



Strathprints Institutional Repository

Robertson, R. John and Dawson, A. (2006) An easy option? OAI static repositories as a method of exposing publishers' metadata to the information environment. In: Proceedings of the 10th International Conference on Electronic Publishing, 2006-06-14 - 2006-06-16. ,

This version is available at <http://strathprints.strath.ac.uk/2373/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to Strathprints administrator: strathprints@strath.ac.uk

CENTRE FOR DIGITAL
LIBRARY RESEARCH



cdlr.strath.ac.uk

An easy option? OAI static repositories as a
method of exposing publishers' metadata to the
wider information environment
Elpub 2006

John Robertson and Alan Dawson
Centre for Digital Library Research
University of Strathclyde

Overview

- Introduction
- Participation in OAI-based services
- The technical overhead problem and OAI-SR solution
- Stargate's solution for publisher metadata
 - Architecture and workflow
- Metadata issues
- Results
- Conclusions and Issues

Introduction

- Stargate (Static Repository Gateway and Toolkit)
 - JISC funded short study as part of the the PALS Publisher Metadata and Interoperability 2 Programme
 - Aim: to investigate the utility of OAI static repositories to enable small publishers to participate in OAI-based services within the Information Environment and beyond
- Intended audience
 - Small journals or publishers
 - minimal technical support, but some form of structured information

Acknowledgements

- Centre for Digital Library Research
 - Stargate Technical Director: Alan Dawson
- Participating journals:
 - Journal of Digital Information
 - Information Research
 - Information Scotland
 - Library and Information Research
- Harvester
 - TechXtra service at Heriot Watt

Participation in OAI-based services

- Publishers should be interested in making their metadata as available as possible – increased dissemination leads to greater impact
- Rise of services that can incorporate OAI-PMH (e.g. OAIster, Connotea, Metalis, Google Scholar/Yahoo)
- Ensuring that the final published version and the ‘correct’ metadata are present in OAI services
 - academic process is dependent on this versions
- Preserving journal identity
- Value of OAI-PMH demonstrated by use by
 - Inderscience (in PALS 1 project)
 - Institute of Physics

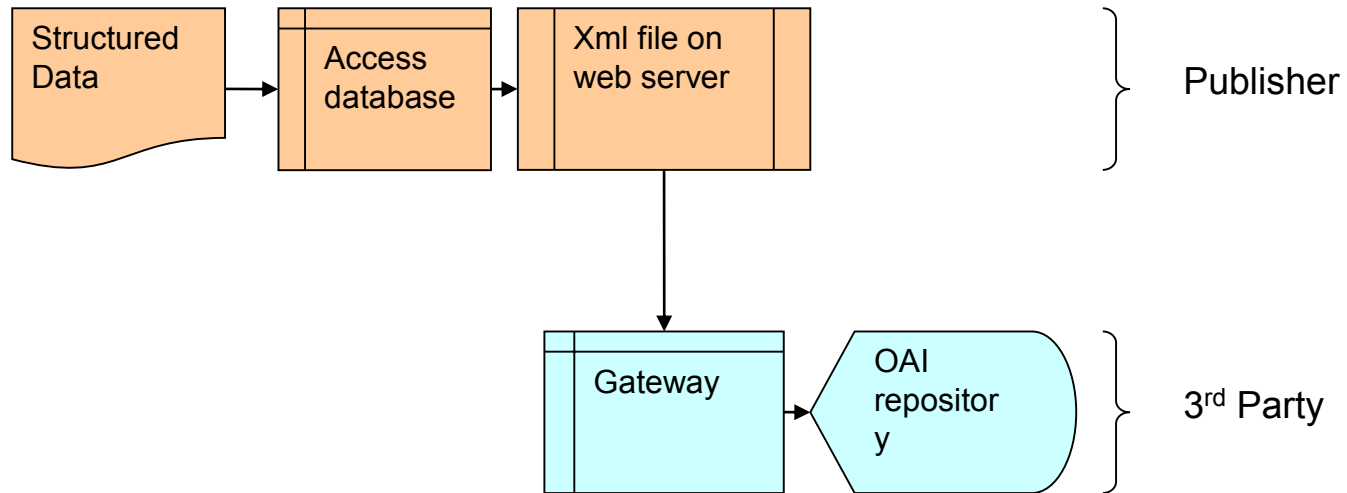
The technical overhead problem and OAI-SR solution

