

10° Fiesole Collection Retreat
Badia Fiesolana, March 27-29 2008

“Scholarly Communication in the Network: Collecting and Collaborating Where our Users are”

The author's workflow and the role of Open Access

Paola Gargiulo
gargiulo@caspur.it

Summary

- OA: what's going on
- OA challenges
- Author's perspective
- Tools and services to support author's deposit, evaluation/assessment activities
- Some examples
- Conclusions

OA: what's going on

- Mandatory policy: constant growth of institutions and funding agencies implementing it

- EC bodies taking stand on OA issues
 - ERC - European Research Council statement(2006);
 - ERC Guidelines (2007)

- Professional and institutional associations are taking action
 - EUA- European University Association Council – Recommendations on Open Access (January 2008)
 - Call for requirement to deposit in institutional archives
 - Call for exploring how university resources can be found for author fees if “author pays model” of OA publishing prevails in the future in some scientific fields

OA: what's going on (2)

- Eu Rectors Initiatives to support OA Scholarship (EurOpenScholarship)
- Infrastructure and services at national and crossnational level
 - JISC Initiatives, SURF/Darenet, DFG, Nordic Countries,
 - DriverI, DriverII
 - OOpen Initiative (Oa publishing in Humanities and Social Sciences)
- Copyright : license to publish (SURF/JISC), Author's Addenda, SPARC and DOAJ licence standard; FAS – Harvard Action
- Increase of support material and tools: guides, tutorials, wiki, websites, blogs, fora
- Increasing number of institutional repositories and OA journals and of hybrid journals

OA: critical issues – Institutional repositories

- Content recruitment
- Authors awareness and support
- Quality control (vetting activities, metadata, etc)
- Copyright issue (adoption of author addendum, license to publish, etc)
- Long term preservation

OA: critical issues – OA journals

- Move from Toll Access to OA journals
- Economic sustainability of OA publishing
- Different business models
- Role of libraries in supporting OA publishing

Studies and surveys on OA: what is missing

- Several articles and surveys report authors' perceptions, attitudes to self-archiving, OA, etc
- Surveys conducted on institutional repositories implementations, testing, planning
- Studies/surveys on librarians perceptions, on publishers, professional society attitude on OA
- Fewer studies on what users really need, want in the OA context
- Even less research has been carried out on authors' needs on tools, services, add-on to help researchers in their work

Author's general attitudes on institutional repositories

- Depositing is cumbersome (submission workflow)
- Resistance to change work habits
- Fear of duplicating work
- Lack of time
- Can someone else do it for me?
- What other benefits can I get beside wider and faster dissemination, increased visibility, more research impact?

Basic requests from authors

- Make deposit mechanism easier, as automatic as possible
 - Batch ingestion of metadata from personal web pages, academic portals, from personal bibliographic databases
 - Easy desk-top application to deposit
 - Integration with National Authority lists
 - Integration with publishers' copyright policy
- Integrating IR with local or national Research Information System
- IR as an updated bibliography, a documents server, a publishing tool

Basic requests from authors (2)

- Export functionalities: bibliographies or other output for research or administrative purposes
- Notification services (e-mail alerts, RSS etc) and Web 2.0 functionalities
- Usage statistics, citations analysis
- More interoperability
- Better discovery and searching tools (semantic web, text mining, data mining)

Making deposit easier and more efficient: a lightweight protocol

- SWORD : **S**imple **W**eb-services **O**ffering **R**epository **D**eposit project funded by JISC Repositories and Preservation Programme (2007)
- Goal: to develop and test a standard mechanism for depositing into repositories and other systems
- Project carried out by Repositories Research Team a collaboration between UKOLN and CETIS
- SWORD has been implemented with Arxiv, with ICE (Integrated Content Environment), ROAD (Robot Generated Open Access Data)
- Interoperable with main IR softwares

SWORD: what it does

- It supports workflows for deposit in different ways which can be user triggered or machine triggered:
 - **A.** Author to multiple repositories via intermediate client
 - **B.** Author to 1 or more repositories,
 - **C.** Deposit of experimental data output from lab machine is started as save as function
 - **D.** Repository to additional repositories
- **A.** Author/user deposits via an easy-deposit desktop application into the institutional repository's mediated deposit queue
- Librarian completes the deposit through repositories
- Librarian invokes deposit of a surrogate into a discipline based archive researchers in their communication workflow

