

Learned Publishing, 24:267–277
doi:10.1087/20110404

PEER: green open access – insight and evidence

Julia M. WALLACE
PEER

Introduction
PEER (Publishing and the Ecology of European Research), supported by the EC eContent^{plus} programme,¹ is investigating the effects of the large-scale, systematic depositing of authors' final peer-reviewed accepted manuscripts (so-called green open access or stage-two research output) with the aim of providing input for evidence-based policy-making in the area of green open access. Launched in September 2008, the project is a collaboration between publishers, repositories, researchers, and funding bodies, and runs until May 2012.

While the publishing and research communities agree on the importance of access to the results of European-funded research, different views are held on whether mandated deposits in open access repositories are necessary and also on what embargo periods would be appropriate for a sustainable system. The question of sustainable embargo periods was also raised in a recent report commissioned by RIN, PRC, Wellcome Trust, RLUK, and JISC in the UK.²

PEER has created an Observatory, operating in real time, to simulate a possible future scenario in which EU-wide, large-scale, systematic archiving is occurring as opposed to the current situation of limited and sporadic archiving. By investigating the effects on scientific publishing and researcher behaviour, PEER will obtain valuable evidence regarding the possible effects of green open access in Europe.

Now entering its final year, PEER has created a robust infrastructure, with substantial volumes of content flowing into the project and out to repositories in support of commissioned research in the areas of article-level usage, author and reader behaviours, and the economics associated with green open access. The project is already generating valuable evidence and experiences that will

ABSTRACT. In support of investigating the effects of systematic archiving of authors' final peer-reviewed, accepted manuscripts (green open access), PEER has developed a robust observatory infrastructure which has already successfully processed over 44,000 manuscripts. Technical challenges successfully overcome by the project include non-uniformity of manuscript files and metadata formats, embargo management, and author authentication for repository deposit. Three areas of research investigating (i) author and reader attitudes and behaviours, (ii) article-level usage at repositories and publisher platforms, and (iii) the economics of large-scale archiving have been commissioned and are producing results. The baseline behavioural survey identified an increasing general awareness of open access, but a lower awareness of institutional and subject repositories. Perceptions were also found to vary depending on whether individuals were responding from an author or reader perspective. PEER itself has seen low uptake from authors when invited to self-deposit into the project. Further results from all three research areas are due before the project ends in May 2012.



Julia M. Wallace

© Julia M. Wallace 2011

Table 1. PEER partners

PEER consortium executive partners	Technical consortium partners
International Association of Scientific, Technical and Medical Publishers (STM)	SURFfoundation
European Science Foundation	University of Bielefeld
Göttingen State and University Library	
Max Planck Society	
INRIA (Institut National de Recherche en Informatique et en Automatique)	

provide insights and guidance for future policy development, and is helping to foster trust and mutual understanding between publishers and the research community.

Participants in PEER

The PEER consortium consists of five executive and two technical partners (Table 1), collectively representing the publishing, repository and research communities and funding bodies. PEER has an overall budget of €4.2m. Up to 50% of overall eligible project costs are being supported by the EC eContentplus programme, with matching funding being provided by the consortium members.

PEER also benefits from the voluntary participation of a number of non-consortium participants, including twelve scholarly publishers, three repositories (in addition to the three consortium partner repositories), a long-term preservation archive and three

individuals who collectively form the Research Oversight Group plus an industry research advisor. PEER also benefits from an 18-member international advisory board with representatives from all key stakeholder groups (Table 2).

The STM publishers participating in PEER are representative of the scholarly publishing community, consisting of large and small publishing houses with a mixture of society publishers and commercial publishers (Table 3). The repositories participating in PEER also provide a representative cross-section of the repository landscape and include institutional, national, and subject-based repositories (Table 4).

