

CENTRE FOR DIGITAL
LIBRARY RESEARCH



cdlr.strath.ac.uk

Online Catalogue and Repository Interoperability Study (OCRIS)

Appendix 3 to OCRIS Final Report

In-depth studies of a sample of IRs and Online Public Access Catalogues

Duncan Birrell, Gordon Dunsire and Kathleen Menzies



JISC

Centre for Digital Library Research, University of Strathclyde
September 2009

In-depth studies of a sample of Institutional Repositories and Online Public Access Catalogues¹

1 Introduction

As part of Workpackage 3, the OCRIS project team devised a workflow by which to investigate Online Public Access Catalogues (OPACs) and Institutional Repositories (IRs) within 10 Higher Education Institutions (HEIs) deemed (both individually and as a group) to be representative of a particular facet of the project and its context. This purposive sampling provided a snapshot of the current situation regarding duplication of both *records* and *scope* between bibliographic/publication systems, the way item types are described and curated, the implementation and customisation of the software being used and the extent to which links are being made between the two systems.

Having identified HEIs with both OPACs and IRs within a previous workpackage, a purposive sample of these (10 in total) was selected. Establishing which item types were in scope for the OPACs and IRs, each sub-set of institutional systems was interrogated at User Interface (UI) level in order to identify and retrieve records for which examples of overlap or duplication might be found (for example, a record for the same thesis, held by both the IR and the LMS). Using these records, many issues and problems stemming from duplication, arrangement and description were explored.

The 10 institutions which OCRIS selected for in-depth study were as follows:

- Cambridge and Glasgow acted as Case Study institutions within Workpackage 3 thus were selected for inclusion.
- Hull is unusual in using Fedora software as the basis of its IR (with a Muradora interface).
- Southampton is notable for having multiple Institutional Repositories (10 in total), many of which pertain to specific subjects or disciplines.
- The National Marine Biological Library was selected as an atypical example for various reasons - its IR is both multi-institutional and subject-specific while its catalogue is based on Open Source software. OCRIS was able to explore its aims and objectives beyond the boundaries of a single University.
- Strathclyde is the "home" library of the OCRIS team and was an early advocate of Open Access repositories – being able to talk directly to staff involved in its development was beneficial.
- The University College of the Creative Arts administrate the Visual Arts Data Service (VADS), a format and discipline specific repository, using custom software.
- The metadata provided for their bi-lingual interfaces was considered in relation to the 3 HEIs which contribute to the Welsh Repositories Network (WRN): the Universities of Aberystwyth, Cardiff and Glamorgan.

The following discussion is structured according to various aspects explored relating directly to the aims and objectives of OCRIS. These are: scoping issues (primarily, duplication of scope at both item and record level as well as information provided to users on scope); the description of item and format types; the arrangement and granularity of subject menus; the application of authority control and levels of ambiguity/clarity; the instantiation of links between IRs and OPACs.

These themes are not explored herein in relation to *every* institution; instead, those examples deemed most illustrative have been chosen. Similarly the discussions vary in depth and length due to factors such as the number of IRs or catalogues within an institution, the amount of duplication occurring or the availability of information concerning any given system.

¹ All information taken from websites was correct at the time when these studies were carried out (July/August 2009). Inevitably some information will have changed since that time.

2 Scopes and item types

2.1 Cambridge University

At Cambridge University, the items searchable in the **Newton** catalogue are given as:

- Book
- Serial
- Electronic journal
- Electronic resource
- Disc (CD/DVD)
- Music Score
- Map
- Non-musical Recording
- Musical Recording
- Archive/Manuscript
- Kit
- Mixed Material/Collection
- Mixed Material
- Visual Material

But naturally the true scope extends beyond this list of 14 MARC-derived fields and is also more complex. Further investigation of the website reveals that the following holdings of Cambridge University Library are recorded in Newton:

- all printed books published from 1978 onwards, with the exception of Official Publications
- selected Official Publications published since 1999
- printed books published before 1978 considered to be of academic importance at the time of acquisition
- all print journals
- all electronic journals
- atlases published after 1977
- maps catalogued since August 2000
- sheet music and recorded music catalogued after 1990
- microfilms and microfiches published after 1977
- audio-visual material published after 1977
- music manuscripts

All printed books and journals in the 4 CUL dependant libraries are also recorded.

An additional note on scope states that "Coverage of books published prior to 1978 in the University Library is incomplete. Books published prior to this date, and considered at the time of acquisition to have been of secondary academic importance, are not included in Newton...[]...Coverage of music manuscripts in this catalogue is complete, and where such manuscripts contain two or more pieces, there are in most cases catalogue entries for each piece. This catalogue does not include theses or non-music manuscripts, which are in a separate Newton catalogue, University Library Manuscripts and Theses."²

Yet it takes some digging for users to get down to this level of detail.

The item types for the CUL's **Pilot Universal Catalogue** are the same as those listed above, but it also allows users to set search limits by format (which it terms "medium"). The range of mediums recognised in the Universal Catalogue is:

- Digital

² http://ul-newton.lib.cam.ac.uk/vwebv/ui/en_US/htdocs/help/about.htm [Accessed 24th august, 2009].

- Globe
- Map
- Microform
- Nonprojected Graphic
- Motion Picture
- Projected Graphic
- Sound Recording
- Text (Eye-Readable)
- Videorecording

DSpace@Cambridge webpages inform users that it was established in 2003, "to facilitate the deposit of digital content of a scholarly or heritage nature, allowing academics and their departments at the University to share and preserve this content in a managed environment". It provides a list of the item types contained within, derived from the standard Dublin Core "dc.item" terms instantiated within the repository. The following are listed in the 'browse by type' menu:

- Article
- Audio
- Book or Book chapter
- BW Image
- Colour Image
- Dataset
- Drawn image
- Image
- Journal Article - Published Version
- Journal Article - Submitted Version
- Map
- Other
- Preprint
- Presentation
- Software
- Table
- Technical Report
- Thesis

Similarly to the OPAC, the scope is in actuality broader than this, with all types of material being sought.

There is then, more potential overlap between OPAC and IR than it may seem from looking at "browse by type" menus. Conversely, overlap may be less likely than it appears once a user unearths sufficient detail. Yet it can be concluded that books, journals, maps and theses are all within scope for both systems at Cambridge. The ways in which that duplication actually occurs is considered in later sections.

2.2 University College for the Creative Arts

Types and Locations listed in "advanced search" delimiters by the **UCCA Library Catalogue** are as follows³:

FORMAT:

Academic research output
Cameras and IT equipment
CD-ROMs and computer files

³ References to administrative aspects such as "reported missing", "in transit between campuses" and "CD-ROM 3 week loan" for reasons of relevance.

