of such data. Data which is useful for research purposes by third parties can be stemming from research as well as from other data, e.g. from other branches of government and health care. It is important to realize that open data of both types can be used for research and that research data can also be used for non-research purposes.

With the Internet the obstacles to data availability to others are decreasingly technical ones but there still remain those caused by policy and practice. In some countries for instance it has been common that government agencies required part of the costs of data production and distribution to be covered through the sales revenue of such data, a perfectly rational approach prior to the emergence of the Web, but now a barrier leading to sub-optimal usage of the data in question.

The case for opening up data sets can be made from different perspectives. First, making the data behind research available is a prerequisite for the independent validation of findings.[34] Secondly, opening up datasets generated as a result of publicly funded activities is a matter of efficient use of public resources. OA to datasets is increasingly required by research funding agencies who want to guarantee the maximum utilization of data and avoid wasting money on redundant data collection.[35] Finally, open data is assumed to accelerate innovation in similar fashion to the way open source software has influenced the field of software development.

Economic effects of Open Access on publishing costs and beyond

Studies of the economic effects of OA have mostly been concentrating on the direct publication and subscription costs, which are the easiest to quantify. There are however costs and benefit effects further down in the chain of events during the life-cycle of a scholarly article, which can span decades. Important costs which are often ignored are the costs borne by scientific libraries of organizing the access to subscription material, which for instance include all personnel time used in negotiating deals with publishers or prioritizing journals to subscribe / not to subscribe to. Such costs seem to entail a sizeable overhead (in some cases over 100%) on top of the direct licence costs and these can be avoided almost entirely in the OA scenario.

Another important cost item is the time researchers spend searching for and retrieving articles. Clearly electronic access has sped up this process considerably and nowadays almost all the reading of scholarly articles is done on versions retrieved electronically. But it is also clear that this would be even more efficient if all material was only one hyperlink away.

All of the above are direct costs, which somehow are quantifiable. More difficult are the opportunity costs of researchers and other experts or lay persons not getting access to the publications they would want to. This is the welfare loss of being denied access (less scientific progress, wrong treatment of patients, fewer innovations in business).

There are four particularly interesting target groups who today often lack direct access by "a mouse-click": academics in other scientific fields, potential readers from less endowed universities (particularly in developing countries), potential readers from industry and the general public. The Houghton et al. report[36] makes an attempt to also quantify the impacts on returns to R&D of OA using a modified macroeconomic Solow-Swan model, and claim significant effects of the increased accessibility to research results. This is so far the only attempt at this type of analysis.

The role of public funding and mandates in the implementation of Open Access

Early OA journals as well as the first subject repositories were mainly founded by individual academics. After 2000 university librarians became a driving force in the OA movement and have heavily engaged in national OA work groups (FinnOA[37], OpenAccess.se[38]). One of the results of this increased awareness has been that the number of institutional repositories has rapidly grown. In the last few years major research funders, such as the NIH in the USA, Wellcome Trust and Research Councils UK in Great Britain, have become an important driving force for OA. Also the EU is getting involved – for instance in the regulations concerning the Horizon 2020 programme.[39]

Funders can promote OA in two ways, firstly by requiring from grantees that they either publish in OA journals or self-archive manuscripts in repositories, and secondly by providing the required funding for paying the APC charges in full OA or hybrid journals. There is currently a debate, particularly in the UK, about possible alternative scenarios for how funding can be implemented and what side-effects alternatives could have on the OA "market".[40]

A general requirement (which in the jargon of the OA debate is called a mandate) that publications stemming from dedicated project funding are made available via OA is increasingly common. The prime example of a mandate is the one issued by the National Institutes of Health in the United States, which allows a maximum delay of one year, in case the authors have chosen the self-archiving route. NIH has reported a compliance rate of 75% in the form of uploads by the authors or publishers to the PMC repository.[41] Gargouri et al. studied the uptake levels (in terms of all published journal articles) for three universities and one research institute with OA mandates and compared those with the uptake levels of a bigger selection of universities without mandates.[42] They found an average uptake of around 60% for institutions with mandates compared to 15–20% for other institutions with voluntary upload.