The PEER Observatory: infrastructure and content

Organizing the content

To provide the observatory with content,

*the PEER
[represents] the
publishing,
repository and
research
communities
and funding
bodies*

Table 2. PEER advisory board

Mr Mayur Amin, Elsevier, UK	Professor Michel Mareschal, L'Université libre de Bruxelles, Belgium
Professor Peder Andersen, University of Copenhagen, Denmark	Mr Cliff Morgan, Wiley-Blackwell, UK
Dr Paul Ayris, University College London, UK	Dr Elisabeth Niggemann, Deutsche Nationalbibliothek, Germany
Ms Stella Dutton, BMJ Group, UK	Dr Sijbolt Noorda, VSNU, the Netherlands
Dr Johannes Fournier, DFG, Germany	Mr John Ochs, American Chemical Society, USA
Dr Elea Gimenez-Toledo, CSIC, Madrid, Spain	Dr Bas Savenije, Koninklijke Bibliotheek, the Netherlands
Professor Jane Grimson, Trinity College, Dublin, Ireland	Mr Wim van der Stelt, Springer SBM, the Netherlands
Mr Robert Kiley, Wellcome Trust, UK	Dr Donald J. Waters, The Andrew W. Mellon Foundation, USA
Professor Norbert Kroo, Hungarian Academy of Sciences, Hungary	Dr Xiaolin Zhang, National Science Library, Chinese Academy of Sciences, China

Table 3. STM publishers participating in PEER

BMJ Group	IOP Publishing	Sage Publications
Cambridge University Press	Nature Publishing Group	Springer
EDP Sciences	Oxford University Press	Taylor & Francis Group
Elsevier	Portland Press	Wiley-Blackwell

Table 4. PEER repositories

eSciDoc.PubMan.PEER, Max Planck Digital Library (MPDL), Max-Planck-Gesellschaft zur Förderung der Wissenschaften.V. (MPG)
HAL, CNRS & Institut National de Recherche en Informatique et en Automatique (INRIA)
Gottingen State and University Library (UGOE)
University Library of Debrecen, Hungary
SSOAR – Social Sciences Open Access repository
GESIS – Leibniz Institute for the Social Sciences
TARA, Trinity College, University of Dublin, Ireland
Long Term Preservation Archive: e-depot, Koninklijke Bibliotheek (National library of the Netherlands)

participating publishers have collectively provided 241 journals within four broad subject areas: life sciences, medicine, physical sciences, and social sciences and humanities.

PEER was keen to investigate usage and behaviour across a broad range of journals, therefore as well as representation across the four subject areas, journals were selected with a range of two-year impact factors³ as reported by Thomson Reuters Science (formerly ISI). A significant portion of journals with high two-year impact factors are included with representation also of average and low impact factor journals plus some journals which had not yet gained an impact factor at the time of selection. The selection process for the participating journals is outlined within the PEER website.⁴

The participating journals are providing EU-authored manuscripts for the project either by publishers directly submitting accepted manuscripts (and metadata), or by publishers inviting authors to self-deposit their accepted manuscripts, with the publishers providing matching metadata for validation and identification purposes.

An additional, similar group of journals

are not participating openly in the project, but are providing logfile information to the usage research team. In this way, the project benefits from being able to distinguish between background effects in the real world and any specific effects seen within the PEER Observatory. In a similar way, the authors in this group of journals also provide a background measure within the behavioural research aspect of the project.

The PEER Depot

A central facility, the PEER Depot, was created at INRIA, which acts as a clearing house, processing unit, and dark archive for all submitted content. Publishers are providing accepted manuscripts as PDFs and metadata is being provided by the publishers in either one step (on publication) or two passes (on acceptance and on publication), following the project guidelines.⁵

Following an invitation from the publisher, authors can submit their manuscripts as PDF files via a specially created centralized author submission interface within the PEER Helpdesk, hosted by the Max Planck Digital Library. Checks are made at this initial stage to ensure the author has selected a valid PEER journal and the submitted manuscript is sent directly to the PEER Depot for further processing and matching with publisher metadata.