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

Copyright cleared photocopies
Dissertations and theses
DVDs, videos and digital games
Journals and newspapers
Manuscripts
Maps
Monographs i.e. books, etc.
Music and sound recordings
Slides

ITEM TYPE:

AV equipment
Academic research output
Audio cassette
BFI ticket - three day loan
Box of slides
British Library article
British Library book
CD
CD-ROM
Canterbury Fine Art Store
Canterbury Media Store
Computer equipment
Copyright cleared photocopy
Digital game
Electronic book
Electronic journal
Epsom Media Store
Information files
Journal
Laptop
MA reading list
Non BL document supply article
Non BL document supply book
Reference book
Scanned material (shadow catalogue)
Slide or group of slides
Thesis

LOCATION:

Academic Reserve Collection
Animated Films Video Collection
Annuals Collection
Archive
Artists' Book Collection
Audio Collection
Berbiers Collection
Canterbury Fine Art Store
Canterbury Media Store
Careers Collection
DVD Collection
Documentaries Video Collection

Feature Films Video Collection
Fiction Collection
Herbert Read Collection
Information Files
Journals Collection
Journals Collection (shadowed)
Look Book Collection
Map Collection
Multimedia Office
New Books Display
Off-air Recordings Collection
Online Collection
Oversize Collection
Professional Collection (in Library Office)
Rare Book Collection
Reference Collection

The OPAC does not appear to include records for single visual/digital images. Formats

The **Visual Arts Data Service (VADS)** is based at the Farnham Campus of the UCCA thus is treated herein as the UCCA's "Institutional Repository." However, VADS is a national service - containing over 100,000 records from many institutions for use in learning, teaching and research - and is aimed at the *entire* academic community.

The VADS "Collections" search offers access to 44 separate image collections, where the material types are marked "images". It also provides guidance and consultancy services in relation to digital projects and preservation.

From the catalogue lists given above it appears that the only *potential* scope overlap between OPAC and IR would be:

- Maps
- Slides
- Photocopies
- Scanned material

Even then this would only apply if they were derived from the two UCCA collections recorded in VADS (the Crafts Study Centre: University for the Creative Arts at Farnham and the Textiles Collection: University for the Creative Arts at Farnham) or if they linked to external digital copies of images.

In the OPAC these would always remain at bibliographic level only while in VADS images would be stored alongside richer, VRA4 bibliographic records.

In any case, there appears to be no duplication as slides/maps/photocopies and scans are not associated with the 2 UCCA VADS collections (whose parent institutions are not listed in the "Locations" menu). These specialist Centres are not libraries and the 30,000 items of the Crafts Study Centre Archive⁴ are described in the JISC-funded *Archives Hub* gateway service (<http://www.archiveshub.ac.uk>) rather than within the UCCA catalogues.

There is also potential overlap with some of the sets of "Resources" available via VADS (<http://www.vads.ac.uk/resources/index.html>) which include a database of selected film titles and viewing notes, transcripts of letters and multimedia collections. But generally these are clearly packaged to act as learning/teaching resources (a variety of "modules" are contained in the resources section) and are quite distinct in purpose to the bibliographic records held in the OPAC, effectively

⁴ <http://www.csc.ucreative.ac.uk/index.cfm?articleid=20211> [Accessed 25th August, 2009].

being parts of a tailored collections. The one exception here may be the electronic copies of the journal *Design*.

VADS is a rich data source and merits further exploration in terms of item description/arrangement. Its search by "Themes" facility allows users to locate items within and across five key subject areas in the arts:

- Applied Arts
- Architecture
- Design
- Fine Art
- Media

The image files within these 5 themes are sub-divided into 20 classifications, with 70 listed material types or techniques. This means the material types depicted in VADS image files can only be discerned by clicking through each of the themes in turn. Whilst effectively placing the onus for building an index of material types on the end user, this type of classification is clearly designed to reflect the ways in which students think about and make use of visual arts materials (and illustrative images of them) within their work. Similarly, items are arranged by Collection (44 in total, 2 of which are UCCA collections). Students can save records to their own "lightbox" area and add tags to items.

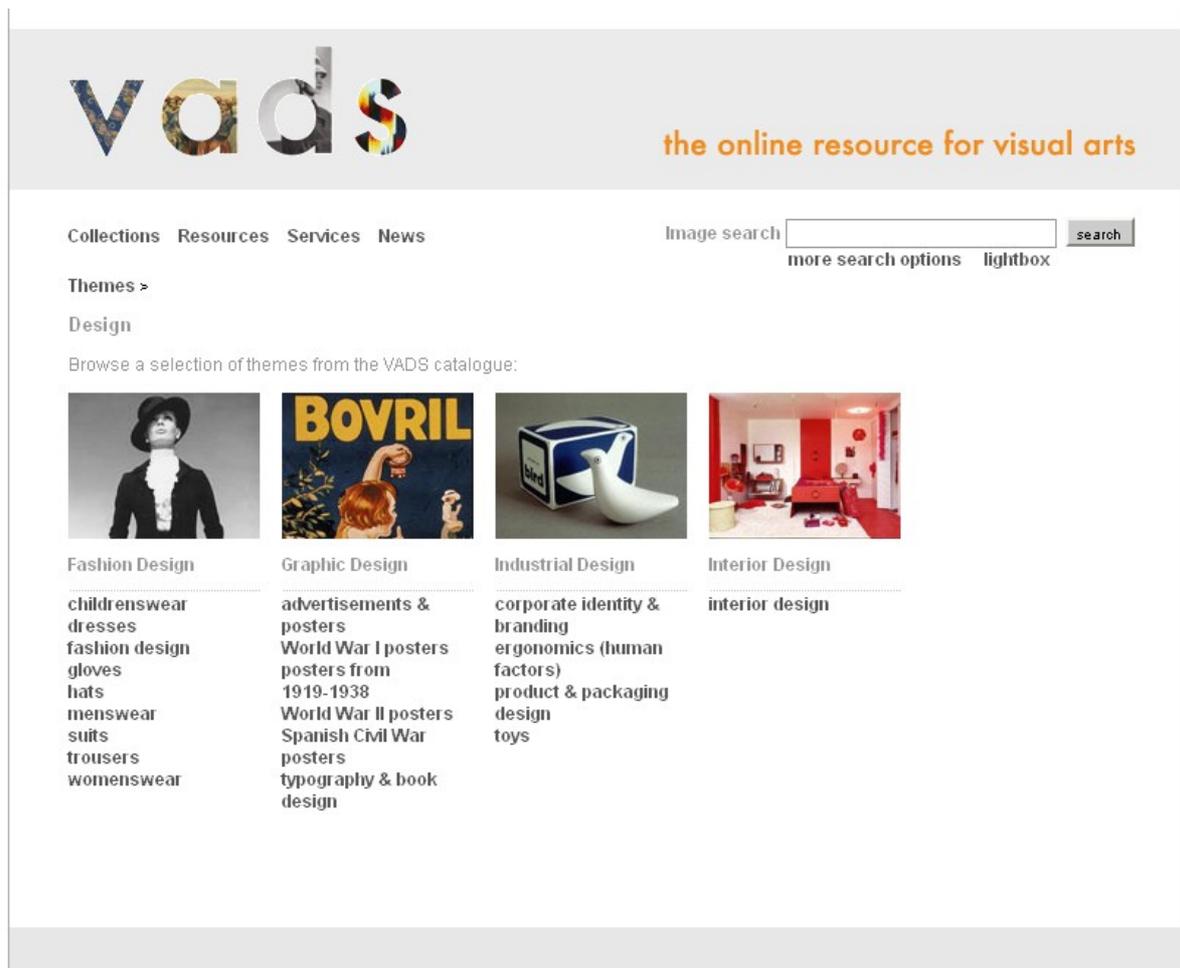


Figure 1 - Just one of the specialised search menus tailored around media type, subject headings and 'themes' within the VADS 'browse by theme' interface. Only upon clicking through to the relevant subject "theme" does it offer the facility to limit searches or group items by material type or scope.