Even with mandates there are problems with getting full compliance and for this reason some funders have started providing "earmarked" funding to pay the APCs of full OA journals and in some cases also hybrid journals. The rationale behind this is that even if APCs are quite commonly defined as eligible costs in research contracts, this means that at the margin the researchers might face the alternative of using the last EUR 2000 of project funding for a conference trip or for paying an APC, and often the paid OA option is

then not chosen. By earmarking the funder provides this funding to an author (or his university) only for the purpose of paying APCs, which essentially means that the paid OA options is "free" for the authors, there is no opportunity cost involved. Wellcome Trust (one of the biggest private funders of medical research in the world) was the first major funder to start offering this type of funding.

Recently the UK Ministry of Science and Education has started providing substantial earmarked APC funding for the universities. For instance Research Councils UK provided block grants of GBP 17 million in 2013–2014.[43] There has been quite a lot of debate in the UK in the last couple of years as to whether government should favour gold or green as the preferred route to OA, and also conserving the conditions for providing funding for the hybrid APC. It is also important to note that scholarly publishing is a big export industry in the UK, and there is a strong publishers lobby trying to influence public policy.

In a wider perspective governments can also effect developments via the normal basic funding granted to universities, which for instance can be used to cover the cost of setting up and running institutional repositories. In some countries government money is also used to support national OA journal portals.

Table 3. National level OA journal portals

Portal	Country	Number of Journals
Scielo	Latin America	1,149
Redalyc	Mexico, Latin America	901
African Journals Online	Africa	130
J-Stage	Japan	1,674
Hrčak	Croatia	353
Biblioteka Nauki (Library of Science)	Poland	382

In many countries governments provide small subsidies to scholarly journals in general, especially in the social sciences and humanities. For instance in Finland the total support is annually around EUR 1 million, which is divided among some 70 journals, and there are very similar structures in Sweden and Norway. The problem is, however, that a monetary budget with income either through subscriptions or by having part of membership fee to the relevant association defray publication costs, is required. This creates a strong bias in favour of traditional subscription fees and steers the support to printing and distribution costs. One exception is Canada where the Social Sciences and Humanities Research Council (SSHRC) has recently changed its rules for supporting scholarly journals so that the subsidy is CAD 850 per published article, with a total maximum of CAD 30,000 per journal and a ceiling of CAD 5,000 for paper or e-distribution costs. These regulations focus the support on peer review and copy-editing costs of the journal production process, which means that OA journals get an equal treatment compared to paper journals.

Support can also be provided particularly for OA journals and activities related to them. Last year the Austrian Science Fund provided grants (after a competitive bidding process) for eight projects aiming

at establishing innovative OA journals (or converting subscription journals) in the humanities and social sciences.[44]

In Serbia in a national project local OA journals, which were struggling to survive economically, received training in setting up procedures for collecting APCs, and as a result of the project five journals have adopted this business model (while continuing to be OA).[45] In Lithuania support and training was provided for journal publishers to take into use the OJS systems.[46]

There are many other ways in which research funders or ministries can positively work towards wider adoption of OA, for example a requirement that only such OA articles (gold or green) can be taken into account in so-called research-assessment exercises where the research outputs of different universities are compared as the basis of their budget funding. For instance such policies are currently being considered in the UK (Higher Education Funding Council for England[47]).

Financing research publications in Poland

As of 2014, over 2,000 research journal titles are published in Poland, and comprise over 75,000 articles annually. Additionally a significant number of books, monographs and conference proceedings is published in Poland every year. Most of them are financed wholly or partly form public funds. In this report we are focusing on this aspect of scientific publishing in Poland, as it is crucial for the implementation of OA.