Once content is received by the PEER Depot, publisher-provided metadata, which can be supplied in a variety of schemas including NLM 2.x, NLM 3.0, ScholarOne, and publisher proprietary schemas, is mapped onto a single and well-constrained TEI structure.⁶

The content also undergoes various checking procedures to ensure:

- journal validity (is it a participating PEER journal?);
- EU author;

PEER was keen to investigate usage and behaviour across a broad range of journals

embargo periods for journals participating in PEER vary from 0 to 36 months

- article type (only research-related content is participating);
- matching metadata is provided for each article.

For the manuscripts which pass this filtering stage and are matched with metadata, further checks are undertaken to ensure all mandatory metadata elements, including DOI and publication dates, are present. If they are not yet available, the manuscript is held by the system awaiting the completion of metadata. If all metadata is provided, the manuscript processing is completed and the manuscript is held by the PEER Depot for the embargo period specified for that particular journal. Embargo periods for journals participating in PEER vary from 0 to 36 months. On expiry of the embargo period, manuscripts and TEI metadata are distributed to participating repositories via the SWORD protocol. The PEER Depot is identified as an authenticated source of content for each of the participating repositories. Figure 1 shows the main steps of the PEER workflow.⁷

In addition to the above processes, the PEER Depot has recently developed the capability to extract metadata from PDFs using the GROBID (GeneRation Of Bibliographic Data) environment,^{8,9} which was trained to match various title page styles in scholarly papers. To date, this system has been used to acquire additional metadata elements, e.g. author affiliation, for over 1,500 manuscripts from one of the participating PEER publishers. Further tests of this system are planned. In principle, this process could be used to enhance the metadata record for any PDF collection of articles which follow the style of a scholarly paper.

Content status

Feeds to the PEER Depot from publishers started at the end of the first year of the project in autumn 2009. To help build a critical mass of embargo-expired content as early as possible in the project, back-content was also provided by a number of publishers in addition to the live content feeds. In many cases, the back-content files had to undergo

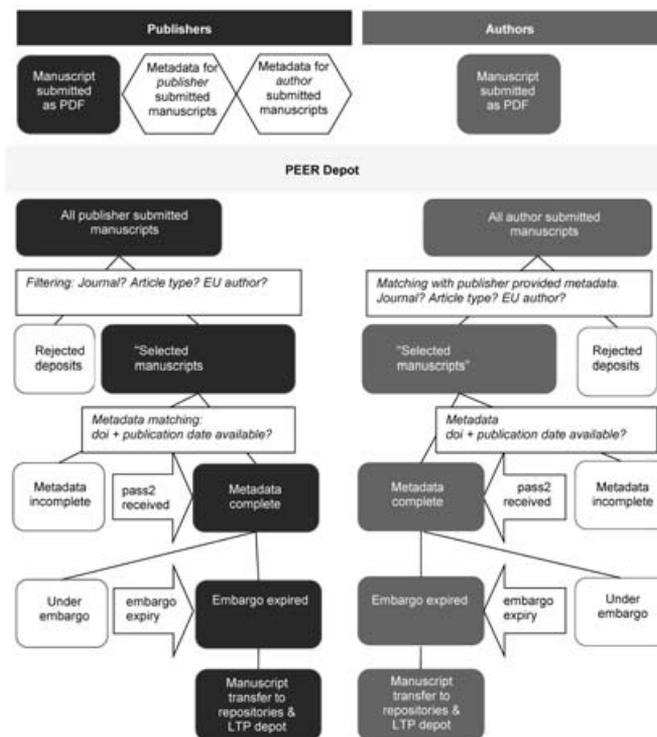


Figure 1. PEER Depot workflow diagram.

