

APE 2011: Academic Publishing in Europe

A short report from the International APE Conference: “Smarter publishing in the new decade” *

11–12 January 2011, Berlin-Brandenburg Academy of Sciences (preceded by the Education and Training Course: “Understanding and being successful in publishing” on 10 January 2011)

Chris Armbruster^a and Svenja Hagenhoff^b

^a *Max Planck Digital Library, Berlin, Germany*

^b *University of Applied Sciences, St. Pölten, Austria*

1. Welcome and opening addresses

In his **opening remarks**, Professor Michael Mabe (CEO, International Association of STM Publishers, The Hague and Oxford) welcomed all participants. He reminded the audience that twenty years of the Internet have changed much in publishing, but that many new formats have historic precursors, e.g., XML the scroll; and PDF the codex. He identified three preeminent themes for 2011 and the years to come: (a) policy development with regard to research data and publications, particularly as new evidence becomes available from a variety of international projects; (b) the (possible) tipping point for eBooks and therefore for all of scholarly publishing to occur online first and primarily; and (c) how to deal with intellectual property rights when, in principle, the first digital copy serves all.

In his **opening keynote**, Professor John Wood (Secretary General of the Association of Commonwealth Universities, London and Chair of the High Level Expert Group on Scientific Data) spoke about *Riding the wave: How Europe can gain from the rising tide of scientific data*. The rising tide drives the globalization and virtualization of research, enabling the world to tackle major challenges like sustainable development in a more democratic way through shared and equal access to data and its extensive re-use. At the level of infrastructure, further large-scale investment in the grid and super computers is paramount, while at the policy level international summits should be hosted to develop common standards and a shared understanding of the technical, linguistic and legal challenges involved. Professor Wood stressed that European projects to build research infrastructures demonstrate conclusively that

* Address for correspondence: E-mail: info@ape2011.eu.

eScience and data-driven research is the future, not only in STM but, equally, in the social sciences and humanities. In all these fields, the integration of publication and data will open up whole new frontiers of research.

In the **second keynote**, Professor Dr. J.J. Engelen (President of the Governing Board of the Netherlands Organisation for Scientific Research – NWO, The Hague; and Vice-President of EUROHORCs, European Heads of Research Councils) highlighted *What academic publishing can do for science*. EUROHORCs and NWO favor a transition to open access publishing – professional and peer reviewed. NWO has established a fund to support open access publishing, not just for journals but also for books. More generally, the main challenge for those responsible for allocating these funds is to find ways to scale up quality open access publishing more rapidly. One route is to flip established journals. A second route is to establish new journals. The latter, however, takes time and is fraught with uncertainty. So fund holders are looking for new ways to promote and secure the launch of quality journals. Publishers are invited to approach fund holders about the launch or conversion of journals.

The first session, **Peer Review**, was chaired by Mayur Amin (Senior VP, Research and Academic Relations, Elsevier, Oxford). He highlighted how peer review was indispensable but also subject to controversial discussion. The focus of the session was to be on improvements to peer review.

Dr. Adrian Mulligan (speaking also for **Dr. Andrew Plume**, both Research and Academic Relations, Elsevier, Oxford) started his talk on *Peer review: The evidence*, by stating that over 1.4 million articles are published in peer-reviewed journals annually. Elsevier conducted a peer review survey between July 28 and August 11, 2009. From a randomly selected sample of 40,000 researchers, about 4000 completed the survey. Overall, researchers are satisfied with the mechanisms of peer review. For the most part, researchers do expect peer review to improve quality, but they are not very confident that the system can verify the originality of a manuscript, select the best or detect fraud. The system is considered sustainable in terms of available peer reviewers. Innovations in peer review were also examined. With regard to open peer review, the British Medical Journal Group conducted a trial and found that it did not improve the quality of the reviews. *Atmospheric Chemistry and Physics* is a journal with open peer review, but three out of four referee comments are posted anonymously, indicating that reviewers prefer this. With regard to cascading peer review, the Neuroscience Peer Review Consortium (NPRC), an alliance of 37 journals, agreed to process manuscript reviews ‘downstream’, but authors agreed to have the reviews forwarded to a second consortium journal for less than five percent of manuscripts submitted. With regard to the idea of publishing rejected manuscripts, *Rejecta Mathematica* accepts manuscripts if the authors are willing to make a case for them, but since the summer of 2009, only six papers have been published.