According to the Act of 30 April 2010 on the Principles of Financing Science, state budget funds allocated to science are managed by the minister responsible for science. As stated in the Article 5 of the Act, science funding shall be allocated, inter alia, to:

- strategic research and development work programmes and other tasks financed by the National Centre for Research and Development;
- basic research and other tasks financed by the National Science Centre;
- activities set forth in scientific unit Charters;
- activities of scientific units of higher education institutions, scientific units of the Polish Academy of Sciences and research institutes;
- large research infrastructure investments and construction projects that serve research or development work;
- scientific collaboration with other countries;
- science dissemination activities.

Research publications, including journals and scientific books, are funded as a part of science dissemination activities. There are two major grant schemes: Science Dissemination Activities (DUN) and Statutory Activities, addressed to different types of institutions.

According to the Article 25 of the Act, the financing of science dissemination activities shall include activities related to promotion, dissemination and popularisation of science. The funds shall be granted to the Polish Academy of Arts and Sciences and entities acting for the benefit of science as well as scientific libraries not included in scientific units. They cannot however be allocated to scientific units themselves (e.g. University units), which receive funding for statutory activities.

According to the Article 18 of the Act, the financing of statutory activities shall include, among other things, the maintenance of the research potential of scientific units as well as activities of scientific units, higher education institutions or other authorised entities related to the maintenance and extension of scientific databases, including the operation and maintenance of the Virtual Scientific Library – the system for making scientific databases and scientific publications available in electronic form. Among activities included in the maintenance of the research potential of scientific units are "science dissemination activities of the scientific unit concerning the tasks set forth in Article 25"; publishing research journals and scientific books is one of such activities.

Both kinds of grants are provided from the state budget and are managed by the Polish Ministry of Science and Higher Education.

The other funding schemes include programmes of the Ministry of Science and Higher Education, such as Index Plus, however their scope and size is smaller than the former two.

As a rule, Polish research funding agencies do not directly subsidize research publishing, however some publications costs (e.g. conference proceedings) can be covered by their specific grants.

According to the recently published results of the survey conducted by the Interdisciplinary Centre for Mathematical and Computational Modelling[48], the most common source of funds for Polish research journals is grants for statutory activities of scientific units, which was indicated by 41.7% of journals. Other sources of funds include subscriptions to the printed version (29%), individual sales of the printed version (26%), grants and donations other than grants for statutory activities and science dissemination activities (24.4%), grants for science dissemination activities (23.1%), author fees (15.3%) and advertising (9.3%).

The most important source of funds is also grants for statutory activities, indicated as one of the two most important sources of funds by 47.4% of journals. The second most important source of funds is individual sales of the printed version (37.3%). The third one is grants for science dissemination activities (30.2%). Less important are subscriptions of the printed version (26%), grants and donations other than grants for statutory activities and science dissemination activities (21.2%), author fees (19.2%) and advertising (7.2%).

As shown by these results, a significant part of the costs of research publishing is recovered through subscription fees or individual sales to research libraries, institutes, or individuals. However, since the majority of these fees is borne by public research institutions, the costs are only transferred to another subject but remain a burden to public funding, although in a non-explicit way. In most disciplines, sales of Polish journals outside Poland is marginal.

The total costs of financing research journals in 2013 DUN grant was PLN 7,892,565.00.

The total costs of statutory activities grants, including also other activities not related to research publishing, is much larger.

The level of statutory activities grants is automatically calculated based on the research status of the beneficiary institution, and while the costs of research publishing are eligible within these grants, the grant scheme does not define a separate part dedicated to these activities. Therefore, the relevant reported numbers are neither being collected nor analysed, and the specific level of funding research publishing through statutory activity grants remains unknown.

A significant cost factor of the non-OA model is the cost of getting access to publications, either through so--called country licenses or individual purchases. These costs are, however, difficult to estimate.

Open Access model in Polish research

As of today, Poland has not assumed a clear OA strategy for research publications and no policies are implemented. A number of institutions however undertakes different, not yet fully coordinated, efforts in this direction.