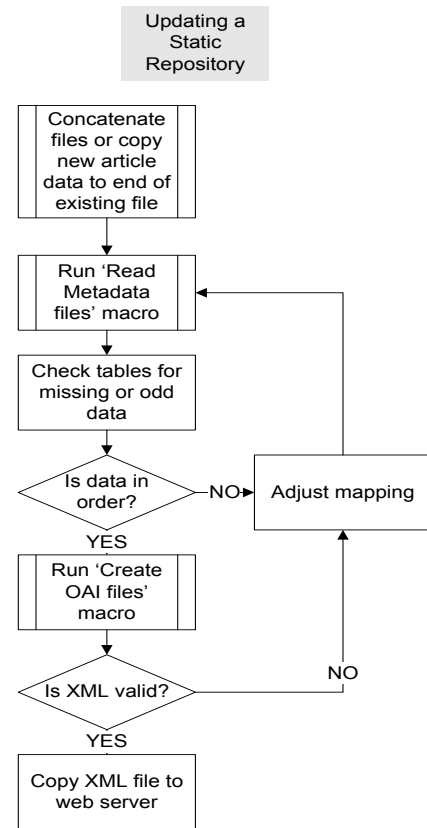
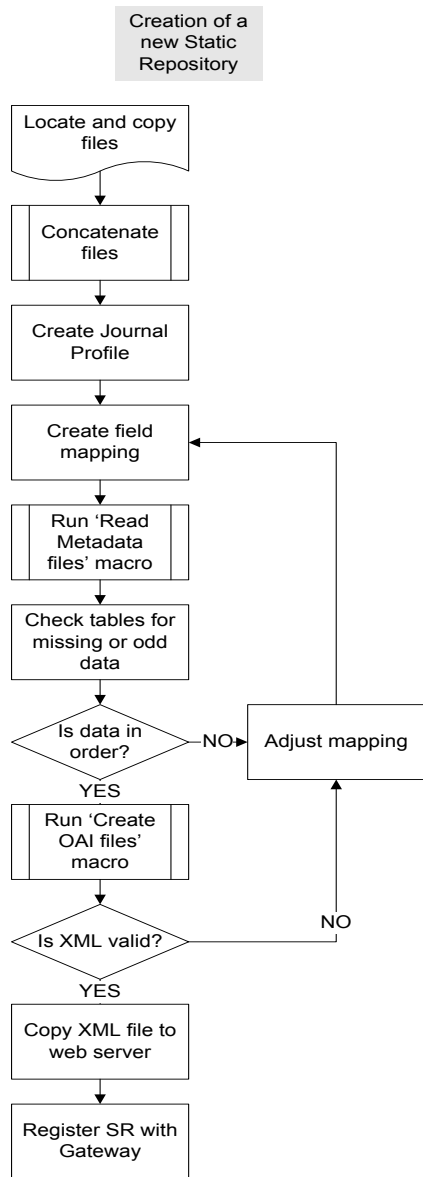
- For a lot of small publishers the required skills to create and manage an OAI repository may be too complex.
 - “knowledge of XSLT transformations, PHP, Java servelets and MySQL”
- A solution exists within the OAI Protocol in the Static Repository
 - In essence a text (xml) file on a webserver
 - The difference between SR and OAI repository is like the difference between a static versus a dynamic website

Stargate's solution for publisher metadata

- As simple as possible actually implies still having a database
 - A database is easier to adjust as users can edit a field not wade through a XSLT
- As simple as possible also implies that ubiquitous novice-friendly software should be used –MS Access and MS Word
- Some initial setup is required but if DC-like structured information exists the process is quick and significantly simpler than supporting a repository

Architecture





Metadata issues

- 1 unusual characters and XML
 - Foreign characters which are valid html cause XML problems. This is resolved through ‘find and replace’ during data import (managed by a table)
- 2 problems with export mappings
- 3 consistency over time / data entry errors
- 4 metadata enhancement
 - By collection level addition (e.g. journal name, publisher name)
 - By extraction (Customise script to extract volume and issue from url)

Results...

Query results in OAlster and TechXtra – provide journal specific retrieval without support for additional metadata

The screenshot shows a search results page from the URL <http://oaister.umdl.umich.edu/cgi/bib/bib-idx?type>. The page is divided into two main sections. The left section lists search results for various institutions, with several entries circled in red. The right section lists various search engines, also with several entries circled in red.