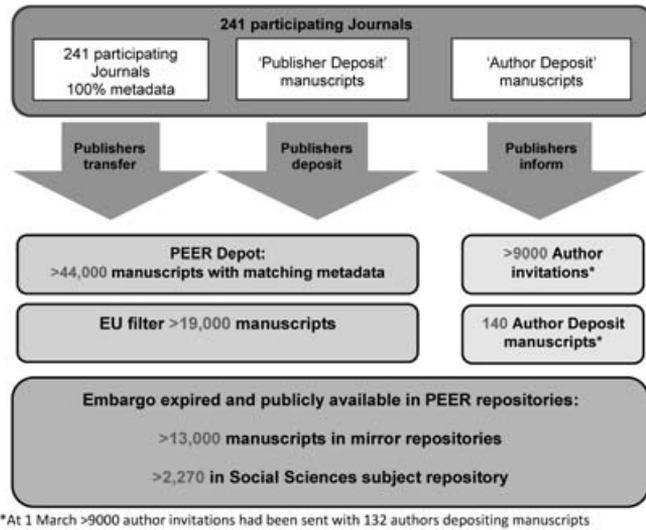


Figure 2. PEER content status at 28 April 2011.

additional processing in order to meet the file format and metadata requirements of PEER.

Figure 2 shows the position as of 28 April 2011, with publishers having submitted over 44,000 manuscripts for processing and filtering by the PEER Depot, of which >19,000 have EU corresponding authors. Of these >13,000 had passed their allocated embargo expiry date and had been sent to the participating repositories. The non-EU, non-research manuscripts which are filtered out are held in the dark archive of the PEER Depot and do not actively participate further in PEER, which focuses on EU-authored research content.

Author deposits

From late 2009 until the end of 2010, within each of the four broad subject areas covered by the journals, 50% of eligible articles were assigned to the publisher deposit route and 50% to the author deposit route. The project had initially expected an author response rate of around 15%;¹⁰ however, throughout the project PEER has observed an overall EU author response rate of under 2%.

Due to the low author response rate observed within the project, at the end of 2010 it was decided to transfer 48 of the journals in the author deposit route over to the publisher deposit route in support of

ensuring a critical mass of content was available in the repositories for the purposes of the usage research. Author invitations by publishers are ongoing for the remaining journals and continue to be monitored.

PEER content in repositories

With the exception of SSOAR, a social sciences 'subject repository', each of the participating PEER repositories hosts all valid PEER content. The availability of PEER content in mirror sites enhances the discoverability by search engines, while the availability of a subset of content in a subject-based repository allows comparisons with usage at institutional and national repositories.

PEER technical infrastructure challenges and sustainability

In the early stages of PEER, it was realized that there were many challenges involved with the transferring of content from publishers or authors directly to repositories in a systematic and scalable way, including:

- non-uniformity of publisher outputs at acceptance stage (file formats/metadata schemas/metadata elements);
- varying requirements by repositories (file formats/metadata schemas/metadata elements);

as of 28 April 2011, publishers have submitted over 44,000 manuscripts

in creating the Observatory, PEER has developed, adapted and implemented a range of tools and technologies, many of which have potential applications outside of the finite duration of the PEER project

- EU and article-type filtering of content;
- embargo management at repositories;
- author authentication for deposit (ensuring authors from outside an institution could deposit);
- non-uniformity of log files.
- format problems with back-content files.

Following extensive consultation with PEER publishers and repositories, solutions were found for all of the above challenges. PEER has successfully developed a robust project infrastructure which has processed over 44,000 manuscripts and has successfully linked feeds from 12 heterogeneous publishers, and over 100 authors (but in principle an infinite number of authors) to 6 heterogeneous repositories.

Publishers provide accepted manuscripts on a daily basis, with content volumes ranging from a few manuscripts to thousands in a given day. After processing, a centralized embargo-management system within the PEER Depot releases content to repositories on a daily basis in accordance with publication dates and embargo periods set for each participating journal.

