



Versions of eprints : understanding researchers' needs

Frances Shipsey, VERSIONS Project, Library, London School of Economics and Political Science

Presentation to CAB International Staff

Wallingford, 8 February 2006

JISC



Outline

- The Library of London School of Economics and Political Science
- Digital repositories
- Issues relating to versions in digital repositories
- The VERSIONS Project
- Other initiatives on versions
- Managing versions in other contexts – examples of approaches
- Possible approaches for versions of eprints

London School of Economics and Political Science – the Library



Redevelopment by Norman Foster Architects in 2001



Open plan layout with clear sightlines



Concentration of study space and access to networked PCs



Printed collections



All collections held at LSE in the Library - recognised for their outstanding national and international importance and awarded 'Designation' status by the Museums Libraries and Archives Council (MLA)

Digital repositories

- Open access principles – Budapest Open Access Initiative (BOAI) - <http://www.soros.org/openaccess/>
 - Purpose of Budapest Meeting in 2001 to ‘accelerate progress in the international effort to make research articles in all academic fields freely available on the internet ‘
- Institutional repositories
- Subject-based repositories
- Repositories based on type of content – theses, images, learning objects
- Directory of open access repositories OpenDOAR <http://www.opendoar.org/>
- Gerard van Westrienen and Clifford Lynch. Academic Institutional Repositories : Deployment Status in 13 Nations as of Mid 2005. D-Lib Magazine 10 (9) September 2005 <http://www.dlib.org/dlib/september05/westrienen/09westrienen.html>

JISC Digital Repositories Programme

- Joint Information Systems Committee
 - 'JISC works with further and higher education by providing strategic guidance, advice and opportunities to use ICT to support teaching, learning, research and administration'
- Digital Repositories Programme builds on earlier programmes including FAIR (Focus on Access to Institutional Resources)
- Digital Repositories Review informed call for proposals in 2005
- 21 Projects funded to start in 2005 – covering a range of issues
- http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories

Issues relating to versions

- Copyright
 - generally cannot use publisher versions of PDFs, but see Cambridge University Press
 - copyright assignment for journal articles is the norm
 - major publishers allow authors to place ‘author final versions’ on personal websites or in institutional repositories – ROMEEO listing at www.sherpa.ac.uk/romeo
 - authors’ rights vary in different EU countries
 - Availability of ‘final version’
 - variations between latest author manuscript and published version
 - final corrections made by hand to publisher PDF proof

Issues ...

- Collaboration and co-authoring
 - naturally leads to proliferation of internal 'non-public' draft versions
 - when collaboration is inter-institutional may lead to proliferation of public versions – paper is placed in working paper series of each institution and on co-authors' personal websites
 - effort needed by all authors to keep versions in line with each other
 - co-authors need to make efforts to obtain final version copies from the Corresponding Author

Issues ...

- Administration of different versions – objects and metadata
 - authors may already be posting their papers in a number of different locations – personal website, departmental website or working paper series, subject-based working paper series, subject collections such as RePEc, university database of research outputs, and now institutional repositories
 - keeping versions up to date in all of these locations is burdensome – authors may take pragmatic decision to update only personal website
 - different versions of documents – single metadata record or separate
 - role for libraries and repository managers – streamlining processes

Issues ...

- Informal communication and time lag in publication
 - 2-3 year wait for publication of journal article is common in economics
 - Preprint and revised versions up to final author postprint are important phases in the scholarly communications cycle
- Impact/visibility and citations
 - Conference, workshop and seminar papers – important for disseminating work quickly and obtaining feedback – but hope for citation of working paper or journal article; hence abstract only conference proceedings
 - Loss of impact if citations are dispersed between published article, working paper and postprint
 - Evidence about the impact of open access articles as compared with non-OA articles. Stevan Harnad and Tim Brody. Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. D-Lib Magazine, 9 (6) June 2004

Issues ...

- Development of a unit of intellectual content
 - Herbert Van de Sompel (et al) (2004) Rethinking scholarly communication : building the system that scholars deserve, D-Lib 10 (9) – ‘Recording the dynamics of scholarship’ – relationships between different versions in the lifecycle of a paper are clear during the process, but difficult to recapture after the event
 - How many versions to keep
 - Role of corrections, comments and discussion in the lifecycle of a paper
 - Journal article definitive and final version, but working papers and technical papers also definitive, can reach a final settled version and have a distinct existence in own right
- Quality issues - peer review
 - Labelling different versions to indicate whether refereed, eg in eprints.org
 - Peer review process – is a mark of quality
 - Open access repositories aim to retain this mark
 - Other quality indications are
 - track record of author
 - reputation of the authors’ institution

Issues ...