VADS developers are well aware of the need to interoperate beyond institutional boundaries and are currently "developing mechanisms to automatically share data with the growing Institutional Repository network and other data aggregators whose users require VADS images for their teaching and research" as part of the JISC Digitisation and e-content Programme⁵.

2.3 University of Glasgow

At Glasgow, the **Encore Resource Discovery Platform (RDP)**, known as the "Quicksearch Service" currently searches the same database as does the main OPAC (or "Classic Catalogue") but uses different relevancy rankings and presents results in a more "Google-generation friendly" way, employing faceted menus, auto-suggest and tag clouds, for example.

It is notable that lists clarifying scope differ in terminology and number of item types listed. The main page of the **OPAC** lists 11. These are:

- Books
- Journals
- Audio-visual materials
- Course materials
- Electronic Resources
- Exam papers
- Maps
- Music Scores
- Rare Books
- Sound Recordings
- Theses

Those listed within the "advanced search options" of the OPAC and on the Resource Discovery Platform interface are:

- PRINT BOOK
- E-BOOK
- PRINT JOURNAL
- E-JOURNAL
- GU EXAM PAPER
- GU THESIS
- MUSIC SCORE
- MAP/ATLAS
- VIDEO
- SOUND
- EXTRACT
- E-EXTRACT
- KITS
- FINDING AID
- ARTWORK
- PHOTO SLIDE(S)
- NOT CODED
- DATABASE
- SOFTWARE
- WEBSITE
- E-SCORE
- E-MAP

⁵ <http://www.jisc.ac.uk/whatwedo/programmes/digitisation/enrichingdigi/vads.aspx> [Accessed 25th August, 2009].

- E-THESIS
- E-VIDEO
- E-SOUND
- DATABASE (CD)

This discrepancy is because of the level of granularity being applied, with "Electronic Resources" and "Audio Visual Materials" broken down into several further subtypes and "Books" becoming "Print Book" and "E-Book".

Enlighten (<http://www.lib.gla.ac.uk/enlighten/>) is the University of Glasgow's Institutional Repository service for published research material including:

- peer-reviewed journal articles
- published conference papers
- books and book chapters

The full list of item types only appears when you select Advanced Search, revealing them to be:

- Article
- Book Section
- Conference Proceedings
- Book
- Patent
- Artefact
- Show/Exhibition
- Composition
- Performance

Although these are probably derived from default EPrints "dc.item" types lists, it is worth observing that those listed on the homepage of Enlighten are different, as this makes a point about in/consistent description. Published conference papers appear to have been conflated with Conference Proceedings (or are they treated as "articles"?), while "book chapters" has become "Book Section".

Nevertheless, the scopes of the 3 Glasgow IRs are well defined and clear. Enlighten acts as a comprehensive publications database with as much content being made freely available as possible (and in such a way that as much of that data can be re-used in as many different places as possible). The **Glasgow Theses Service** has a self-evidently narrow scope, whilst the **Glasgow DSpace Service** (intended to make "grey" and other non-peer reviewed outputs visible) is gradually being sidelined in favour of Enlighten.

However, remnants of the IRs' histories as part of the "Data Providers for Academic E-content and the Disclosure of Assets for Learning, Understanding and Scholarship" (DAEDALUS) project (when scoping and other issues were less certain)⁶ can still be glimpsed in some aspects of Enlighten's User Interface (UI) as it continues its journey towards becoming a central publications system.

For users accessing the homepage, it may appear that the **Glasgow Theses Service** functions as a subset or specialised partition of Enlighten (Theses being referred to in its search options menu as visible in Figure 1) or that the three IRs are inter-linked. However, although there are web links provided on the Enlighten interface, allowing you to navigate and select easily between the additional Theses and DSpace Services - and conveying a seamless "one stop shop" - the underlying services are technically quite separate.

⁶ See: <http://eprints.gla.ac.uk/115/> [Accessed 26th August 2009].

home > services > Enlighten

Enlighten

- Enlighten Home
- About Enlighten
- Search**
- Simple
- Advanced
- Theses & Other Material
- Our Top 100 Searches
- Browse Papers by**
- Year
- Subject
- Journal
- Author
- Faculty/Department

Search Enlighten, Theses and Other Research Material

It is possible to search all of the material in Enlighten using this local Google search and other research material.

Google™ Custom Search

What are the most popular searches?

- [Top 100 search terms and phrases from Google and Google Scholar](#)

Search for

These pages will let you search our individual repositories:

- [Research publications including journal articles and books \(Enlighten\)](#)
- [Theses \(Glasgow Theses Service\)](#)
- [Other Research Material \(Glasgow DSpace Service\)](#)

Figure 2 - Enlighten links to the other Glasgow IRs in various ways.

2.4 University of Hull

Hull is unusual because it uses Fedora Commons Repository Software for its IR. Known as **eDocs**, it is being customised and developed by the University's e-Services Integration Group (working in close co-operation with other members of Academic Services) and uses Muradora⁷ as its web-based "front end".

eDocs, it is indicated, can manage "any type" of digital material generated through research, education and administration, stating that "Materials currently being managed through the repository include images, project reports, undergraduate dissertations, teaching materials, datasets, collections of past exam papers and doctoral theses." There are plans for future development which would expand its range of content: "possible content for the future include published articles and conference papers, book chapters, and other forms of written research output".

However, the scope is somewhat unclear. While eDocs currently manages some of the research and teaching materials of academic staff and the outputs of postgraduate students it mainly focuses at present on administrative material, the repository being used to explore issues relating to institutional records management and storage, personal metadata, workflows and the use of web services.⁸

Underpinning this, it seeks to support the work of both the teaching and learning communities as well as the University's administrators. The areas of potential overlap between the IR and OPAC are many.

The OPAC (<http://library.hull.ac.uk/>) does not provide a statement on scope, but by selecting advanced keyword search users can search by "Material Type". These types are as follows:

- ARTEFACT
- BOOK
- BOOK WITH DISK
- CASSETTE

⁷ <http://www.muradora.org/muradora> [Accessed 25th August, 2009].

⁸ See: <http://www.jisc.ac.uk/whatwedo/programmes/preservation/remap.aspx> [Accessed 26th August, 2009].

- CD AUDIO
- CHARTS
- COMPUTER FILE
- DVD
- EBOOKS
- EJOURNALS
- KIT
- MICROFORM
- MS MAP
- PERIODICAL
- PLAYBILL
- PRINT MAP
- PRINTED MUSIC
- SLIDE
- SOUND RECORD
- SPOKEN RECORD
- TAPE WITH BOOK
- THESIS
- VIDEO
- XEROX

A list of the "Material types" supported by the OPAC at Hull

Despite the planned integration of the IR with other institutional data services, the only duplication as yet appears to be of "Theses". However, this depends on what is classified as "Teaching Material", which is somewhat unclear.

