INVITATION TO TENDER   22 December 2008

Usage Research: Journals and Repositories

Deadline for proposals: 17 February 2009

Summary

PEER (Publishing and the Ecology of European Research) **is a pioneering collaboration between publishers, repositories and the research community, by which at least 16,000 peer reviewed manuscripts destined to become journal articles in ISI ranked journals will be made available for archiving every year for three years**. The aim is to investigate the effects of the large-scale deposit (so called Green Open Access) on reader access, author visibility, and journal viability.

At the heart of the project is research to gather evidence about the impact of systematic archiving of stage-two research outputs, i.e. the author's final peer-reviewed manuscript – in contradistinction to the working paper or pre-print (before any submission to a journal) and the final published version. There are three strands to this research: **behavioural, usage and economic.** PEER will fund this research.

This is an invitation to tender for a study that will **investigate usage**. What impact will large-scale deposit have on usage patterns? What is the source and nature of usage of deposited manuscripts? Do the usage patterns reveal new information of relevance to the research community, publishers and/or repositories? May new research information services be derived from usage data? For the purpose of this study **access to log files** of participating publishers and repositories **will be granted**.

The study will run from May 2009 to August 2011 continuously. Tenders are invited from suitably qualified organizations based within the European Union as well as elsewhere in the world, subject to the terms and conditions mentioned below.

Funding

1. Funding in the range of EUR 290,000 to 450,000 (inclusive of VAT and expenses, possibly more for exceptional bids) is available for this study. A suitable profile of payments of the funding will be agreed with the successful bidder following the award of contract.

Background

2. The overall aim of PEER is to develop **an Observatory that will act as a controlled experiment** to compare an evolving scenario of large-scale and systematic archiving with the current situation of limited and sporadic archiving. From 2008 to 2011, at least 50,000 manuscripts will become available for archiving (so-called Green Open Access) and many more manuscripts might become available. Half of the peer-reviewed manuscripts will be **deposited by publishers** in repositories. For the other half, the **authors will be alerted to the possibility of self-archiving** the

manuscripts in repositories and encouraged to do so. This should raise awareness of the issue across Europe.

3. PEER is a consortium formed by the International Association of Scientific, Technical and Medical Publishers (STM), Fondation Européenne de la Science / European Science Foundation (ESF), Georg-August-Universität Göttingen (State and University Library), Max-Planck-Gesellschaft (MPG), Institut National de Recherche en Informatique et en Automatique (INRIA), with the technical partners SURF and Bielefeld University. PEER is supported by the EC eContent*plus* programme.

4. The **key issue** is: What will be the impact of archiving the stage-two research outputs in repositories, if implemented on a broad and systematic scale, on journals and on the wider ecology of scientific research in Europe? Stage-two research outputs are the authors' final manuscripts (after peer review) - the so-called Green Open Access policy. The impact on journal viability of the open access archiving of stage-two outputs (peer reviewed manuscript, also known as the author's final manuscript) has been the subject of lively debate among publishers, learned societies and proponents of Green Open Access.

5. The **proposed solution** is for publishers and the research community to collaborate in developing an "Observatory" that will enable them to monitor the impacts of the deposit of stage-two research outputs. The data provided by the Observatory can then be used to provide an evidence-based foundation for discussion on future policy. Significant about the PEER project is the cooperation of the various stakeholders in the scholarly publishing cycle without prejudice.

6. The study shall contribute to the **Observatory**.

7. PEER has the following **project objectives**:
   - Determine how the large-scale deposit of stage-two research outputs in repositories will affect journal viability;
   - Determine whether the large-scale deposit of stage-two research outputs in repositories increases access;
   - Determine how the large-scale deposit of stage-two outputs in repositories will affect article usage and citing behaviour;
   - Determine whether the large-scale deposit of stage-two research outputs in repositories will affect the broader ecology of European research;
   - Determine the factors affecting the readiness to deposit manuscripts in institutional and disciplinary repositories and measure the associated costs;
   - Develop a model to illustrate how traditional publishing systems can coexist with self-archiving.