- Authors' control over versions of research outputs
 - Placing a copy on personal website – authors retain (fairly) full control
 - Posting a copy in a repository – who decides about removing earlier versions, authors, repository managers, or university administration
 - Awareness that once posted on the web, papers cannot easily be withdrawn
- Persistence
 - Pre-prints and postprints commonly cited – possible implications for cited author or citing author if the referenced work changes or disappears
 - Revision vs persistence – a tension
- Digital objects but still mirroring print process
 - Multiple versions – not well linked to each other – still following model of print - snapshot PDF views of evolution of the paper

The VERSIONS Project

- VERSIONS : Versions of Eprints – user Requirements Study and Investigation of the Need for Standards
- Funded by the Joint Information Systems Committee (JISC) under the Digital Repositories Programme
- London School of Economics and Political Science (LSE) - lead partner
- Nereus – consortium of European research libraries specialising in economics – associate partner
- July 2005 to January 2007

A library network



Aims of the VERSIONS Project

- to clarify the position on different versions of academic papers in economics available 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

Focus of the project

- Economics
 - Established pattern of using preprints
 - Importance of journals coupled with long lead times for publication
 - Builds on existing experience and partnerships
- Eprints
 - Builds on experience of other projects and programmes
 - Not looking at data or other object types
- Europe
 - Cultural and IPR differences worth investigating
 - Builds on existing experience and partnerships

Approach of the VERSIONS Project

- User requirements study
 - Talking to authors, researchers, librarians, repository software developers, relevant standards communities, and other stakeholders
 - Attitudes and current practice
 - Interviews, online survey and evaluation of user study
- Publications lists analysis
 - Analysis of publications lists of 70 economists in Economists Online repository, looking at availability of full text
 - By date of publication
 - By country
 - By publisher/self-archiving policy
 - Location of full text copy

Guidelines and standards

- Reaching consensus with stakeholders
- Development of guidelines on good practice
- Production of a toolkit for researchers
- Recommendation of standards

Interviews

- Interviews held during October 2005 to February 2006
- Held at
 - KU Leuven, Belgium
 - Universite Libre de Bruxelles, Belgium
 - WU-Wien, Austria
 - Maastricht, Netherlands
 - ZBW-Kiel, Germany
 - LSE, UK
- Interviewees
 - Professors
 - Lecturers
 - PhD students
 - Library staff
 - Secondary publisher
 - Research information service staff

Initial results from interviews

- Authors retain many versions of their work
- Most of these are not seen as public versions
- May actively seek to keep some versions out of the public domain or to control their use
 - Very early drafts circulated between co-authors
 - Results that are early or tentative (conference presentations)
 - Evidence of rejected journal submissions
- Delay in publication of peer-reviewed articles contributes to the use of other dissemination outlets

Initial results from interviews

- Posting papers in multiple locations – administrative effort to update in each location
- Collaboration with co-authors requires additional effort to manage versions
- Use of date is crucial; simple way to identify latest version of others' work would be excellent – a 'non-obvious task' at present
- Essential to be able to identify the definitive versions and to point to journal article for citations
- As authors citing others' work in publications
 - Adds weight to argument through citing published articles rather than work in progress
 - Administrative burden of checking citations when article is submitted and again after acceptance
- Trade-off between wide dissemination and control over versions
- As readers – finding broken links in articles is problematic

Other initiatives on versions

- NISO/ALPSP Working Group on Versions of Journal Articles
 - Policy paper from February 2005 by Sally Morris
 - Two groups – technical working group and review group
 - Developing use cases
 - http://www.niso.org/committees/Journal_versioning/JournalVer_com.html
- JISC scoping study
 - RIVER : Scoping Study on Repository Version Identification, commissioned by the JISC Scholarly Communications Group – to report in March 2006. Lead partner: Rightscom Ltd. Partners Library of LSE and Oxford University Computing Services

Versions in other contexts

- Literary criticism
 - Shakespeare – quarto and folio editions
 - Textual transmission
 - Variorum editions have existed in the print age – new initiatives to produce electronic variorum editions
- Software development
 - Concurrent versioning systems used by developers
 - Release numbering systems
 - Divergence of software – local customisation by or for clients
- Continuous updating
 - A problem in print environment, eg legal encyclopedia looseleaf publications – earlier versions lost
 - A solution in the fully digital environment? - earlier versions can be retained and compared
 - Cf Wikipedia

Possible approaches

- Labelling
 - Reach consensus on terminology – naming conventions
 - Numbering systems, cf software release
- Describing
 - Author annotations and free-text descriptions in metadata
 - Date
 - Cover sheets – standard template
- Linking
 - MARC21 Linking entry fields (76X-78X)
 - IFLA Functional Requirements for Bibliographic Records (FRBR)
 - Dublin Core element - relation
 - Dublin Core refinements
 - hasVersion, isVersionOf
- Comparing text
 - Formats
 - Tools
- Signposting
 - To published journal version (for citations)
 - To author-approved latest version of pre-print (for full elaboration of argument, proof, supporting data)

Consultation

- The VERSIONS Project is seeking views and opinions on handling of multiple versions from:
 - researchers and authors
 - software developers
 - metadata community
 - librarians
 - publishers
 - university research information staff
- Comments and views welcome

Contacts

- VERSIONS Project
 - versions@lse.ac.uk
 - www.lse.ac.uk/versions
- The Library at LSE
 - www.lse.ac.uk/library
- Nereus
 - www.nereus4economics.info
- JISC Digital Repositories Programme
 - http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories

