

Louise Allsop
Project and Communications Officer, VERSIONS Project
The London School of Economics and Political Science

VERSIONS

Versions of Eprints – user Requirements Study and Investigation Of the Need for Standards

Introduction to the VERSIONS Project

The UK Joint Information Systems Committee (JISC)-funded VERSIONS Project addresses the issues and uncertainties relating to version identification in digital repositories and open access research paper collections. The project forms part of JISC's Digital Repositories Programme¹ and is running from July 2005 to January 2007.

Previous work on institutional repositories had suggested that the question of version identification could be worthy of further investigation. Stephen Pinfield and Hamish James raised the possibility of preserving only specific versions – postprints rather than preprints, for example. They also noted that the demand for certain versions will probably differ by subject area citing the likelihood of 'local priorities for preservation'². Rachel Heery and Sheila Anderson's 2005 Digital Repositories Review, used to provide a background for participants in JISC's Digital Repositories Programme, noted 'issues as to how different repositories deal with versioning of preprints', calling for further investigative work in the area³. Anecdotal experience from LSE has suggested that authors may be unsure about what versions are required or allowed for deposit, or that they may not have retained an appropriate postprint version following publication in a refereed journal. Equally readers may be unsure about exactly which version of a paper they are reading online. The Research Councils UK Position Statement on Access to Research Outputs, proposed in 2005, cited this issue and called for repositories to make clear distinctions between versions. They recommended that postprints and different preprints be labelled discretely and appropriately⁴.

The VERSIONS Project is looking at how academics produce, archive, disseminate and access electronic versions of papers at different stages in their lifecycles, and at researcher's views on the current situation. The project has a focus on eprints in the subject discipline of economics, and takes a comparative view by drawing on established partnerships and experience with European libraries specialising in economics.

¹ JISC Digital Repositories Programme

http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories

² Pinfield, S. and James, H. (2003) The Digital Preservation of e-Prints. D-Lib Magazine 9 (9), September 2003 <http://www.dlib.org/dlib/september03/pinfield/09pinfield.html> (Accessed 14/06/06)

³ Heery, R. and Anderson S. (2005) Digital Repositories Review

http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf (Accessed 14/06/06)

⁴ RCUK Position Statement on Access to Research Outputs (2005)

<http://www.rcuk.ac.uk/access/statement.pdf> (Accessed 14/06/06)

The project is being led by the London School of Economics and Political Science, with the Nereus Consortium of European research libraries in economics⁵ as associate partners. Nereus is a collaborative project of leading European economics research libraries, seeking to integrate access to a valuable set of information resources for economists. The consortium consists of sixteen partners representing nine European countries, and is committed to developing institutional repositories and an economics subject gateway – Economists Online. Nereus places a large emphasis on the provision of open access content and is providing valuable support to the VERSIONS Project team.

Project aims and objectives

Aims:

- to clarify the position on different versions of academic papers in economics, and their availability for deposit in digital repositories, in order to help build trust among academic users of repository content
- to produce a toolkit of guidelines about versions for authors, researchers, librarians and others engaged in maintaining digital repositories
- to propose standards on versions to JISC to inform discussions and negotiations with stakeholders

Objectives:

- to learn about researchers' current understanding of different versions of academic papers, and their attitudes towards electronic archiving, open access, and dissemination at key stages during the process
- to gather specific information about current and anticipated future repository use; are existing repositories being well used, by whom, and if not, what are the reasons for this?
- to uncover current practice among academic researchers relating to retention of their own author versions, and the availability of these versions for deposit into open access repositories
- to develop a set of best practice guidelines relating to version identification and propose as a standard to JISC.

Overall approach

Current attitudes and practice in relation to version identification are being investigated via two main work packages:

- A user requirements study
- A publications list analysis

Information gathered during the above investigations will be used to shape the next work package:

- Development of guidelines and standards

⁵ Nereus Consortium <http://www.nereus4economics.info>

Progress to date

Interviews with a number of European academics, repository staff and students have been completed. The interviews were designed to ascertain information about current usage of, and attitudes towards, open access repositories. In addition, information about authors' retention of their own work was obtained.

The main findings were as follows:

- Authors retain many versions of their own work, most of which are not intended as public versions. Authors may actively seek to keep some draft, early or tentative versions out of the public domain. However, the time lag between article submission and publication in a peer-reviewed journal may encourage the use of other dissemination outlets – this is particularly true in the case of economics.
- Dating versions is crucial. The date is a simple way to identify the latest version of others' work. At present it is not always included / obvious in online documents.
- It is essential to be able to identify definitive version(s) and point to the journal article for citations. Citing the journal article version benefits the reputation of authors and adds weight to the argument of those wishing to cite the research.
- Authors do not always possess an electronic verbatim copy of the peer-reviewed published article, as final corrections may have been made on hard copy proofs.
- Authors often keep all versions but cannot quickly identify or locate the latest manuscript version.

The interview findings were used to develop the more wide-reaching online questionnaires. Responses have now been collected from two surveys; one addressed to researchers as authors and as readers; the second addressed to other interested parties such as library staff, the repository software community, publishers, research funders and university senior management.

The questionnaire for researchers looked into the researcher as an author and as a reader. The author-based questions focused on thoughts and actions in relation to storage and dissemination of their own research, and depositing their own work in open access repositories. As readers, the researchers were asked about their attitudes towards locating and using open access online versions, and to assess possible ways of making version identification easier for the reader.

The survey for other interested parties focused on the level of importance of version identification, and different aspects of version identification, for different user groups. Respondents were invited to consider the value of potential solutions to the versioning issue and the performance of current repository software with regard to version identification.

Both surveys presented possible solutions to the version identification issue that included the value of a standardised terminology for versions, a date-based approach to file naming, possible linking of related documents, and a free text descriptive approach to version identification. Respondents were asked to assess the above and asked to explain any of their own ideas about how to improve on the current position.

At the time of writing more than 450 participants had responded to the online questionnaires, most of the respondents being academic researchers. Analysis of the

preliminary results has suggested that version identification is an area of confusion for many, authors and readers alike. A number of respondents have expressed a dissatisfaction with the way that they organise their own revisions and versions of papers. Though many others report having developed sophisticated systems of archiving, most do not have a full understanding of which versions of papers can be disseminated at which times, and a significant number feel that it is not easy to establish the status of papers that they locate online.

The full results are currently being collated and analysed.

The publications lists analysis looks at the research collected together in the Nereus consortium's Economists Online gateway⁶. The study establishes which journal articles are available in Economists Online as open access full text copies, and goes on to look at the source of the document posted (postprint supplied by author, scan of published article etc.). It aims to draw conclusions on full text availability, date of full text material, provenance of full text versions and copyright positions of relevant publishers, and to compare the situation in different EU countries. Findings are currently being finalised.

Future work packages

The survey has yielded a large amount of numeric data, along with many free text comments. During the next few months the findings will be analysed and a findings report drafted. The findings will be used as a basis to draft and publish a toolkit of guidelines advising academic researchers on best practice for version control. These guidelines will be proposed as standards to JISC.

All work packages will be evaluated and all results will be disseminated widely to the JISC community.

Further Information

For further information, please visit the VERSIONS website at www.lse.ac.uk/versions or contact the project team (Frances Shipsey – Project Manager and Louise Allsop – Project and Communications Officer) directly at versions@lse.ac.uk.

⁶ Economists Online <http://www.nereus4economics.info/eonline.html>