8. **Participating publishers** have agreed to collectively contribute 200 to 300 journals from a wide range of types and subjects; setting, however, embargo times for each journal appropriate to the discipline and individual journal economics. The PEER project offers a number of **features** that bidders could utilise in the design of the study:
   - PEER is based on the selection of 200 to 300 ISI ranked journals, from which manuscripts are selected for deposit. Publishers hold a control group of equivalent journals from which no manuscripts will be deposited.
   - From journals selected for deposit, only the manuscripts with European based (lead) authors are selected for archiving, leaving all other articles/manuscripts as an alternate group.

- Half of the manuscripts will be deposited directly by the publisher, but the other half will require action by the author before archiving.
- Authors will be invited to deposit in repositories participating in the PEER project.

9. This study is expected to contribute to the following **outcomes**:
    - Greater understanding by <u>both publishers and the research community</u> of the effects of large-scale deposit of stage-two research outputs in open access repositories, in particular of the access, use, and economics of journals, but also of the broader ecology of research in Europe;
    - <u>Evidence-based guidance</u> for the evolution of policy in this area;
    - A model illustrating the interaction of publishing systems, repositories and libraries to <u>maximise potential benefits</u>;
    - <u>Trust and mutual understanding</u> between the publisher and research communities to assist in the achievement of the ambitious development goals for science in the <u>European Research Area</u>.

Details of the study

10. The **objectives** of the usage research are to:
    - <u>Determine usage trends at publishers and repositories.</u> For participating publishers and participating repositories, usage data will be available for articles in which case a) the manuscript was deposited by the publisher; b) the manuscript was deposited by the author; c) no manuscript was deposited.
    - <u>Understand source and nature of use of deposited manuscripts in repositories</u> (so called Green Open Access). The harvesting of log files will be facilitated at participating publishers and repositories.
    - <u>Understand how embargo periods affect usage patterns.</u> As traditional publishers request an embargo for stage-two outputs, usage (downloading, reading, citing) is delayed.
    - <u>Track trends, develop indicators and explain patterns of usage.</u> A contribution to the new field of usage research is expected.

11. **Design of the study:** Bidders are encouraged to **develop the design most suitable** for the task, including new ideas and innovative methods. While PEER communicates to participating publishers and repositories that complete, raw and unfiltered data are best suited to study usage, it should not be assumed that a universal standard for logfiles exists. PEER is aware that mining repository logfiles is not (yet) common. Accordingly, bidders should make plans, highlight opportunities and assess risks.

12. The design should include specific **research questions** that are to be addressed. The PEER consortium offers the following nonexclusive research questions as reference:
    - Will usage of archived articles be higher than expected for non-archived articles?
    - Will archived articles with an embargo [or by length of embargo] receive less use that those without an embargo?
    - What is the relationship between usage and age of article? How does that differ by subject discipline?
    - Who are the users of repository materials vs publisher site materials? Which countries, which institution types, type of users (transitory or serious)?

13. **Description of work**:
   - Confirm specification for bulk harvesting of log files from repositories and publishers (including identifying risks of not getting minimum standards of requirement from logfiles);
   - Secure the quality of data (for validity and reliability of later analysis);
   - Develop a frame to track usage trends;
   - Develop a frame to understand source and nature of repository usage (possibly including a search for similarities and differences to journal usage);
   - Probe initial data for explanations and indicators of usage;
   - Identify data and trends that are difficult to explain (and resolve by means of further research, e.g. interviews);
   - Consolidate harvesting and frames of analysis
   - Move towards modelling usage for scenarios in which peer reviewed manuscripts are or are not deposited in repositories on a large (or small) scale.