Usefully, the OPAC provides links to electronic copies of theses held within eDocs, as can be seen below:

Author	Mung'au, Franklin.
Title	Channel estimation for gigabit multi-user MIMO-OFDM systems.
Publication Info	2008.

Location/Availability	Find Similar Items
-----------------------	---------------------------

Click on the following to:
[Go to electronic thesis](#)

LOCATION	SHELVED AT	STATUS
BJL Short Loan Theses	T/H 2008 Ph.D. M9	AVAILABLE

Figure 3 - The "Go to electronic thesis" hyperlink takes you to the copy held in the Hull eDocs repository.

2.5 National Marine Biological Library

LibWeb is used by the **National Marine Biological Association** (a UK Learned Society promoting scientific research into marine science) as the basis for its OPAC.

Its homepage states that: " LibWeb, the NMBL's online catalogue includes references to selected journal articles, books, book chapters, reports, conference papers, CDs, etc. It contains all items catalogued since 1988, the entire book collection, and a growing number of records from the old card catalogue, with an emphasis on the historical literature."

Only by contacting staff at the NMBL can users readily discover that "the selection of journal articles, book chapters, etc. (for both the old card catalogues and LibWeb) is based on the research interests of scientific staff between 1888 and about 2000." Further, "Since then only books, monographs, staff publications, and any articles relevant to British waters or marine pollution have been catalogued and indexed."⁹

However, a useful collection overview is provided:

"The NMBL holds:

- 13,527 books
- 80,000 reprints
- 147 expedition report series
- 1,200 conference proceedings
- 4,668 serial titles ([click here to search serials catalogue](#))
- 1,052 current serials
- [Special collections](#) which include: Browne, Lebour, Harmer, Bidder and Young Libraries and Lankester Pamphlets
- [MBA archive](#) including papers, letters, notebooks and other documents, illustrations, photographs and pictures
- Comprehensive collection of maps, charts, atlases, expedition reports, microfilms/microfiches and CD-ROMs

Description of collection:

- Extensive coverage of marine biological and ecological literature
- Long runs of periodicals
- Wealth of historical literature
- Extensive reference materials
- Reports and grey literature from all parts of the world
- Worldwide coverage of research literature
- Books classified by subject, including finely illustrated, rare and valuable books from the 19th century and earlier
- Periodicals arranged in subject groups, subdivided geographically
- Rare material e.g. pre-1914 publications from Japan and the Far East, Russia, the Ukraine and other Eastern European countries
- Expedition report series dealing with deep-sea, offshore, coastal and terrestrial environments, many now difficult to locate in libraries."

The OPAC offers access then to a major research collection with much archival material. However, it is hard to get an overview of collection strengths or item types helpful to the formulation of a search strategy; this is partly because the Open Source software on which LibWeb is based is not simple or particularly user friendly. For example, a user cannot search for *phrases* ("climate change" would not find anything but "climate*change" retrieves 698 records) and the interface is quite unlike the interfaces of most OPACs or resource discovery websites.

If wishing to limit a search by year, or item type, a user must type certain commands into the "search terms" box - for example, "0909" will limit the search to items added to the catalogue from September 2009 onwards while "book" will search only for books. But which other terms might be appropriate to limit searches by item type? A list of item types is not provided thus users may have to guess whether "expedition report" or "technical paper" for example would be recognised terms.

Although the LibWeb help pages contain detailed and extensive guidance to assist users in the formulation of search queries (http://www.mba.ac.uk/nmb/llib_help.htm) it is undeniable that

⁹ Personal correspondence via email, July 2009.

discerning finer details of scope and searching the catalogue accordingly is - at least, for a novice - time-consuming.

The **Marine & Ocean Science EPrints Archive @ Plymouth** administered by the NMBL is "a digital archive providing access to papers produced by the staff of the Marine Biological Association of the United Kingdom, Plymouth Marine Laboratory and the Sir Alister Hardy Foundation for Ocean Science." It provides services to, and is shared by, all three organisations. Further, it:

"Is also an historical archive containing digital copies of early papers from the *Journal of the Marine Biological Association of the United Kingdom*." This makes its scope atypical.

The item types that can be searched are listed as:

- Article
- Book Section
- Monograph
- Conference or Workshop Item
- Book
- Thesis
- Patent
- Other
- Multimedia
- Maps

Clearly some of these correspond to the item types referred to in the OPAC, partly because of the emphasis within both of collecting historical/archival/Special Collections items and records. The "Browse by year" facility goes back as far as 1887.

As well as such items, current research imperatives are very much recognised by the repository. Details of any research output resulting from Natural Environment Research Council (NERC) funding must be added each year to the externally administered NERC Research Output Database (ROD). The OCRIS team were informed¹⁰ that extra fields have therefore been added to the EPrints repository for collecting information required for this annual exercise (e.g. publication type (ISI journal, non-refereed journal, etc), type of co-author (overseas institution, UK private sector, etc)). The hope was that, eventually, a programme would be developed so that records from the IR could be migrated directly to ROD.

Yet there are actually very few recent records in the IR. This, staff state, is due to a new version of the EPrints software having become available, offering far greater functionality than the version that is currently in use at NMBL. "The aim is to update the system before encouraging research staff to routinely deposit their records. Unfortunately, due to lack of IT support, the upgrade has been pending for some time now"¹¹.

2.6 University of Southampton

The OPAC at Southampton lists the following formats as being in scope:

- Bibliographies
- British Official Publications
- Exhibition catalogues
- Medical projects
- Scores
- Statistics
- Theses
- Periodicals

¹⁰ Personal correspondence via email, July 2009.

¹¹ As above.

This appears limited and rather mixed in its degree of specificity. For example, "Medical projects" is quite specific, certainly at discipline level (although, what type of project is left unstated) while "Statistics" might relate to any subject. It neglects entirely to mention books or book items, which are also recorded. The "Mediums" listed are more extensive and are listed in full below:

- Audio Visual
- CD (music)
- CD Rom
- Cine Film
- DVD disk
- Database
- Diskette
- E-book
- E-journal
- Electronic Academic Reserve item
- Map
- Micro card
- Microfiche
- Microfilm
- Microprint
- Multimedia
- OHP
- Poster
- Print
- Print plus diskette
- Print plus tape
- Slide
- Software
- Sound Disc
- Sound Tape
- Tapeslide
- Video

The "Items" and "Mediums" lists may be pre-loaded lists which came with the LMS software; alternatively staff may be attempting to reflect the range of specialist types of material available via their OPAC - but which is the case cannot be readily discerned.

The numerous **Institutional Repositories at Southampton** are as follows:

1. Ted Nelson e-Prints repository
2. 6Journal
3. Serpent
4. Portfolio
5. Cogprints
6. Language Box
7. e-Prints Soton, the University of Southampton's Research Repository
8. Edshare
9. ECS EPrints
10. eCrystals

It was only feasible for OCRIS to select a subset of these for closer investigation. Those selected were: **e-Prints Soton**, **EdShare**, **eCrystals** and **ECS EPrints**. Summaries of the scopes of all 10 IRs and explanations of how and why a sub-set was selected are however provided below.