Dr. Bernd Pulverer (Head of Scientific Publications at EMBO and Chief Editor of the EMBO Journal, Heidelberg) spoke on *Transparent Peer Review*. He re-affirmed peer review as crucial to quality assurance, but believes that the process must be overhauled to make it faster and more transparent. In 2009, EMBO established a new review process based on Web 2.0 technologies. The review platform enables the reading of the anonymous reviews, and accelerates communication with authors. In effect, the manuscripts and the reviews are published together. Initially, some five percent of authors backed out, but this number has dropped. In autumn 2010, the new review process was rolled out for all EMBO publications. No differences have been observed in the quality of reviews and the ratio of submissions to acceptance, but the reviews are available for public inspection.

Prof. **Maurizio Marchese** (Department of Information Engineering and Computer Science, University of Trento, Italy) gave a talk about *Liquid publications: Scientific communication meets the social web*. He started with the observation that no systematic correlation could be found for articles between

peer review judgment prior to publication and the number of citations received after publication. Moreover, peer reviewers mainly demonstrate strong agreement when clearly rejecting or fully accepting manuscripts, but not for the majority of manuscripts that fall in between. LiquidPub is about linking scientific networks and social media to develop a scholarly communication system for the Internet era – with usage and re-use as important components. LiquidPub gives authors the freedom to write until they are finished (no standard length of text), embed data, slides and other material, update and improve the text and kick-start viral dissemination through sharing. The relevance of a contribution is determined based on usage analysis, and publications can be ranked, establishing a recommendations system.

The second session, **The data deluge, to drown or swim?** was chaired by Robert M. Campbell (Senior Publisher, Wiley Blackwell, Oxford). He focused on the fact that the possible integration of data with publishing is still an open issue with an unknown outcome.

Dr. Herbert Grüttemeier (Immediate Past President, ICSTI and Head of International Relations, INIST-CNRS, Vandoeuvre-lès-Nancy) reported on *Helping to ride – a look at data sharing and access policies*. He noted a wide variety of quite similar policies on data sharing and preservation across Europe and North America. This is a major concern and efforts are underway to establish integrated data platforms. It is anticipated that this will promote openness, encourage re-use and lead to new and unexpected research questions. However, faculty is often reluctant to pro-actively share and/or hand over data. Researchers have invested heavily in their data, and, since they may derive a competitive advantage from what they have produced, will want to exploit them in publication, and may not trust others to preserve them adequately. This means that data sharing policies are currently beset by the problem of enforceability.

Mrs Eefke Smit (Director, Standards and Technology, International Association of STM Publishers) spoke on *Taking the current when it serves – research data from the publishers' perspective*. She introduced her talk by stating that research is becoming increasingly data-intensive. The researchers surveyed indicated that most of them store data on their computer, mobile storage carriers, and/or servers of the institute. Asked about the storage of future data, most of them mentioned the digital archives of their organization as well as external web services like Google. She also pointed out that publishers have no coherent approach, let alone a standardized one, for including data with publications. Some allow data supplements, some do not; some ask researchers to include temporary or permanent links to the data, some do not. She recommended that publishers should face up to the challenge of data storage. She stressed that requiring the availability of supporting research material should be an editorial policy, and that publishers should ensure that data is stored, curated and preserved by secure and reliable means. Bi-directional links and persistent identifiers are required, and a best practice for peer reviewing data must be developed.

Dr. Jan Brase (Chairman, International DOI Foundation and Managing Director, DataCite, German National Library of Science and Technology, Hannover) spoke about *DataCite – improving access to research data from the perspective of libraries*. DataCite is a consortium of national libraries from Germany, Denmark, France, the Netherlands, Canada, Great Britain and the United States (15 members, shared infrastructure) to facilitate access to data and establish them as legitimate and citable contributions to the scientific record. The consortium is also aiming for data archiving that verifies results and helps to re-purpose data for further study. DataCite is a DOI registration agency and cooperates with CrossRef. DOIs provide the permanent link, and the metadata kernels are collected in a central database.