14. Ongoing projects, which focus, in part, on some of the areas addressed by this study and which bidders might find useful to review:
   - COUNTER - http://www.projectcounter.org/
   - SUSHI - http://www.niso.org/workrooms/sushi
   - MESUR - http://www.mesur.org/
   - JISC Usage Statistics Review – http://ie-repository.jisc.ac.uk/250/1/Usage_Statistics_Review_Final_report.pdf

Methodology

15. While for PEER the key issue is the impact of depositing manuscripts, care should be taken in the design of the usage research to **deliver results that are of interest to the research community** in related fields such as scientometrics, informetrics, webometrics and so on. Usage data, for researchers, mean information in real time on usage **trends, patterns and metrics**.

16. Assessment of the experimental design: **Risks for the usage study** (given PEER's experimental design, the harvesting from repositories) should be assessed and mitigating strategies proposed (see above on features of the project, objectives and study design).

17. The main focus should be on **analysis and explanation**, not evaluation or recommendations. In the best scenario, data and analysis will be of real utility to publishers, repositories and research organisations in the sense of helping them to understand the potential long-term consequences of large-scale self-archiving.

18. It is envisaged that the study will be carried out in **five stages**:
   - Interaction with participating repositories and publishers to clarify logfile configurations and harvesting specifications;
   - Background (state of the art, desk research): Initially to be undertaken at the start of the project, but to be kept up-to-date until the final report by digesting new results and literature published in the field;
   - Research (design, raw data): Collection of data should be undertaken so that the complete and original record is available separately;
   - Analysis (hypothesis, data, results): A scientific analysis is required, but it is not expected that policy recommendations will be derived;
   - Reporting (on all of the above): An interim report, a working report and a final report are required (see below).

19. It will be of fundamental importance to **unpack the notion of usage**, particularly in relation to other important concepts of scientometrics (informetrics, webometrics) such as citation and access. It should not be assumed that the meaning of usage is always identical for publishers, repositories, libraries and researchers, but comparability of data is important.

20. Throughout the research stage it is essential that **a full and unadulterated record of all data is kept and stored.**

21. Bidders are expected to provide evidence of existing, constructive links with universities and the research community and **identify imaginative ways of ensuring participation** and involvement of this category of users.

Outputs and Timetable

22. The outputs and timetable for this study will be:
    - **Interim report**: report on state of the art (knowledge, method, theory); report on spec for harvesting and frames of analysis; initial data quality; pointers for analysis. Timing: 11/2009
    - **Working report**: The PEER Observatory requires, for the purpose of model building in the final months of the project, a full working report on background, research and analysis, especially an analysis of the results against the research data obtained and in light of the prior state of the art. Timing: 01/2011
    - **Final report:** As usage data continues to come in, the working report should be updated into a final report. It must include the final and complete set of raw research data. Timing: 08/2011

23. One **aim of the Observatory is to integrate the results of the studies on behaviour, usage and economics**, which are to be utilised for building a model of the interaction between publishing systems, repositories and libraries in an effort to enhance the ecology of European research. A joint workshop of the research teams, expert groups and the PEER consortium will be scheduled mid-way in the project (approx. March 2010). The research teams will be invited to comment on the model building (approx. May 2011).

Terms and conditions

24. STM is the leader of the PEER work package on usage research.

25. The research will be overseen by the **Observatory**, led by Chris Armbruster, Max Planck Digital Library, Max Planck Society. It is expected that the conducted research will lead to high-quality publications and may inspire further frontier research on the question of usage for digital scholarly communication.

26. Bidders should be aware that issues pertaining to data access, confidentiality of information and anonymity (users, data providers) will be arising in the course of the research. Handling of these issues will be specified in the contract.

Structure and format of proposal

27. **Tenders should be no more than 15 pages in extent**, plus CVs. They should include:
    - A description of the proposed work;
    - The proposed methodology for carrying out the study;

- A detailed programme of the work to be undertaken, including the methodology and a work plan showing deliverables;
- A summary of relevant experience to undertake this project and evidence of engagement with the stakeholder community;
- An assessment of the risks associated with the project and how these will be managed;
- The total cost and a breakdown of costs according to task, including a statement on VAT. Staff costs should be broken down into the estimated number of days to be contributed to the project by each person / percentage of FTE. Calculation of other costs, such as travel and subsistence; hardware and software (if applicable) should also be clarified. Any institutional contribution should be clearly identified (see the budget spreadsheet template attached);
- A covering letter from a senior member of the organisation (Head of Dept. or higher in the case of an HE institution) indicating support for the bid and strategic fit with the organisation/institution remit;
- Short CVs (one page max.) of key personnel proposed for this project, including relevant experience and qualifications.