1. The **Ted Nelson EPrints repository** has not yet been set up fully (appearing perhaps to be a test installation). It contains only 31 items at present - some of which have the item type "thing" assigned to them. Thus, it was not worth exploring further.
2. **6Journal** was established as part of an European Union Sixth Framework Programme (EU FP6) funded project and distributes information about Internet Protocol version 6 (IPv6). It is extremely unlikely that this material would be in scope for description by either the OPAC or any other Southampton IR.
3. This is also the case with **SERPENT (Scientific and Environmental ROV Partnership using Existing Industrial Technology)**, a media archive holding all images and video footage associated with a specific deep sea research project "exploring [the] undersea environment" to make "cutting-edge industrial ROV technology and data more accessible to the world's science community" while working in collaboration with key players in the oil and gas industry.

SERPENT has an extremely detailed subject classification scheme in place, reflecting, among other things, the biological classification and taxonomies of the sea creatures being examined. The format types held in the repository are: "Image", "Sample", and "Video". While some project outputs might be describable in the OPAC an initial search suggests that there is no duplication occurring between the systems. It is most likely that outputs are only deposited within SERPENT, rather than being itemised as bibliographic records with the library.

4. **Portfolio** states: "this repository is being developed as a portfolio for students to exhibit their work." Again, it has well defined scope, with the following item types listed:
 - GDP Report (3)
 - IRP Report (10)
 - MSc Dissertation
 - Other (7)
 - Part III Project Report (11)
 - Software (5)

Clearly content is scarce at present and - as a showcase for student work - it can be assumed that none of these items will be recorded in the OPAC, as an initial search appears to confirm.

5. **Cogprints** welcomes users with the following detailed description:

"Welcome to Cogprints, an electronic archive for self-archive papers in any area of Psychology, Neuroscience, and Linguistics, and many areas of Computer Science (e.g., artificial intelligence, robotics, vision, learning, speech, neural networks), Philosophy (e.g., mind, language, knowledge, science, logic), Biology (e.g., ethnology, behavioural ecology, socio-biology, behaviour genetics, evolutionary theory), Medicine (e.g., Psychiatry, Neurology, human genetics, Imaging), Anthropology (e.g., primatology, cognitive ethnology, archaeology, palaeontology), as well as any other portions of the physical, social and mathematical sciences that are pertinent to the study of cognition."

This clearly scoped yet highly inter-disciplinary repository was investigated further by the OCRIS team.

6. **The Language Box** was developed as part of the EdShare initiative by Southampton's Faroes project (<http://www.faroes.ecs.soton.ac.uk/>) and is based on EPrints. It represents "an attempt to reinvent learning object repositories for the Web 2.0 world" and to extend the scope of EPrints installations beyond research into teaching and learning. Further, it is positioned as:

"A place where students and teachers of languages can publish and share their learning materials, resources and links on the web. [They] can use the resources directly, or create new activities to put your own twist on things. The Language Box is funded by JISC and designed, built and run by the

Faroës project team at the University of Southampton and the University of Portsmouth" and "is for anyone who teaches languages and who wants somewhere to host and manage their materials, it's also for language students that want to record and share good stuff that they've found."

It is stated on its website that those developing the "Learning Box" (as part of the EdShare initiative) have "re-examined old assumptions about learning objects and learning metadata" to enhance development of subject based "community repositories".

Item types given are "Activity" and "Resource". Content associated with these item types are quite specific and unusual (compared with the scopes of other institutional systems) - for example, the following items are stored within The Language Box:

- Character matching game (Mandarin Chinese)
- Daily routine by two native speakers of French
- Deutsche Zungenbrecher
- Images of Salzburg (taken by Dr. David Millard on a summer trip in 2005).

These are not likely to be catalogued in an OPAC or any other Southampton IR.

7. **e-Prints Soton** "is the University of Southampton's Research Repository and contains a growing collection of literature from the University." Further, "the service provides a simple mechanism for enabling researchers to deposit the full text of their work (self-archiving) or to give it to the central service to deposit for them (assisted archiving). It offers advice and guidance on issues such as metadata, formats and copyright. It will aim to ensure the quality of the metadata being added. It provides a complementary service to traditional catalogues and learning resources."

ePrints Soton be browsed by "Year" (going back to 1920) or by "Library of Congress Subject Area" (42291).

These Subject Areas come bundled with EPrints. Because e-Prints Soton focuses on acting as a repository for all types of research outputs, there is clearly potential overlap with some of the other Southampton IRs, regardless that their scopes are far narrower and more tightly defined (whether in terms of purpose or type of departmental output sought).

8. **Edshare** states that it is "A resource for collaboration and sharing of materials used in teaching and learning". There is no note on Scope. An Advanced Search lets you select format but doesn't specify item type. It has a variety of 'Web 2' features such as controlled tags, tag clouds, details of the person who uploaded, semantic mark-up and a comments facility, as illustrated in Figure 3 below:

The screenshot shows the EdShare repository interface. At the top left is the EdShare logo (University of Southampton) and a 'Login' button. A search bar is located at the top right. The main content area is titled 'Exercise sheet 2 pdf' and contains a list of PDF documents under the heading 'Exercises and solutions in PDF'. The documents listed are:

- Normal distribution and Brownian motion**: Format: PDF, Size: 17Kb. Buttons: Download, Request a copy, Link to...
- independent Brownian motions**: Format: PDF, Size: 24Kb. Buttons: Download, Request a copy, Link to...
- normally distributed risky assets and their mean and variances**: Format: PDF, Size: 36Kb. Buttons: Download, Request a copy, Link to...
- correlated and uncorrelated assets**: Format: PDF, Size: 20Kb. Buttons: Download, Request a copy, Link to...

On the right side, there is a 'Tags' section with a list of keywords: *Exercise, independent Brownian motions, normally distributed risky assets mean variances, correlated assets, Normal distribution Brownian motion, Financial Mathematics, mathbank, Normal Distribution*. Below this is the 'University Structure' field, which is 'School of Mathematics'. Further down is the 'About this share' section, which includes:

- Link**: <http://www.edshare.soton.ac.uk/2284/>
- Embed**: `<div class="embed_box"><te`
- Shared with**: Selected School/Unit
- Added by**: [Mathbank](#)
- Added on**: 22 Apr 2009 13:20
- Last modified**: 09 Oct 2009 03:04

On the left side, there are three sections: 'User Tools' (Create a share, Create a collection, Bulk upload, EdShare Organiser (beta), Latest additions, Tag cloud), 'Share tools' (Email this share..., Report a problem), and 'Send us your feedback!' (Fill-in our online survey, Email us).

Figure 4 – a typical 'record' for a "share" items within the EdShare repository.

9. **ECS EPrints** is the IR for the School of Electronics and Computer Science at Southampton. In 2001, ECS was the first academic institution in the world to mandate self-archiving of research output. It is fitting then that the *complete* research output of all members of the School is available and fully searchable in their online repository. It can be browsed by Year, Research Group and searches can be limited by Creator, Editor, Keywords, or Full text status. It is the "Advanced Search" facility which offers a break down of the Item Types held by the IR:

Item Type:

- Article
- Conference or Workshop Item
- Book
- Book Section
- Technical Report
- Thesis
- Other

Monograph Type:

- Technical Report
- Project Report
- Documentation
- Manual
- Working Paper
- Discussion Paper
- Other

Presentation Type:

- Paper
- Oral Presentation
- Poster
- Other

It is interesting to note that ECS EPrints has a 'Modify RAE Information' tool although this is password protected. It would be interesting to discover what links exist between the ECS IR and research systems at Southampton, or how it may make use of bibliographic tools to assist in the preparation and gathering of research outputs.