Dr. Malte Dreyer (Head, Research and Development, Max Planck Digital Library, Munich) spoke on *Managing publication and research data: The eSciDoc eResearch infrastructure*. eSciDoc is an open

source e-research environment which allows scholarly communities to collaborate globally in an interdisciplinary manner. It provides a set of services, tools and content models as well as a growing set of applications, such as publication management, digitized resources and image data. It is an infrastructure that ensures enduring access to research results and the data resources of scientific institutions and research organizations.

The panel discussion which followed centered on the sustainability of data archiving and long-term access. Members of the audience were particularly concerned about costs. It was suggested that fund holders as well as libraries would have to be engaged more strongly. The issue of whether scholarly publishers or research institutions would be better suited to manage data archiving was also discussed. The crucial questions which emerged, however, were the need to capture metadata early and decisions about what will be archived and for how long.

The **APE Lecture** was delivered by Professor Werner Kuhn (Director, Institute for Geoinformatics, University of Münster). In *The APE lecture: Information in space and time*, Prof. Kuhn stressed that academic publishing should be open access by default and should deliver more than just papers. Discussions about cost and ownership detract from the real issue, which is to create value. From his point of view there are significant untapped opportunities to create value in academic publishing. For example; there is an opportunity to create and curate vocabularies for linked data, and to exploit location and time for access, filtering and reasoning. We currently have silos with a lot of data about different topics such as climate change, health, migration etc. The real challenge is to exploit this data across silos by linking data and models. The LODUM project was established for this purpose. LODUM stands for Linked Open Data University of Muenster. The University of Muenster is the first university which has decided to go open with its research data to support spatio-temporal and semantic reasoning. There are still some challenges surrounding the ambitious aims of the LODUM project, like creating incentives to open up research data, accessing rights management, the lack of shared vocabularies as well as trust and reputation modeling.

The second day of the conference began with two parallel sessions:

(A) **The Future Group** (Workshop moderated by Ehrhardt F. Heinold, Heinold, Spiller and Partner, Hamburg) explored the possible futures for scientific publishing with regard to potential changes of the product and, consequently, the role of publishing houses. Will the rise of data-driven research, with continuous data streams, lead to the emergence of complementary, continuous forms of publishing with shorter, fluid and/or updated versions, or will the traditional discrete forms of publishing by article and book prevail? Will Web 2.0 publishing (social media, wiki publishing) support more continuous publishing with fluid versions for different audiences? There was an animated discussion of these controversial questions and different answers were offered, but no consensus prevailed. Equally controversial was the question of whether publishing houses should primarily be serving authors, communities or readers, and what the preferred mechanism of payment should be. To sum up, there was no agreement on the future role of publishers, suggesting that it is either too early to make predictions, or alternatively, that divergent models of scientific publishing might be emerging.

(B) The session: **Going Out! International Chinese Publishing** was chaired by Dr. Matthias Wahls (m.w. publishing consultancy, The Hague), who gave a brief introduction to the Chinese strategy of 'Going Out'. Since 2005, financial support has been available for translations of Chinese works into more widely used languages such as English and German. The strategy follows a three step approach: firstly, the Chinese want to increase copyright trading and co-publishing agreements ("to sail in a borrowed boat"); secondly, their aim is to enable native publishers to invest outside the Peoples Republic ("to set out in a self-built boat"); and thirdly, Chinese capital should be invested in foreign publishers ("to

purchase a boat for going out”). He explained that western publishers in China (the early movers), are competitors rather than partners, as they tap directly into Chinese scholarship. They utilize their competitive advantage in e-publishing, journal publishing, and international marketing and sales. The current situation of Chinese publishers is characterized by a great variety of co-operations, impressive copyright trading and co-publishing activity – as well as by an improved international presence at events like FIBF and other book-fairs; English websites and English speaking staff. Culture and language, coupled with – from a western perspective – management, could be the hurdles which impede the Chinese aim of going out.