Since PEER is a practical experiment, many of the solutions identified for the project can potentially be applied in 'real life' scenarios. In creating the Observatory, PEER has developed, adapted and implemented a range of tools and technologies, many of which have potential applications outside of the finite duration of the PEER project, including:

- implementation of the SWORD protocol to allow application-level deposit of material into repositories;
- automated metadata extraction from manuscript PDFs (GROBID);
- establishment of a unique exchange format of metadata (publishers/repositories) by means of a TEI customization plus the mapping of different metadata schemas (e.g. NLM and proprietary schemas);
- the creation of viable workflow models for content submission, filtering, processing, and repository ingest;
- a central embargo-management facility to correctly manage the different embargo periods assigned to each journal;
- additional filtering of subject based con-

tent only for ingest by a subject repository (in the context of PEER this has been applied to social sciences content for SSOAR);

- author deposit interface;
- online helpdesk with ticketing system where queries can be submitted and will be directed to the appropriate PEER project participant for a prompt response;
- the ability to match author manuscripts with publisher-provided metadata;
- the development of a bug tracking and reporting workflow.

These technological developments provide valuable practical outcomes from PEER.

The PEER Observatory: research

PEER has commissioned independent research teams to undertake behavioural, usage, and economic research. Collectively they are addressing such central issues as:

- how large-scale archiving may affect journals;
- whether it increases access;
- how it will affect the broader ecology of European research;
- which factors influence the readiness to deposit in institutional and disciplinary repositories – including perceptions, motivations, and behaviours – and what the associated costs might be

The outcomes from the research, combined with the experiences arising from PEER itself, will inform the development of scenarios to illustrate how traditional publishing systems can coexist with self-archiving.

PEER has benefited from the expertise of independent research experts who form the Research Oversight Group (ROG) plus an industry advisor (Table 5). These experts voluntarily advise the PEER Executive and the research teams on the research tenders, team selection and the design, methodologies and implementation of the research activities. They also validate the research analysis.

The appointed teams undertaking the research are;

- *usage research*: CIBER, University College London, UK;

Table 5. PEER Research Oversight Group

Current PEER Research Oversight Group	Former ROG members
Professor Tomàs Baiget, Pompeu Fabra University and Statistical Institute of Catalonia	Professor Henk Moed, Leiden University
Dr Chérifa Boukacem-Zeghmouri, University Charles de Gaulle Lille 3	Professor Justus Haucap, University of Erlangen
Professor Carol Tenopir, University of Tennessee	
Industry Advisor: Mayur Amin, Elsevier	

- *behavioural research*: Department of Information Science and LISU at Loughborough University, UK;
- *economics research*: ASK Research Center, Bocconi University, Milan, Italy

The final usage research report from CIBER will be made publicly available towards the end of the project (May 2012), while the final behavioural and economics reports arising from the project will be made publicly available via the reports section of the PEER website from September 2011 onwards.

Usage research

CIBER has been commissioned by PEER to undertake the first large-scale and comparative collection of article-level usage. The primary aim of the usage research is to measure any effects on usage of the final published version on the publishers' websites that can be attributable to the mass deposit of earlier versions of the same articles simultaneously across a number of European institutional repositories. In other words: what is the 'commercial' impact (if any) of green open access?

The EU-authored content provided by participating publishers models a hypothetical future scenario in which an EU-wide mandate on self-archiving with a significant level of compliance has become the norm. Three sets of usage data are being collected on a monthly cycle:

- downloads from the publishers' websites for the journals that are part of the PEER experiment;
- downloads from a carefully matched control set;

- downloads from the participating repositories.

The volumes of data are considerable: by April 2011, 860 gigabytes of raw log files comprising hundreds of millions of lines of code had been collected. CIBER is using well-established deep-log analysis techniques to parse and make sense of this data flood, most importantly identifying and screening out visits by spiders and robots that often account for the majority of 'use' of websites.

As well as monitoring monthly patterns of download activity to see broad patterns in the use of licensed and freely available earlier versions, CIBER is addressing three further aims:

- to model any interaction between the length of embargo periods and the use of Stage II versions in each of the four broad subject areas (life sciences, medicine, physical sciences, and social sciences and humanities);
- to try to answer a question for which there truly is remarkably little information in the public domain: does green open access attract new and different audiences for research information?
- to learn more about information-seeking behaviour under green open access conditions and the drivers of repository use.

An original objective of the usage study was to model 'voluntary' self-archiving as well as the mandatory model referred to earlier; however, as reported in this paper, author take-up has been minimal, so this aspect of the study cannot take place.

Throughout the project, material has gradually been building up in the repository.

by April 2011, 860 gigabytes of raw log files comprising hundreds of millions of lines of code had been collected

tories and has recently accelerated to the point where a sufficient critical mass has developed, with over 11,000 embargo-expired manuscripts available in repositories by the end of February 2011. This allows usage to be measured over a 12-month period, since the use of research information follows a highly seasonal pattern (ironically in the Internet age, virtual scholarship still follows the pattern of the agricultural calendar!). CIBER will be preparing an interim report on the first 6 months of green exposure later this year and will follow up with 9- and 12-month reports during 2012.

The live experiment comprises roughly equal proportions of manuscripts from the four broad subject areas and a mix of publication years and embargo periods (mostly 12 months or less, but with a leavening of 18-, 24-, and 36-month periods).

There are already early signs that usage is healthy and picking up momentum in at least two of the repositories. During March, at the HAL/INRIA repository CIBER saw 5,032 full-text downloads of 2,814 PEER Stage II manuscripts; and 4,772 downloads of 2,153 manuscripts at the University of Göttingen. While these levels are small compared with activity at the publisher sites, they are not negligible (no value judgement is being made since it is not known to what socially beneficial ends the manuscripts are being used, just that in absolute terms repository usage is still relatively small).

Untangling the differential effects (if any) of publication year, embargo period plus a whole raft of other indicators such as impact factor, will necessitate some sophisticated experimental design and statistical modelling using a range of multivariate and other techniques. A very relevant and interesting *Learned Publishing* article by Ronald Snijder¹¹ deals a problem not dissimilar to that of the usage research within PEER: what is the impact of the free accessibility of e-book content via institutional repositories and enhanced visibility through the Google Book Search programme on the sales, usage, and citation impact of those e-books? CIBER will be drawing on some of the ideas in Snijder's paper, notably his use of multiple regression to test out various formal hypotheses, as well

as developing further the ideas from their original PEER proposal.

Behavioural research

The behavioural research being undertaken by Loughborough University has been addressing the role of Stage II, accepted manuscript repositories in the scholarly communication system, by exploring perceptions, motivations, and behaviours of authors and users. This complements the usage research which is monitoring actual behaviours at publisher and repository platforms. The information-gathering aspect of the behavioural research has been completed, with analysis of the results underway.

The research has been split into two phases, both of which have now been completed. In the first phase of their research, for which the report is available from the PEER website,¹² the behavioural research team used a combination of a survey and focus groups to explore perceptions, motivations, and behaviours of authors and users towards journal publishing and depositing in repositories. The Baseline Report outlined findings from the first phase of the research, based on >3,000 questionnaire responses and four focus groups, one in each of the broad subject disciplines identified by PEER. For privacy issues, the author surveys were sent by publishers on behalf of the research teams.

Some common perceptions arising from the research:¹³

- General awareness of OA
 - growing compared to previous studies;
 - still considerable ignorance, with over one-quarter unsure of institutional repository, and almost one-half unaware of subject repository.
- Concern over quality
 - authors: reputation of the repository;
 - readers: quality of content.
- Perceptions change depending on whether author or reader perspective

The baseline report established the general attitude and behaviour of authors *vis-à-vis* repositories and journals for a scenario of green open access, and the second phase of the behavioural research has sought to

the baseline report established the general attitude and behaviour of authors vis-à-vis repositories and journals for a scenario of green open access

widen and deepen the understanding of authors' attitudes as well as finding out more about users' perceptions.