10. **eCrystals** has a very clear scope, being the archive for Crystal Structures generated by the Southampton Chemical Crystallography Group and the EPSRC UK National Crystallography Service.

As the homepage of eCrystals makes clear, "The information contained within each entry of this archive is all the fundamental and derived data resulting from a single crystal X-ray structure determination, but excluding the raw images."

Accordingly, the repository supports a variety of file formats not commonly used outwith certain disciplines or fields of study - for example, JMOL allows the structures of the crystals to be modelled and the resulting models can "rotate" on screen. "Data collection parameters" and "refinement results" tables provide fields for highly specific, numeric metadata associated with a given crystal. Authors are described as "sample originators".

The "About" pages of the website illustrate the types of specific and sophisticated need which a local (as opposed to centralised), format or discipline-specific repositories can serve, and the particular arrangement and formatting of records (and the IR itself) which attendant user communities may find best fitted to their needs:

"The information contained within an entry in this archive is all the underlying data generated during the course of a structure determination from a single crystal x-ray diffraction experiment. This information is freely available and allows the reader to a) assess the validity of the dataset or b) repeat the experiment or c) use the data for further studies.

An individual entry consists of three parts:

1. Core bibliographic data, such as authors, affiliation and a number of chemical identifiers,
2. Data collection parameters that allow the reader to assess at a glance certain aspects of the crystallographic dataset,
3. Files available for download. These files are: visualisations of the raw data (.jpg), the raw data itself (.hkl), experimental conditions(.htm), outputs from stages of the structure determination (_xs.lst, _xl.lst & .res), the final structural result (.cif & .cml) and the validation report of the derived structure (_checkcif.htm)."

The repository uses Dublin Core, qualified Dublin Core to incorporate additional chemical information. It uses the eBank Schema for the exchange of crystallography metadata (an application profile) and the locally developed "eCrystals Federation Data Deposit Model" (based on a model proposed by UKOLN).¹²

Staff are currently involved in the JISC-funded "eCrystals Federation" which "will make a major contribution to the development of a repository-based e-infrastructure for research data by establishing a solid foundation of crystallography data repositories across an international group of partner sites, with metadata harvested by a number of aggregator services. The issues of

¹² Lyon, Dr. L., and Coles, Dr. S. (2008). *eCrystals Federation: Open Repositories for global Open Science*. Presentation given at the DRIVER Summit, Gottingen, January 2008. Online at: <http://www.driver-repository.eu/Summit1/Lyon%20driver-summit-jan08.ppt> [Accessed 5th October 2009].

interoperability, repository deployment and preservation in a federation of data repositories are to be addressed through the testbed that this project will develop."

A typical eCrystals record is shown below:

eCrystals UNIVERSITY OF Southampton

Home | About | Browse by Year | Browse by People | Search

Login | Create Account

trans-3-(Tetra-O-acetyl-D-arabino-tetritol-1'-yl)-2-(2-phenyl-4-benzyl-1,3-dioxo-2,4-diazapentyl)-2-phenylthiirane

Sample Originator: Mark E. Light, J. L. Jimenez, Michael B. Hursthouse, J. C. Palacios, R. Gordillo, F. R. Clemente, P. Cintas, R. Babiano and M. Avalos.

$C_{36}H_{38}N_2O_{10}S$

InChI=1/C36H38N2O10S/c1-23(39)45-22-30(46-24(2)40)31(47-25(3)41)32(48-26(4)42)33-36(49-33,28-17-11-7-12-18-28)34(43)38(29-19-13-8-14-20-29)35(44)37(5)21-27-15-9-6-10-16-27/h6-20,30-33H,21-22H2,1-5H3/t30-,31-,32+,33+,36+wm1/s1

Identification Number: 10.3737/ecrystals.chem.soton.ac.uk/131
 Date Created: 20 April 2001
 Deposited On: 21 Jan 2008 15:29
 Deposited By: A.N. Admin

Depositor Comments
 This manuscript describes a one-pot, selective synthesis of optically active thiranes by [3+2] cycloaddition of mesoionic dipoles with sugar aldehydes. Overall, the result is the formation of monosaccharide mimetics that display numerous functional groups for molecular recognition.

Data collection parameters

Chemical formula	C36 H38 N2 O10 S
Crystal morphology	
Crystal system	Monoclinic
Space group symbol	P2(1)
Cell length a	7.8231(2)
Cell length b	12.9445(3)
Cell length c	17.3686(4)
Cell angle alpha	90.00
Cell angle beta	95.6680(10)
Cell angle gamma	90.00
Data collection temperature	180(2)

Refinement results

Solution figure of merit	0.0947
R Factor (Obs)	0.0418
R Factor (All)	0.0475
Weighted R Factor (Obs)	0.1006
Weighted R Factor (All)	0.1037

Citation: Light, Mark E. and Jimenez, J. L. and Hursthouse, Michael B. and Palacios, J. C. and Gordillo, R. and Clemente, F. R. and Cintas, P. and Babiano, R. and Avalos, M. (2001) University of Southampton, Crystal Structure Report Archive.
 (doi:10.3737/ecrystals.chem.soton.ac.uk/131)
 Export as: EndNote BibTeX ASCII Citation

Available Files

Final Result

01esp001_CIF	19k
01esp001.cml	11k

Validation

01esp001_checkcif.htm	9k
---------------------------------------	----

Refinement

01esp001.res	12k
01esp001_xl.lst	58k

Solution

01esp001.PRP	5k
01esp001_xs.lst	41k

Processing

01ESP001.htm	6k
01esp001.HKL	443k

Other Files

01ESP001.doc	790k
01esp001.mol	7k
01esp001_ellipsoid.gif	25k

Jmol

Figure 5 – eCrystals is an impressive example of a subject-specific repository supporting a range of file types and modelling tools and with a clear focus on interoperability.

The numerous IRs present within Southampton appear reflect a quite deliberate strategy being enacted within the University; this would be worth exploring through further work. Some repositories (as mentioned above) are unlikely to overlap regardless of scope, as researchers will choose to deposit in their local IR (with subject headings or metadata standards more suited to their outputs) rather than the central publications repository. This can be a complex issue in terms of metadata sharing and interoperability.

2.7 University of Strathclyde

Strathprints (<http://strathprints.strath.ac.uk>) is the institutional repository at the University of Strathclyde. On the "Policies" page, linked to via the "About Strathprints" page, it is stated that:

- This is an institutional or departmental repository.
- Strathprints holds all types of materials.
- Papers may include:
 - unpublished pre-prints (not peer-reviewed)
 - final peer-reviewed drafts (post-prints)
 - published versions
- Papers are individually tagged with their peer-review and publication status."