OA journals currently constitute the most widespread form of OA in Poland. According to the research conducted by the Interdisciplinary Centre for Mathematical and Computational Modelling[49], 947 Polish research journals, which is 49.2% of all the journals included on the list of journals rated by the Ministry of Science and Higher Education[50], provide their contents in an OA form. The prevailing type of OA as regards the vast majority of journals is gratis OA; only 1.35% of them provide libre OA (CC-BY and CC-BY--SA licences). In humanities 40.8% of the journals are OA; in social studies – 50.7%; in science – 63%; in biological sciences – 64%; in technological sciences – 59.4%; in agricultural, forestry and veterinary sciences – 68.6%; in medical, health and sport sciences – 45%; in the arts – 36.6%. The journals rated highest by the Ministry of Science and Higher Education (those with 20 or more rating points) include the highest percentage of OA journals – 64.7%. Different journals are available on their own websites, publishers' websites or platforms, in bibliographic databases or/and on journal platforms. Among the latter the biggest one is the Library of Science (Biblioteka Nauki) which provides OA to more than 400 journals.

Polish network of OA repositories remains, in comparison, rather underdeveloped. Currently there are 22 working repositories in Poland. 21 of them are institutional repositories, one is a disciplinary repository. Most are managed by public higher education institutions – universities (8) and universities of technology (5). One repository is managed by a department of the university. Private higher education institutions have two repositories. Repositories are also managed by (one in each case): an institute of the Polish Academy of Sciences, a consortium of the institutes of the Polish Academy of Sciences, a consortium of two universities, a project consortium, a scientific society and a scientific association. Most repository managers are librarians. Not all of the repositories are open in a strict sense; only in half of them (11) all content is OA. Many of them mix their traditional self-archiving functions with those of a journal platform. Most of the repositories are aggregated by the CeON Aggregator , six of them are harvested by the Open Access Infrastructure for Research in Europe (OpenAIRE). CeON Repository serves the researchers whose institutions do not have their own repositories. There are currently no data repositories in Poland.[51]

As far as publications of Polish authors in foreign journals are concerned, Springer *Open Choice* programme for Polish authors is worth mentioning. The programme allows each author affiliated with a Polish academic, educational or scientific research institution to freely opt out not to transfer the copyright of his OA publication in Springer hybrid journals to the publisher.

Current and possible ways of funding the Open Access model in Poland

A number of individual, yet uncoordinated activities related to OA in Poland is currently funded through different means. These include the funds described above, as well as own funds of institutions or European Funds such as the Programme Innovative Economy grants. In particular, the OA infrastructure costs are commonly funded by the individual institutions, European Projects or European Funds. Certain coordination actions as well as enabling and promotional activities are funded through the Ministry's of Science and Higher Education Virtual Library of Science grant scheme. Since 2010 the above mentioned Springer *Open Choice* programme has been funded from the same source. As for Polish scientific publishers, no specific incentives are currently available within the Science Dissemination Activities or Statutory Activities grant schemes neither for Open Access journals nor traditional journals for making their publications available in OA mode.

The generally available Polish sources for funding OA in research are mostly confined to one-time or short term grants. This setting, coupled with the lack of coherence of OA related funding leads to a situation where it is difficult for any stakeholder institution to assume a long term OA strategy.

A possible, more sustainable funding could be set up as a dedicated programme of the Ministry of Science and Higher Education. Such programme could cater for different types of OA-related actions, including e--infrastructure development and maintenance, support for migration to OA, support for OA journals, APCs, coordination or promotional activities.

It seems also possible to adjust the currently available funding schemes to provide better support for implementing OA, by specifically enabling and promoting the related kinds of expenses. For example, the current "automatic" way of assigning Statutory Activities grants means that any possible expenses related to OA for a given institution have to compete with its other more immediate costs within the same grant's budget, thus not providing any real incentive for promoting OA. Similarly, APCs could be considered as eligible costs in research grants but they should not compete with other costs in a given grant's budget.