At the time of writing, the surveys and a workshop for the second phase of the behavioural research have just been completed. This second phase used a combination of surveys aimed at different groups of authors plus users of PEER content at repositories users followed by a participatory workshop. In this second phase, building from the initial baseline report, the key areas being investigated are:

- to explore the influence of context on reader behaviours in relation to repositories, such as the purpose of reading, e.g. current awareness, proposal and article writing;
- to further investigate reader behaviours in relation to repositories and the characteristics of those readers;
- to investigate the values that authors place on aspects of scholarly communication, such as peer review, prestige, and speed of communication, how these influence current practices in relation to repositories, and how they are shaped by coarse-grained characteristics such as discipline.

For the author-focused survey, there are three main groups of authors. The first two groups of authors have had manuscripts accepted for publication in either (i) a PEER publisher submission journal or (ii) a PEER author submission journal. Both of these groups will have received notification about PEER from the participating publishers. The third author group consists of authors from the control set of journals which are not publicly participating in PEER. These authors will have received no direct communication about PEER from the publishers in advance of the survey.

Economics research

The information gathering for the economics research being undertaken by Bocconi University, Milan has been completed, with analysis now underway. The research has been investigating what the costs associated with archiving Stage II articles under differ-

ent business models are for the various stakeholders involved in article deposit. In support of this, the team has been collecting evidence about the structure and value creation for repositories and publishers in providing green open access.

The research is addressing the following questions:

- *What are the costs associated with research outcome publication and dissemination under different business models?*

This requires analysis of the process of research publication and dissemination and the activities put in place to ensure stability of flow, quality control of the process, accessibility and visibility for the end user, minimization of exceptions, preservation of research outcomes.

- *Which conditions affect costs?*

The cost of performing an activity depends heavily on the cost structure of whoever performs the activity (the impact of fixed and variable costs, the size of the organization, make versus buy choices on specific activities) and on the amount of articles handled (for scale and learning economies). Therefore, the research aims at determining costs for players with various activity configurations and sizes.

- *What are the emerging business models and how they interact within scholarly publishing industry at large?*

Scholarly publishing and dissemination have witnessed significant changes in the past years. The research compares business models coexisting in the industry in order to assess the sustainability and likely evolution of the industry. In addition to the cost structure of different players, this requires taking into consideration resource scarcity, and allocation priorities by different players.

The research team has been gathering data from publishers and repositories in the context of PEER in order to identify cost drivers associated with different activities and the critical issues facing different players within the scholarly communications ecosystem as far as sustainability is concerned. In addition, the team has been assessing the existing evidence and models available from the literature. A preview of some findings

the research compares business models coexisting in the industry in order to assess the sustainability and likely evolution of the industry

regarding the competition or complementarity of publisher and repository platforms and platform costs was presented at the APE 2011 conference.¹⁴

Further information regarding the research commissioned for PEER can be found on the PEER website: www.peerproject.eu/peer-research/

What we have learned so far and what we aim to achieve

PEER is organizationally and politically complex in nature and has benefited from the sustained collaboration of all partners and participants despite any underlying differences in policy they may have regarding green open access. This approach is also fostering respect and trust between participants who represent the wider stakeholder communities involved in scholarly communication.

The proactive collaboration of PEER participants has enabled the project to create a robust and scalable infrastructure in support of a unified and standardized ingestion and distribution service. Although some manual intervention is still required in instances where content is incomplete or delivered with errors, the vast majority of processing is automated and makes use of the impressive array of technical solutions implemented in support of the project infrastructure. The PEER Depot can process thousands of manuscripts in a day, and to date has processed over 44,000 accepted manuscripts for the project.

The behavioural and economics strands of the commissioned research have completed information gathering and following analysis will publicly report on findings around September 2011. A significant volume of accepted manuscripts have been processed, with enough embargo-expired content available to enable the usage research team to measure the full 12 months of activity required for valid and reliable results.

As well as obtaining the results from the commissioned research, PEER will draw from the experiences of project participants to develop scenarios illustrating the effects of archiving on traditional publishing systems. The research evidence and practical

project experiences from PEER will then be available to stimulate discussion and debate on how to maximize the benefits of both green open access and traditional publishing with the aim of supporting the development and refinement of evidence-based open access policies and practices.

Further information on PEER is available from the project website: www.peerproject.eu

Acknowledgements

PEER is receiving matching funding in the amount of up to €2,120,000 from the European Commission eContentplus programme under Grant Agreement ECP 2007 DILI 537003 PEER.

The author would like to thank colleagues within PEER for their help with preparing the article plus the three research teams commissioned for PEER who provided input for the research section.

References

1. The eContentplus Programme. http://ec.europa.eu/information_society/activities/econtentplus/closedcalls/econtentplus/programme/index_en.htm
2. Heading for the Open Road – Costs & Benefits of Transitions in Scholarly Communications. <http://www.jisc.ac.uk/media/documents/publications/reports/2011/dynamicsoftransition.pdf>
3. The impact factor is a measure of the frequency with which the ‘average article’ in a journal has been cited in a particular year or period. For further information see: http://thomsonreuters.com/products_services/science/free/essays/impact_factor/
4. Selection Process – PEER Participating Journals. <http://www.peerproject.eu/about/participating-journals/>
5. D3.1 Guidelines for Publishers and Repository Managers on Deposit, Assisted Deposit and Self-archiving. <http://www.peerproject.eu/reports/>
6. TEI and Scholarly Publishing – Experience from the PEER Project. Laurent Romary, INRIA & HUB-ISDL. http://dho.ie/sites/default/files/events/teipublishing/TEIScholPub_Dublin.pdf
7. Schematic of PEER Depot Workflow after PEER Annual Report D9.8 (2010). http://www.peerproject.eu/fileadmin/media/reports/D9_8_annual_public_report_20100930.pdf
8. Bretel, F., Lopez, P., Medves, M., Monteil, A., and Romary L. INRIA & HUB IDSL, *Back to Meaning – Information Structuring in the PEER Project*, manuscript, published in TEI Conference, 2010. www.peerproject.eu/fileadmin/media/ppt_about_peer/PEERBreakingNews.pdf
9. Lopez, P. GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications. In *Proceedings of ECDL 2009*, 13th European Conference on Digital Library, Corfu, Greece.
10. Steven Harnad, a noted proponent of open access has quoted this figure in several papers and interviews, e.g. Harnad S. 2006. Self-Archiving and Journal Subscriptions: Flawed Method and No Data. Technical Report, ECS, University of Southampton
11. Snijder, R. 2010. The Profits of Free Books: An Exper-

this approach is also fostering respect and trust between participants who represent the wider stakeholder communities

- iment to Measure the Impact of Open Access Publishing. *Learned Publishing*, 23(4): 293–301. <http://dx.doi.org/10.1087/20100403>
12. Fry, J., Oppenheim, C., Proberts, S., Creaser, C., Greenwood, H. Spezi, V., and White, S. *D4.1 PEER Behavioural Research – Baseline Report*. www.peerproject.eu/fileadmin/media/reports/Final_revision_-_behavioural_baseline_report_-_20_01_10.pdf
 13. Excerpt from a presentation made by Dr J. Fry on behalf of the behavioural research team to the PEER Advisory Board and PEER Executive, Berlin, 30 June 2010.
 14. Dubini, P. *Complementary Article Dissemination via Journals and Repositories: Economic Evidence from the*

PEER Project, presented at APE 2011 – Academic Publishing in Europe, Brandenburg Academy of Sciences, Berlin, 11-12 January 2011. www.peerproject.eu/fileadmin/media/ppt_about_peer/APE_presentation.pdf

Julia M. WALLACE

Project Manager, PEER

*International Association of Scientific,
Technical and Medical Publishers (STM)*

Prima House, 267 Banbury Road

Oxford OX2 7HT, UK

Email: wallace@stm-assoc.org