28. Proposals should adhere to the following format, **with the headings listed below**:
- Concept and objectives
- Methodology
- Workplan (including timeline and deliverables)
- Risk assessment
- Management structure
- Participants
- Resources to be committed
- Expected results
- Contribution of outcome to PEER

Evaluation Criteria

29. Submitted proposals will be evaluated by PEER. The final decision on whom to award the contract is taken by the Executive of the PEER consortium. The following criteria will be used for evaluation:
- Knowledge and understanding of the area and related issues;
- Engagement with the community;
- Methodology;
- Innovativeness of approach and method;
- Potential to contribute to PEER objectives and outcomes;
- Project management and project plan;
- Relevant experience;
- Risk assessment & management;
- Value for money.

30. Tenders should demonstrate:
- Knowledge and understanding of the area of work, issues to be considered and the community to be surveyed;
- Methodologies to be adopted;
- Clear project management set up;
- Clear outputs and deliverables;
- An appropriate mix of skills for the project, including rationale for collaboration in the case of consortia bids;

- The risks associated with the project have been considered;
- An outline of the areas to be covered in the final report;
- Clear, easy to understand costs for the project, with a clear rationale provided.

Submitting a proposal

31. The deadline for receipt of submissions is 17:00 on Tuesday, February 17, 2009. Late proposals will NOT be accepted. It is the responsibility of the bidder to ensure that the proposal has arrived by the deadline stated. **Bids must be submitted in a sealed envelope containing five paper copies of the proposal and at least one electronic copy of the bid** (stored on a machine readable device like a CD ROM or USB stick). All envelopes will be opened on 18 February 2009. Bidders should send the sealed envelope by registered mail, marked as "PEER usage," to the following address: Max Planck Digital Library, Invalidenstrasse 35, D-10115 Berlin.

Award of contract

32. It is anticipated that bidders will be notified of the outcome of this tender exercise on or before 31 March 2009.

33. PEER will expect to work with the successful bidder to agree the work plan. The kick-off meeting with the PEER consortium is anticipated for April 2009. The selected team will be introduced to the Research Oversight Group of the PEER project:
- Justus Haucap, Professor of Competition Policy, University of Erlangen. Prof. Haucap chairs the German Monopolies Commission;
- Henk Moed, Senior researcher at the Centre for Science and Technology Studies, Leiden University. Dr. Moed has been the recipient of the Derek de Solla Price Award;
- Carol Tenopir, Professor of Information Sciences, University of Tennessee. Prof. Tenopir has received the International Information Industry Lifetime Achievement Award.

34. PEER shall be under no obligation to accept the lowest, or any tender, and tenderers shall submit offers on the basis of so doing at no cost to PEER.

35. Proposals received from tenderers outside an EU Member State or an associated State will be considered, but in accordance with the eContent*plus* policy, would only be successful if it can be demonstrated that no other subcontractor from a Member State or associated State can provide a service comparable to the one offered by the subcontractor from a third country.

36. Bidders are advised that tendering for one strand of PEER research does not preclude a bid for another strand of PEER research, but a separate bid must be made for each tender. PEER endeavours to select the best bid for each strand of research.

Further information

37. For enquiries regarding this study, including information on participating publishers and repositories, please contact the PEER Observatory: Chris Armbruster, Max Planck Digital Library, Max Planck Society. Email: armbruster@mpdl.mpg.de Tel: +4930 288867788

38. For enquiries regarding the bidding process, please contact Julia Wallace,
E-mail: wallace@stm-assoc.org Tel: +44 1623 472 91