Additionally, Polish research funding institutions could include provisions to promote OA publishing as well as OA to research data produced.

To summarise, besides the currently available fund sources, a dedicated programme focused on supporting OA in research could be created, while the other research funding schemes could be adapted to explicitly promote implementation of OA policies.

Glossary of Open Access terms**

Article processing charge (APC) – a fee often charged by full OA journals and always by hybrid journals. These charges are the major source of revenue for professional OA publishers.

Bundled or "Big deal" subscription contracts – electronic access to a broad package of journals from a single publisher or aggregator; it is negotiated with a university consortium or a single university, generally including a substantial discount compared with the list price of access to the journals in the package.

Creative Commons License – a standard set of licenses that are widely used to regulate electronic use of copyrighted material. These licenses are widely used for OA journals and provide options for specifying what types of uses are allowed. The CC-BY license, which only requires appropriate attribution, is widely used by APC-funded OA journals and is required by many funding agencies.

Delayed OA – subscription journals which make articles available at no charge after a delay of typically one year. Many of these journals are high impact, for instance *New England Journal of Medicine*.

DOAJ – Directory of Open Access Journals. An index of OA scholarly journals including various descriptive meta-data and currently containing almost 10,000 titles.

Gold OA – scholarly articles available openly immediately upon publication at the publisher's site. Full OA journal can either charge APCs to obtain revenue, they can be subscription journals making the electronic version only OA, or they can be electronic only journals which are free both for authors and readers and the resources for publication are provided by other means. Hybrid OA articles can also be seen as form of gold OA.

Green OA – green OA is a complement to gold OA. Some version of the manuscript, often not the final published version of article, is freely available. The term "green OA" is often used as a synonym for self-archived OA, though in some cases the publisher archives the green copy.

Hybrid OA – a form of OA where a publisher opens up and makes freely available an individual article in an otherwise subscription based journal, whose author, their institution or funder pays an article's processing fee.

** Adapted from Bo-Christer Björk, David Solomon, *Developing an effective market for open access article processing charges. Report*, Wellcome Trust, UK, March 2014, http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtp055910.pdf.

Journal Citation Reports – a report published yearly by Thomson-Reuters reporting the impact factors of journals indexed in Web of Science.

Open Access (OA) – OA is a practice of providing unrestricted access via Internet to scholarly research reports, most commonly scholarly articles but also to other types of written material and research data.

Open Journals System (OJS) – an open source platform for running scholarly journals.

OA mandate – regulation issued by a university or research funder, stipulating that articles must be available either as in full OA journals, as hybrid OA articles or as a manuscript self-archived within a maximum period in a repository.

Publisher embargo – a period after which an author, in the copyright agreement with the publisher, is allowed to self-archive an OA version of the article's manuscript in a repository. Often 12-month long.

Institutional repository – database for storing and dissemination of the publications, theses, teaching material, data sets etc. of a university or research institution.

Subject repository (also: disciplinary repository) – a data base of self-archived manuscripts of publications from a particular scientific field, for instance high energy physics, biomedicine or economics.

Self-archiving – author of an article makes a manuscript version (before or after the review process) available for free somewhere on the Web, typically in a subject or institutional repository.

Web of Science – index of journals and citation statistics owned by Thomson-Reuters. The inclusion criteria for new journals are significantly more restrictive than Scopus.

Scopus – index of journals and citations owned by Elsevier. Includes about 50% more journals than Web of Science.