Source: J. Allison, S. Francois, S. Lewis. SWORD: Simple Web Service Offering Repository Deposit, "Ariadne" January 2008

NTU –IR Metadata Importing tool

- Nanyang Technological University Institutional Repository – Metadata Import Tool
- Extraction metadata and digital resources from academic information portals
- Ingestion into IR
- Semi-automatic ingestion
- Batch ingestion into IR

NTU –IR Metadata Importing tool (2)

- Extract metadata from staff webpage
- The Url object establish a connection to the staff webpage
- Specify sections to extract
- Extract metadata (author, date of publication, Title of publication, url of digital resource) to an excel file using Java

NTU –IR Metadata Importing tool (3)

- Extraction from excel to DSpace Archive format
 - Main archive folder name is specified at run time
 - A connection is made to the Metadata Excel Sheet
 - Archive folders and files are automatically generated from the Excel Sheet for each publication
 - Dublin core file
 - Content file
 - Actual item retrieved from a web page
- NTU -IR Import tool generates Archive Folders and files
- Importing metadata and Content in NTU –IR
- Future developments: extraction from generalized citation styles, move to a full automated extraction method

SURplus: integrated system for evaluation

- Sur+ is an integrated system for research management that aims to control processes and data related to research activities and their outputs in the academic environment developed by CILEA, Italy
- Sur+ enhances the internal and external visibility of research activities and provides real-time monitoring and reporting tools
- Sur+ is formed from several interoperable application modules that can be adopted separately.

Source: Susanna Mornati, Andrea Bollini. New opportunities for institutional repositories: the evaluation challenge. A case study. DSUG Meeting, Rome October 17-19 2007

SURplus OA: Open Archive Module

- The Open Archive module is a repository (DSpace) that collects, describes and indexes, and distributes the intellectual output of all institutional research activities
- Data are imported from bibliographic sources (e.g WOS, PubMed,...for initial metadata population)
- Data are exported via OAI-PMH and web services to the web and to other external targets such as any CRIS (current Information Systems)
- Data are transferred to or used by other Sur+ modules for further elaboration

SURplus OA: workflow (1)

- o Authors connect through SUR+ authentication system
- o Authors' data (names, affiliations, IDs) are automatically extracted from SUR+staff entities modules
- o Customized workflow to organize submission and review according to the university departments
- o Only institutional authors appear in the browse list (authority file), all authors are searchable in the word indexes

SURplus OA: workflow (2)

- Submission according to material type (customised templates follow metadata requirements for the Italian RAE)
- Controlled list of journals ISSNs (for each exact matching with yearly IF to be used in other modules)
- Automatic building of citations from metadata

Conclusions

- More holistic view is required
 - (IR cannot be separated from authors researching/communicating/publishing habits, from authors and institutions web publishing, from institutional assessment and evaluation activities, etc)
- Authors/Researchers needs should be better addressed, enhanced web services and authoring tools should be implemented
- More integration and interoperability
 - Aggregation of different type digital content with datasets, reference codes of objects, experimental sources

Conclusions

- o More progress is required in document format, authors are still working with traditional formats (Word, TeX, html, PDF)
- o Newer document formats adoption (Microsoft's Open Office XML or the OASIS OpenDocument) is needed
 - o to extract semantic information, to enable semantic analysis to allow automated interlinking to external sources, machine to machine linking of items, at sub-item level

Thank you

Open Access and digital repositories

- Promotion and Advocacy
 - Training courses
 - SCOAP – Working Group
 - SPARC
 - Participation in conferences in Italy and abroad
 - Hosting conference in Italy

- Digital Repositories – Institutional Archives
 - Consultancy on technological issues
 - Hosting and maintenance of DSpace and E-prints instances in some consortial universities
 - Personalization and customization of these instances

- PLEIADI – Portal for the Italian Scientific Electronic Literature
 - Digital platform to access OA content deposited in Italian institutional and disciplinary archives in collaboration with CILEA