Stephen Boume (Chief Executive, Cambridge University Press) presented a publisher’s experience in his talk *East meets West: Practical problems, practical solutions*. Firstly, he stressed that because of extensive support for scholarly publishing, Chinese circulation figures are much higher, for example, they can be as high as sales of 50,000 for an important book as compared to 500 in the EU and US. Also, Chinese books often cover very specialized topics, so while a book about a Chinese tea ceremony might not have many buyers in the west, it is an appropriate topic for the Chinese market. An unusual aspect is how “truth” is dealt with. A book on the history of China was translated and printed for the western market by Cambridge University Press. Introductory chapters were written by western scholars to provide readers with context. The Chinese agreed to this solution and the Chinese president personally authorized Cambridge University Press to print the translated book. More generally, Chinese publishers or the Chinese government approach western publishers and ask for a translation of their books. A preferable scenario would be one in which western publishers could approach Chinese academics directly and ask them to write up their results for the western market. Currently, the format of Chinese publications presents a challenge. For example, Chinese authors do not use bibliographies. Neither is the notion of copyright well understood (or adhered to). Moreover, it is difficult to find professional translators capable of providing a fluent translation from Chinese to English, so translation has to be a two-step process. First, a technically perfect translation of the text into English is provided – by a Chinese translator. Then an English native speaker transforms the text into idiomatic language. Most of these difficulties could be circumvented if Chinese scholars wrote directly for the international market.

In her talk *Academic publishing in China: Tailored partnerships for common goals* **Dr. Anke Beck** (Senior VP publishing, De Gruyter, Berlin) explained why De Gruyter is involved in the Chinese publishing market. She stressed that China is the fastest growing content provider because Chinese scholars have become fully-fledged members of the scientific community. China’s growing economic and geopolitical influence, accompanied by a desire to promote its culture and language internationally, makes the humanities sector an interesting market with much potential for growth. De Gruyter is looking for long-term strategic partnerships rather than short-term revenue through licensing or distribution agreements and aims to develop common projects from scratch. This means starting with an English product without any need to solve the problems of translation and to be involved in a project from the conceptual stages onwards, ensuring that the eventual product will meet high quality standards. Censorship is circumnavigated by avoiding ‘touchy’ topics. A particular project is the Chinese language atlas. A comprehensive Language Atlas of China was produced 15 years ago on the basis of accumulated fieldwork in China and with UNESCO support. Currently an updated, revised English version, as well as a second volume on minority languages, is being developed in cooperation with Chinese publisher Commercial Press. The editorial board as well as the contributors will consist of both Chinese experts and western sinologists.

Matthias Kaun (Head, East Asia Department, Staatsbibliothek zu Berlin) spoke on *Collection building reality: Challenges and new needs for a Chinese collection in Europe*. Global material is collected by major libraries in Germany, with support from the German Research Foundation (DFG). The Berlin

State Library is responsible for law, Slavic studies and literature, foreign newspapers, topographic maps and cartography, and East- and Southeast Asia. The East and South East Asia Collection includes acquisition, e-resources, digitization projects and technical solutions. For China alone, in 2009, about 580 publishing houses produced more than 300,000 printed items, of which nearly 170,000 are new titles; nearly 240,000 books, of which nearly 150,000 were new books; 9851 journal titles and 1937 newspapers. Moreover, prices are increasing and shifts in the exchange rate further increase acquisition costs. Secondly, new business models are on the rise (print & e, e-only). Furthermore, cooperative projects with Chinese companies are being launched (cataloguing, digitization, input of metadata, special services, building of new databases etc.). However, stable and reliable services from e-providers – especially those who are not the copyright holder – are needed. Yet questions of intellectual property protection and/or open access are unresolved. Hence the challenge of building a collection in the west is a very big one.

The session **Innovation** was chaired by Eefke Smit (Director, Standards and Technology, International Association of STM Publishers, Amsterdam).

Adam Marshall (Group Head of Marketing and Customer Services, Portland Press, London) spoke on *Bringing the journal alive: The Semantic Biochemical Journal*. He presented the idea of Utopia Documents, software through which PDFs come alive by linking to live resources on the web and turning static data into live interactive content. A user can comment on the article and explore article content with an integrated semantic search bar. Live data can be inspected by interacting directly with curated database entries. Together with the University of Manchester, Portland Press has transformed the Biochemical Journal into a Semantic Biochemical Journal.

Victor Henning (Co-founder and CEO, Mendeley, London) talked about *Mendeley: Building the World's largest Semantic Research Database – from Scratch!* Mendeley is a reference manager and academic social network tool that helps researchers to organize their activities, to collaborate with others, and to integrate bibliographies and data. Mendeley enhances collaboration and transparency by inviting researchers to share the knowledge on their computer (which otherwise remains locked up). While basic services are free, the premium version has to be paid for. After 24 months, Mendeley counts more than 700,000 users providing real-time information (e.g., readership statistics, related research) on 58 m research papers. The semantic data created is provided under a creative commons license, allowing the development of apps that mash-up this data with further information (more than 50 apps are under development).

In his speech *Anytime, anywhere: Strategies for mobile content delivery* **Kevin Cohn** (Vice President of Operations, Atypon, Santa Clara) talked about technologies options. He underlined the advantages of using HTML 5 for application development compared to proprietary solutions. He presented WebKit, which is an open source web browser engine developed by Apple, RIM, Google and others. It uses HTML5, CSS3 and JavaScript with high performance and stability. It is integrated in the Safari browser, Google Chrome and every modern smart phone operating system. Cohn acknowledged that scientific publishing on mobile devices is still in the early stages of development, but pointed out that there is a rapidly rising demand in places where devices such as iPads and smart phones are commonly used.

The afternoon session **Open Access: The facts** was chaired by Dr. Salvatore Mele (Head of Open Access, CERN, Geneva). He emphasized that the presentations which followed were all based on rigorous study and empirical data.

Professor **Stuart M. Shieber** (Director, Office for Scholarly Communication, Harvard University Library, Cambridge, MA) spoke on *The rational and empirical basis for open-access initiatives*. According

to Shieber, the empirical evidence available indicates that article-processing charges were highly conducive to the emergence of an efficient market in a way that big subscription deals never were or will be. He added that it was not individual actors who were to blame, but that a scientific information market based on subscriptions tends to be monopolistic, exhibit little demand elasticity, and result in significant moral hazards. As regards a market based on APCs, he suggested that the asking price should correspond to the level of service and quality offered, with cheap and free journals constituting a bottom feed, and that there was competition among publishers with regard to quality, service and price.

Bettina Goerner (Manager Open Access, Springer Science+Business Media, Heidelberg) highlighted some results of *SOAP – Study of Open Access Publishing: What researchers want, what publishers offer* (funded by the EC FP7 – Science and Society). The study analyzes and compares supply and demand for open access publishing, based on a comprehensive analysis of the Directory of Open Access Journals and a global survey with more than 40,000 responses. Currently, the overwhelming majority of open access publishers are very small (one or a few journals, collectively publishing much more than half the total output) and less than ten large publishers can be identified. Fifty percent of the publishers are in STM, but 2/3 of all journals and 3/4 of all articles. Many of the small publishers do not charge, but in the case of larger publishers, income was derived from article processing charges, membership fees and advertisement – with APCs being by far the most important source. The key findings of the survey were that ninety percent of scientists think open access publishing is preferable, but only eight to ten percent of all articles are published in this way. The main barriers seem to be lack of funds (40%) and journal quality and prestige (30%).

Dr. Caroline Sutton (President, Open Access Scholarly Publishing Association) spoke on *Open Access Publishers: Retaining core values while challenging non-core practices*. She said that all publishers were committed to the core function of scholarly publishing, namely the registration, evaluation, dissemination and preservation of knowledge claims. The integrity of the publisher, rigor in evaluation, and a commitment to quality are all essential to this task. Much of what is controversial, such as the business model, metrics and publishing formats, is not actually a core feature. Nevertheless, these ‘additional’ features should be examined to determine how they might best be structured to serve core functions. According to Sutton, this means that publishers are free to experiment with the business model (and other features) to preserve and enhance the core functions for 21st century Internet publishing. OASPA is committed to developing the open access business model as highly complementary to the core functions.

Dr. Mark Patterson (Director of Publishing, Public Library of Science (PLoS), Cambridge) spoke on *Open Access Publishers: Breaking even and growing fast*.