References

1. Ware, Mark and Michael Mabe, *The STM Report: An Overview of Scientific and Scholarly Journal Publishing*, http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf.
2. FINELIB, *Katsaus FinELib-konsortion vuoden 2011 toimintaan* (in Finnish, yearly report 2011 of the FinELib--konsortium), http://www.kansalliskirjasto.fi/attachments/5l4xoyz0b/67VJr66W5/Files/CurrentFile/FinELib_raportti_2011_final.pdf.
3. Poynder, Richard, "Open Access in Serbia, Interview with Biljana Kosanović", „Open and Shut" blog, Newsletter, 13.9.2013, http://poynder.blogspot.fi/2013/09/open-access-in-serbia-interview-with.html.
4. Panitch, Judith and Sarah Michalak, *The Serials Crisis, A White Paper for the UNC-Chapel Hill Scholarly Communications Convocation*, http://www.unc.edu/scholcomdig/whitepapers/panitch-michalak.html.
5. Van Noorden, Richard, "Open access: the true cost of scientific publishing", *Nature* Vol. 495, Issue 7442 (2013), http://www.nature.com/news/open-access-the-true-cost-of-science-publishing-1.12676.
6. Budapest Open Access Initiative, 2002, http://www.budapestopenaccessinitiative.org/read.
7. Björk, Bo-Christer, "The hybrid model for open access publication of scholarly articles – a failed experiment?" *Journal of the American Society for Information Sciences and Technology* Vol. 63, Issue 8 (2012), pp. 1496–1504, http://onlinelibrary.wiley.com/doi/10.1002/asi.22709/pdf.
8. Laakso, Mikael and Bo-Christer Björk, "Delayed open access – an overlooked high-impact type of openly available scientific literature", *Journal of the American Society for Information Science and Technology* Vol. 64, Issue 7 (2013), pp. 1323–1329, http://www.openaccesspublishing.org/delayed/laakso_bj_rk_delay_preprint.pdf.

   Laakso, Mikael and Bo-Christer Björk, "Anatomy of open access publishing – a study of longitudinal development and internal structure", *BMC Medicine* 2012, http://www.biomedcentral.com/1741-7015/10/124.
9. DOAJ, Directory of Open Access Journals, http://www.doaj.org/.
10. Archambault, Eric, Didier Amyot, Philippe Deschamps, Aurore Nicol, Lise Rebout and Guillaume Roberge, *Proportion of Open Access Peer-Reviewed Papers at the European and World Levels – 2004–2011*, report, Science-Metrix, http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf.
11. Author's current estimate, based on triangulating several sources and extrapolating results to the current year.
12. Swan, Alma, *The Open Access citation advantage: Studies and results to date*, Technical Report, School of Electronics & Computer Science, University of Southampton, http://eprints.soton.ac.uk/268516/2/Citation_advantage_paper.pdf.

Wagner, A. Ben, "Open Access Citation Advantage: An Annotated Bibliography", *Issues in Science and Technology Librarianship* 2010 (60), http://www.istl.org/10-winter/article2.html.