Independent of software package used, such statements are largely conventional, offering only minimal information on the range of material types and publication versions (pre-prints/post-prints) that are deposited in the IR. This particular scoping statement originated on OpenDOAR (<http://www.opendoar.org/tools/en/policies.php>) and it suggests that essentially *all* item types ("all types of materials") are in scope. However, there is no comprehensive, specifically-tailored note on the scope of Strathprints as a unique repository, and end users have to dig into the search options to find what item types are currently included. On the "Depositing in Strathprints" page it states that "the following types of item can currently be deposited:

- Journal articles
- Book chapters
- Monographs
- Conference papers
- Patents
- Reviews
- Theses"

Item types which can be searched by "advanced search" are:

Item type:	<input type="checkbox"/> Article	<input type="checkbox"/> Show/Exhibition
	<input type="checkbox"/> Book Section	<input type="checkbox"/> Composition
	<input type="checkbox"/> Monograph	<input type="checkbox"/> Performance
	<input type="checkbox"/> Conference or Workshop Item	<input type="checkbox"/> Image
	<input type="checkbox"/> Book	<input type="checkbox"/> Video
	<input type="checkbox"/> Thesis	<input type="checkbox"/> Audio
	<input type="checkbox"/> Patent	<input type="checkbox"/> Dataset
	<input type="checkbox"/> Review	<input type="checkbox"/> Experiment
	<input type="checkbox"/> Artefact	<input type="checkbox"/> Teaching Resource
		<input type="checkbox"/> Other

Figure 6 - an extract of the 'Advanced Search' page of Strathprints.

The OPAC's new SUpRimo interface, which replaces the Voyager service (<http://zen.lib.strath.ac.uk/>) does not allow users to search by format, even as part of an advanced search. The following statement is given: "Use our **SUPRIMO Service** to search our Library collections, including: books - journal titles - electronic resources - theses - exam papers - media resources". Thus scope cannot be readily

discerned. However, it is clear that book chapters/monographs and journal articles are common to the IR and OPAC.

3 Journal Articles

3.1 Cambridge University

Browsing the Journal Articles in **DSpace@Cambridge** for both published and submitted versions of articles there are found to be only 3 of each. There is currently no duplication with journal articles made available via the **Newton OPAC**. Examples of partial duplication (i.e. bibliographic/item level) can however be found, even with such a small sample set.

Taking the published version of the journal article "Alive and kicking: a progress report on Open Access, institutional repositories, and health information," published in the electronic journal "He@lth information on the internet" by author Peter Morgan, a search in **DSpace@Cambridge** retrieved the full Dublin Core metadata record for the item. Searching **Newton** for "Electronic resource" using the "Journal title" field for 'He@lth information on the Internet', one result is returned.

This is illustrated in the record transcripts below:

```
000 01674cms a22004577a 450
001 4278727
005 20080916100040.0
006 m d
007 cr n
008 981016c19989999enkbr p s 0 0eng
016 7_ |a 9808211 |2 DNLM
016 7_ |a SR0092289 |2 DNLM
022 ___ |a 1460-4140
035 ___ |a (WaSeSS)ssj0013504
040 ___ |a DNLM |c DNLM |d OCoLC |d OCoLC |d NLGGC |d CU-S |d WaSeSS
050 14 |a R118.2
060 00 |a W1 |b HE35225
082 _4 |a 610.285
130 0_ |a He@lth information on the Internet (Online)
245 00 |a He@lth information on the Internet |h [electronic resource].
246 3_ |a Health information on the Internet
260 ___ |a London : |b Royal Society of Medicine Press, |c c1998-
300 ___ |a v.
310 ___ |a Six times a year
362 0_ |a No. 1 (Feb. 1998)-
500 ___ |a Title from caption.
506 ___ |a Access restricted to authorised University of Cambridge users
516 ___ |a Text (electronic journal).
550 ___ |a Issued by the Royal Society of Medicine in association with the Wellcome Trust.
650 _0 |a Health |x Information services |v Periodicals.
650 12 |a Information Services |v Periodicals.
650 12 |a Information Storage and Retrieval |v Periodicals.
650 12 |a Internet |v Periodicals.
650 17 |a Websites. |2 gtt
650 17 |a Geneeskunde. |2 gtt
655 _0 |a Electronic journals.
710 2_ |a Royal Society of Medicine (Great Britain)
```

Project Acronym: OCRIS
 Document: Appendix 3 to OCRIS Final Report
 Version: 1
 Contact: Kathleen Menzies or Gordon Dunsire
 Date: September 2009

776 1_ |t He@lth information on the Internet (Online) |w (OCoLC)49849128
 850 ___ |a DNLN
 852 7_ |2 cammain |b online |h Online
 856 40 |z Full text available from Royal Society of Medicine Press: 01/06/2001 to present |u
<http://hii.rsmjournals.com>

Figure 7 – The MARC record from Newton for "He@lth Information on the Internet."

Full metadata record

DC Field	Value	Language
dc.contributor.author	Morgan, Peter	-
dc.date.accessioned	2008-02-06T16:11:12Z	-
dc.date.available	2008-02-06T16:11:12Z	-
dc.date.issued	2007-08	-
dc.identifier.citation	He@lth Information on the Internet, no.58 (August 2007), pp.6-8	en
dc.identifier.issn	1460-4140	-
dc.identifier.uri	http://www.DSpace.cam.ac.uk/handle/1810/195215	-
dc.description	An updated version of a talk given at the CILIP HLG Annual Conference, Eastbourne, 11 July 2006.	en
dc.description.abstract	The Open Access movement has promoted two parallel strands of information dissemination - OA publishing, and self-archiving in OA repositories. But while the original focus was peer-reviewed scientific research papers, repositories have extended their role well beyond the original OA concept, generating a lively debate about the role of the OA movement. This paper reviews recent developments internationally, examining different OA strategies, and looking in particular at activities within the field of health information.	en
dc.language.iso	en	en
dc.publisher	Royal Society of Medicine Press	en
dc.subject	open access	en
dc.subject	Institutional repositories	en
dc.subject	Health information	en
dc.title	Alive and kicking: a progress report on Open Access, institutional repositories, and health information	en
dc.type	Journal Article - Published Version	en
Appears in Collections:	Publications from University Library staff	

Figure 8 – The Dublin Core record retrieved for an article in "He@lth Information on the Internet".

3.2 University College for the Creative Arts

An advanced search for *Design* magazine using the Journal title field of the UCCA library catalogue retrieves collections located in each of the 5 libraries throughout UCCA: Canterbury, Epsom, Farnham, Maidstone and Rochester (Figure 9). Holdings information for the magazine is available at: <http://librarycatalogue.ucreative.ac.uk/uhtbin/cgiirsi.exe/APY2qAWNzW/CANTERBURY/245710023/2/1000>.

The Library OPAC retrieves records for volumes of the magazine published from 1949 as can be seen in Figure 10:

Publication info: London: Design Council.

Holdings [Change Holdings Display](#)

Canterbury Library	Copies	Material	Location
DESIGN	1	Journal	Journals Collection

Epsom Library	Copies	Material	Location
DESIGN	1	Journal	Journals Collection

Farnham Library	Copies	Material	Location
DESIGN	1	Journal	Journals Collection

Maidstone Library	Copies	Material	Location
DESIGN	1	Journal	Journals Collection

Rochester Library	Copies	Material	Location
DESIGN	1	Journal	Quiet Study Room - Rochester

Figure 9 – UCCA holdings information for "Design" magazine.