PLoS, and some other open access publishers, have been breaking even and are still growing fast. In fact, the successes of PLoS – one in particular – have been noticed by other publishers, several of whom are striving to re-create this new type of journal. In 2009, PLoS received the ALPSP award for publishing innovation for its editorial process. Editors and reviewers do not seek to ascertain how important a work is and what the particular audience might be, but focus on a technical review of the manuscript. Post-publication tools are then used to sort and filter scholarly content. As regards PLoS on the whole, the number of submissions from 2003 to 2010 has risen from a few hundred to more than 20,000 annually, and the number of publications to about 8000.

Professor Paola Dubini (Director, Art Science Knowledge (ASK) Research Center, Università Bocconi, Milano) spoke on *Complementary article dissemination via journals and repositories: Economic evidence from the PEER project*. PEER (Publishing and the Ecology of European Research) is a collaborative project involving publishers, repositories and the research community, which is investigating

the effects of the large-scale, systematic depositing of authors' final peer-reviewed accepted manuscripts (supported by EC eContentplus). Business models and cost data from eleven publishers and eight repositories were examined on a case study basis, and that of seven large platforms used by publishers and repositories was examined in more detail. It emerged that, for publishers and repositories alike, competition is increasingly based on the platform. Platform setup costs vary widely, maintenance costs much less, and in all cases publishing costs can be significantly reduced if a larger critical mass of content is reached.

The **closing keynote** was given by H. Frederick Dylla (Executive Director and CEO, American Institute of Physics, College Park, MD) on *After the Scholarly Publishing Roundtable: The evolution of the debate on public access*. Mr. Dylla presented his view on the report of the Scholarly Publishing Round Table in the United States, which he co-authored as a representative of publishers. The core recommendation of the Roundtable is that "each federal research funding agency should expeditiously but carefully develop and implement an explicit public access policy that brings about free public access to the results of the research that it funds as soon as possible after those results have been published in a peer-reviewed journal". Most stakeholders expressed respect for the approach and were supportive of the general principles as well as the recommended consultative process. Societies and for-profit publishers had reservations about intellectual property and copyright issues as well as government intervention. In particular, they are concerned about unfunded public access policies, which they believe might threaten the viability of the scholarly publishing enterprise. Also controversial is whether an embargo period of 12 months is sufficient to recoup publisher investment. Additionally, there is no agreement concerning the recommendation that publishers provide access to the article's Version of Record. Advocates of a legislative public access mandate expressed their disappointment because there were no specific recommendations for a public access mandate. They also disagree with the notion that public access articles could be hosted by publishers on their own sites.

The **closing panel** was moderated by **Dr. Herman P. Spruijt** (Immediate Past President, International Publishers Association – IPA, Geneva). Participants were **Mrs Eefke Smit** (Director, Standards and Technology, International Association of STM Publishers), **Robert M. Campbell** (Senior Publisher, Wiley Blackwell, Oxford), Professor **Michael Mabe** (CEO, International Association of STM Publishers, The Hague and Oxford) and **Dr. Matthias Wahls** (m.w. publishing consultancy, The Hague). In his conference synopsis Robert M. Campbell commended APE for the strong program. He stressed that publishers should together develop more consistent data policies. Eefke Smit supported this and said that it was time to move on from experiments and data collection towards sustainable solutions. Matthias Wahls argued in support of downsizing the symbolic value of the journal impact factor and/or replacing it. But he also noted that those funding the research lacked a clear alternative concept and workable mechanisms for measuring quality. Michael Mabe stressed that he found all the sessions of the conference relevant. Yet, he felt that the issue of intellectual property rights – and legal developments at the global and national level – deserved more attention next time. Herman Spruijt said that what interested him was the great age variance among the users of social media, and that increasing usage of this media confirmed the degree to which science is a social activity. Given what we were observing with regard to the Internet – the rise of data-driven science and open access publishing – the journal article seems still to have a very bright future, even though this may be in an embedded form used in novel ways.

Berlin, St. Pölten, Heidelberg, February/July 2011

Please note: APE 2012 will be held 24–25 January 2012.