Search Results (Left):

- 51 records
- Information Research: 55 records
- Information Scotland: 129 records
- Institut Francais d'Etudes Andines (IFEA) Publication Series: 1 record
- Institute of Physics (IOP): 157 records
- The International Journal of Education and Development using Information and Communication Technology (IJEDICT): 3 records
- Iowa Publications Online: 61 records
- IR at RUG, Rijksuniversiteit Groningen: 26 records
- IU Bio-Archive: 22 records
- IUScholarWorks: DSpace at Indiana University: 2 records
- Jefferson Digital Commons: 7 records
- JeLit (Journal of eLiteracy): 9 records
- Johnson Technical Reports Server (JTRS), Johnson Space Center, NASA: 1 record
- Journal of Chemical Education (JCE) Digital Library: 320 records
- The Journal of Community Informatics: 1 record
- Journal of Digital Information (JoDI): 26 records
- Utrechtse Hogeschool Delft: 1 record

Search Engines (Right):

- Directory of Open Access Journals (DOAJ) [description website](#)
- DSpace at MIT [description website](#)
- EEVL EJournal Search Engine (EESE) [description website](#)
- EEVL - Best of the Web [description website](#)
- EEVL Website Search Engine [description website](#)
- ePrints UK [description website](#)
- Euclid Mathematics and Statistics Journals [description website](#)
- Geotechnical, Rock and Water Resources Library (GROW) [description website](#)
- Inderscience Journals [description website](#)
- Information Research [description website](#)
- Information Scotland [description website](#)
- IoP Electronic Journals [description website](#)
- Journal of Digital Information [description website](#)
- Library and Information Research [description website](#)
- NACA Technical Reports [description website](#)
- NASA Technical Reports [description website](#)
- Networked Digital Library of Theses and Dissertations (NDLTD) [description website](#)
- National Engineering Education Delivery System (NEEDS) [description website](#)
- OneStep Jobs [description website](#)
- OneStep Industry News [description website](#)
- Pearson Education - Books [description website](#)

Conclusions

- Findings
 - Static Repositories are an easy and effective, low tech option
 - Need for a dedicated Gateway service
 - Metadata diversity will still be an issue (elements available, no author format)
- Questions for service development
 - Journal targets: what impact would this have for users?
 - Open Access materials only or everything?
 - Possibility for dedicated service and community-specific metadata (parallel to OLAC)
- Other uses
 - OAI for a Virtual museum (<http://gdl.cdlr.strath.ac.uk/springburn/>)
 - Parallel methodology for EBook production (<http://gdl.cdlr.strath.ac.uk/scotlandebooks.html>)

Contact

Questions, comments

R. John Robertson

Centre for Digital Library Research

Department of Computer and Information Sciences

University of Strathclyde, Livingstone Tower

26 Richmond Street, Glasgow, G1 1XH, UK

Tel: +44 (0) 141 548 5854

Website: <http://cdlr.strath.ac.uk/stargate>



```
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />
[...]
```

Comunicación de conocimiento. ¿Habilidad de los profesores universitarios?

```
<meta name="dc.title" content="Comunicación de conocimiento.
¿Habilidad de los profesores universitarios?" />
<meta name="dc.creator" content="J. Licea de Arenas, J.V. Rodríguez,
J.A. Gómez, M. Arenas" />
<meta name="dc.subject" content="Comunicación de conocimiento.
¿Habilidad de los profesores universitarios?" />
<meta name="dc.description" content="Introducción: [...]" />
<meta name="dc.subject.keywords" content="bibliometría,
alfabetización en información, profesores universitarios,
Española" />
<meta name="dc.subject" content="Comunicación de conocimiento.
¿Habilidad de los profesores universitarios?" />
<meta name="robots" content="all" />
<meta name="dc.publisher" content="Professor T.D. Wilson" />
<meta name="dc.coverage.placename" content="global" />
<meta name="dc.type" content="text" />
<meta name="dc.identifier" scheme="ISSN" content="1368-1613" />
<meta name="dc.identifier" scheme="URI"
content="http://InformationR.net/ir/11-1/paper243.html" />
<meta name="dc.relation.IsPartOf" content="http://InformationR.net/ir/11-
1/infres111.html" />
<meta name="dc.format" content="text/html" />
<meta name="dc.language" content="es" />
<meta name="dc.rights" content="http://creativecommons.org/licenses/by-nd-
nc/1.0/" />
<meta name="dc.date.available" content="2005-10-15" />
```

```
<dc:title> Comunicación de conocimiento. ¿Habilidad de los profesores
universitarios? </dc:title>
<dc:creator> J. Licea de Arenas, J.V. Rodríguez, J.A. Gómez, M. Arenas
</dc:creator>
<dc:type>text</dc:type>
<dc:publisher>Professor T.D. Wilson, University of Sheffield</dc:publisher>
<dc:format>text/html</dc:format>
<dc:language>es</dc:language>
<dc:description> Introducción: [...]</dc:description>
```

PubId	PubStatus	PubTitle	Publisher	PubURL
infres	active	Information Research	Professor T.D. Wilson, University of Sheffield	http://infor mationr.n et/
	PubFormat	RecordDelim	FieldDelim	UpdateItemId
	html	newrecord	<meta name="	yes

Translations

PubID	OldText	NewText
infoscot	Description.abstract	description
infoscot	Subject.titleUniform	omit