13. RIN, *Heading for the open road: costs and benefits of transitions in scholarly communications*, Research Information Network, http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/heading-open-road-costs-and-benefits-transitions-s.
14. Houghton, John, Bruce Rasmussen, Peter Sheehan, Charles Oppenheim, Anne Morris, Claire Creaser, Helen Greenwood et al., *Economic implications of alternative scholarly publishing models: Exploring the costs and benefits. A report to the Joint Information Systems Committee*, JISC 2009, http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf.
15. Houghton, John, *Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark,* Knowledge Exchange, http://www.knowledge-exchange.info/Default.aspx?ID=316.
16. http://bibliotekanauki.ceon.pl/
17. Thomson Reuters, *Thomson Reuters Collaborates with SciELO to Showcase Emerging Research Centers within Web of Knowledge*, http://thomsonreuters.com/press-releases/102013/SciELO-Collaboration.
18. DeGruyter*, DeGruyter acquires Versita, increasing their open-access publishing business,* http://www.libraries.wright.edu/noshelfrequired/2012/01/09/degruyter-acquires-versita-increasing-their-open-access-publishing-business/.
19. Solomon, David, Bo-Christer Björk, "A study of open access journals using article processing charges", *Journal of the American Society for Information Science and Technology*, Vol. 63, Issue 8 (2012), pp. 1485–1495, http://onlinelibrary.wiley.com/doi/10.1002/asi.22673/pdf.
20. Björk, Bo-Christer, David Solomon, *Developing an effective market for open access article processing charges*. *Report*, Wellcome Trust, UK, March 2014, http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtp055910.pdf.
21. Beall, Jeffrey, „Beall's List of Predatory Publishers 2013", „Scholarly Open Acceess" blog, http://scholarlyoa.com/2012/12/06/bealls-list-of-predatory-publishers-2013/.
22. Björk, Bo-Christer, David Solomon, *Developing an effective market for open access article processing charges. Report*, op. cit.
23. Laakso, Mikael, Bo-Christer Björk, "Delayed open access – an overlooked high-impact type of openly available scientific literature", *Journal of the American Society for Information Science and Technology*, Volume 64, Issue 7 (2013), 1323–1329, http://www.openaccesspublishing.org/delayed/laakso_bj_rk_delay_preprint.pdf.
24. Björk, Bo-Christer, Mikael Laakso, Patrik Welling, Patrik Paetau, "Anatomy of Green Open Access"*, Journal of the American Society for Information Science and Technology*, Vol 65, No 2, (2014) pp. 237–250, http://www.openaccesspublishing.org/apc8/Personal%20VersionGreenOa.pdf.
25. Björk, Bo-Christer, "Open Access Subject Repositories – an Overview", *Journal of the American Society for Information Science and Technology*, e-published ahead of print, November 2013, http://www.openaccesspublishing.org/repositories/Subject_Repositories.pdf.
26. Cave, Mike, Sean Loughna and John Pilbeam, "Open Access Repository System for Forced Migration Online"*, Association of Librarians and Information Professionals Quarterly*, 3(4) (2008).
27. Armbruster, Chris, Laurent Romary, *Comparing Repository Types*, http://arxiv.org/ftp/arxiv/papers/1003/1003.4187.pdf.
28. Driver, project home pages, http://search.driver.research-infrastructures.eu/.
29. Hagerlid, Jan, „Elsevier tries to block institutional OA mandates", „Open Access i Sverige" blog, Swedish

Royal Library, 31.1.2012, http://openaccess.blogg.kb.se/?s=Elsevier+tries+to+block+institutional+OA+mandates.

30. Laakso, Mikael, "Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed", *Scientometrics*, http://link.springer.com/article/10.1007%2Fs11192-013-1205-3, self-archived version at www.openaccesspublishing.org/.
31. Houghton, John, Bruce Rasmussen, Peter Sheehan, Charles Oppenheim, Anne Morris, Claire Creaser, Helen Greenwood et al., op. cit.
32. *PEER Economics Report*, Centro ASK, Università Bocconi, http://www.peerproject.eu/fileadmin/media/reports/PEER_Economics_Report.pdf.
33. Murray-Rust, Peter, "Open Data in Science", *Serials Review*, Vol. 34, Issue 1 (2008), pp. 52–64, http://confoa08.sdum.uminho.pt/apresentacoes/peter_murray_rust_npre20081526-1.pdf.
34. Boulton, Geoffrey, Michael Rawlins, Patrick Vallance, Mark Walport, "Science as a public enterprise: the case for open data", The Lancet Vol. 377, Issue 9778 (2011) pp. 1633–1635, http://www.thelancet.com/pdfs/journals/lancet/PIIS0140673611606478.pdf.
35. *OECD Principles and Guidelines for Access to Research Data from Public Funding*, OECD 2007, p. 49, http://www.oecd.org/sti/sci-tech/38500813.pdf.
36. Houghton, John, Bruce Rasmussen, Peter Sheehan, Charles Oppenheim, Anne Morris, Claire Creaser, Helen Greenwood et al., op. cit.
37. http://www.finnoa.fi/
38. OpenAccess.se – Programme for Open Access to Scholarly Publishing, http://www.kb.se/openaccess_english/.
39. EU, *Open Access in Horizon 2020*, OpenAIRE homepages, https://www.openaire.eu/oa-in-h2020/oa-in-h2020/h2020/.
40. Björk, Bo-Christer, David Solomon, *Developing an effective market for open access article processing charges*, op. cit.
41. Poynder, Richard, *Open Access Mandates: Ensuring Compliance*, "Open and Shut" blog, 18.5.2012, http://poynder.blogspot.fi/2012/05/open-access-mandates-ensuring.html.
42. Gargouri, Yassine, Chawki Hajjem, Vincent Larivière, Yves Gingras, Les Carr, Tim Brody and Stevan Harnad, "Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research", PLoS ONE, 5(10) (2010), http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0013636.
43. *Accessibility, Sustainability, Excellence: How to Expand Access to Research Publications. A Review of Progress in Implementing the Recommendations of the Finch Report*, October 2013, http://www.researchinfonet.org/wp-content/uploads/2013/02/Final-version.pdf.
44. *Austrian Science fund funding the metamorphosis of HSS open access journals*, Europe PubMEd Central blog, 13.1.2014, http://blog.europepmc.org/2014/01/austrian-science-fund-funding.html.
45. Sipka, Pero, *Supporting Serbian Journal Publishers to Switch to Article Processing OA Model – A Case Study from Serbia*, Eifl report, February 2012, http://ceon.rs/pdf/oa_case_study_serbia.pdf.
46. Dagiene, Eleonora, *Promoting Open Access Through Implementation of Open Journals Systems in Lithuanian Research Journals Project*, Eifl report, February 2012, http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCYQFjAA&url=http%3A%2F%2Fwww.eifl.net%2Fsystem%2Ffiles%2F201202%2Foa-case-study-lithuania-oa-ojspublishing-final.pdf&ei=MNgrVLXhNYOkygPvloHgCg&usg=AFQjCNFaOnyULP0R3wJXuBSlw6MiKNV1xQ&sig2=nkyJi6YcVBx_-0-7hslRBQ&cad=rja.

47. HEFCE, *Open Access and research assessment*, https://www.hefce.ac.uk/whatwedo/rsrch/rinfrastruct/oa/.
48. *Otwarta nauka w Polsce 2014. Diagnoza*, ed. J. Szprot, Wydawnictwa ICM, Warszawa 2014, pp. 65–66, http://pon.edu.pl/index.php/nasze-publikacje?pubid=13.
49. Ibid., pp. 54–80.
50. The research covered journals from the list which had been published on the 20th of December 2012 and had been in force until the 17th of December 2013, when the new list was published.
51. *Otwarta nauka w Polsce 2014. Diagnoza*, op. cit., pp. 30–42.

Biographical notes

Bo-Christer Björk is Professor at the Hanken School of Economics in Helsinki; he has degrees in Engineering and Economics from three different universities. Previously he was Professor at the Royal Institute of Technology in Stockholm for seven years. He is one of the leading researchers studying the effects of the Internet on publication and information sharing in science as well as an active participant in the societal debate on the subject. He was the first chairman of the FinnOA (Finnish Open Access) committee from 2003 to 2008. He published over 50 articles in peer-reviewed journals.

Wojtek Sylwestrzak works at the Interdisciplinary Centre for Mathematical and Computational Modelling at the University of Warsaw, where he established and heads the Centre for Open Science. His past experience involves deployment of a number of milestone Polish Internet services since 1993, including data repositories, large scale search engines and distributed systems. His current interests include Big Data analytics and text mining, particularly of scholarly publications.

Jakub Szprot works at the Interdisciplinary Centre for Mathematical and Computational Modelling at the University of Warsaw, where he leads the Open Science Platform. Since 2006 he has been involved in various Polish and European projects concerning scientific infrastructure, digital scientific resources and open models of scholarly communication.