Design [Change Catalog Display](#)

ISSN: 0011-9245

Title: [Design](#).

Publication info: London : Design Council.

Current frequency: Quarterly, 1995-

Volume/date range: Rochester: 1961-1999 (Quiet Study Room)

Volume/date range: Canterbury: 1960-1998

Volume/date range: Maidstone: 1949-1999 (nos 164, 166-167, 176, 184, 187, 191, 193, 201, 208-209, 211-212, 218-219, 225, 232-233, 241, 243-249, 460-461, 464, 467, 485, 525, 534-536, 542, 544 missing.)

Volume/date range: Epsom: No.2 Feb 1949 to No.545, May 1994 (incomplete)

Volume/date range: Farnham: 1949 - 1951 then 1962 – 1998

General Note: Indexed in ABM, Art Abs, DAAI.

General Note: Publication ceased with no. 545, May 1994, resumed 1995, ceased after Spring 1999.

Figure 10 - Bibliographic records for volumes of the magazine *Design*, published from 1949 – 1999

Design Online is an electronic library containing a digitised record of *Design* magazine for the years 1965 to 1974. Through Design Online, VADS provided access to digitized copies of *Design* magazine between these dates, provided by the London College of Communication. It is part of the Digitisation in Art and Design (DIAD) project (<http://www.vads.ac.uk/resources/DIAD.html>) which, according to its webpages aims "to provide a digitised record of seminal journals in design and applied arts subjects within the Higher Education sector".

Design Online only retrieves records from the years 1965 to 1974, however, it supports access to "around 100 pages in each magazine, which are available as full screen size black-and-white or colour images". The magazines are therefore a searchable resource through the VADS web interface.



Figure 11 - the Search by year interface of Design online.

Although there may be some duplication in scope between IR and OPAC here, the OPAC provides more extensive coverage to this journal whilst the IR provides enhanced functionality to end users.

3.3 University of Glasgow

Taking a pre-print titled "Changing gender relations and environmental knowledges in Upper Egypt" from the Glasgow DSpace Service repository (<https://DSpace.gla.ac.uk/handle/1905/12>) which contains working papers, technical reports and other grey literature, Glasgow's other IRs and OPACs were interrogated for records of the same item.

As a **pre-print** it can be located in DSpace. As a **post-print**, it can be found in Enlighten (<http://eprints.gla.ac.uk/170/>) - which wouldn't store the "pre-print" version as it is a *publications* database.

The University Library's OPAC also has a record for the *Geographical Journal* in which the article was published.¹³ The post-print is then available electronically by various means (Expanded Academic ASAP, the JSTOR Arts and Sciences II Collection, the SocINDEX with Full Text and the Wiley-Blackwell Full Collection):

Geographical journal -- Expanded Academic ASAP 01 July 1994- Full
 details of this resource

Geographical journal -- JSTOR Arts and Sciences II Collection 01 Jan.
 1893- 31 Dec. 2003 Full details of this resource

Geographical journal -- SocINDEX with Full Text 01 Mar. 2001-
 Full details of this resource

Geographical journal -- Wiley-Blackwell Full Collection

Figure 12 - journal holdings available from Glasgow for the "Geographical Journal."

Interestingly, the title has changed between the pre- and post-printing. The pre-print is titled:

Changing gender relations and environmental knowledges in Upper Egypt,

While the post-print is:

Changing women's roles, changing environmental knowledges: evidence from Upper Egypt.

There is clear overlap between the scope of both the IRs and the OPAC in relation to information about and copies of journal articles (not just here, but across the board). Partly this foregrounds the necessity for clear notes on *version* with all IR deposits (this is already occurring at Glasgow and within EPrints software generally). It also flags up the fact that ensuring links exist between pre- and post-print versions and records can help end users find what they are looking for more readily, navigating through a "seamless web" of information rather than encountering "digital dead ends".

3.4 University of Southampton

The following examples of overlap have been drawn from ECS ePrints and TDNet: The University Library's Journal Catalogue.

Below is an Electronic Journals Catalogue record for the Journal *Electrophoresis* on TDNet.

The screenshot shows the University of Southampton TDNet website. The header includes the university logo and navigation links: Home, Search TDNet, My Profile, eBook Manager, About TDNet, Login, and Library Website. Below the header, a search bar indicates the search criteria: "You searched for: Journal Title: Electrophoresis;". The search results are displayed in a table with columns for Journal Title, Online Coverage, Publisher, Full Text Access, and Print Holdings.

Journal Title	Online Coverage	Publisher	Full Text Access	Print Holdings
Electrophoresis Wiley InterScience	1999 -	John Wiley and Sons, Inc.	<input checked="" type="checkbox"/>	
Journal of Electrophoresis DOAJ - Directory of Open Access Journals	2005 -	Japanese Electrophoresis Society	<input checked="" type="checkbox"/>	

Below the table, there is a "Jump To Page:" section with a dropdown menu set to "Electrophoresis --- End of Page" and a message: "2 record(s) retrieved for this search." A "New Search" button is located at the bottom of the search results area.

Figure 13 - Electronic catalogue record for the Journal *Electrophoresis*.

¹³ <http://eleanor.lib.gla.ac.uk/search~S6?/t%22the+geographical+journal%22/tgeographical+journal/1%2C5%2C9%2CB/frameset&FF=tgeographical+journal+online&1%2C1%2C/indxsort=->

Records for articles from this journal can be found in the **ECS eprints** repository. Figure 14 shows a record for one such article, by Cui, L., Holmes, D. and Morgan, H. and published in *Electrophoresis* in 2001. Only bibliographic data is available from the repository, however, the record does not direct end users to the TDNet catalogue to facilitate retrieval of the full text articles via the library's own journals subscription.

The screenshot shows the University of Southampton School of Electronics and Computer Science website. The main content area displays an article record for 'The dielectrophoretic levitation and separation of latex beads in microchips' by Cui, L., Holmes, D. and Morgan, H. (2001). The article is published in *Electrophoresis*, 22 (18), pp. 3893-3901. The record indicates that the full text is not available from this repository. Metadata includes: Creators: L. Cui, D. Holmes, H. Morgan; Item Type: Article; Research Group: Nanoscale Systems Integration Group; Deposited On: 24 May 2004 by Holmes, David; ID Code: 9377; Last Modified: 16 Nov 2007 03:51; Performance Indicator: EZ~03~02~11; Citations: Google Scholar: 47. The left sidebar contains navigation menus for ECS Research, Publications, and Intranet Tools. The bottom of the record includes a 'Tools' section with a link to 'Look up in Google Scholar', a 'Metadata' section with a 'Pageflow' dropdown and an 'Export' button, and a 'Corrections' section with a note that ECS staff and postgraduates may modify the record.

Figure 14 - ECS eprints IR article record.

Although the IR has implemented features such as "RSS Feeds" and "Look up in Google Scholar" as well as a Google Scholar Citation count, no link has been created to highlight the availability of relevant journals via the OPAC. Users are simply told "Full text not available from the repository".

Implementing SFX or another type of link based on the metadata about where the article appeared, or pulling in data via OAI-PMH would clearly be advantageous.

3.5 University of Strathclyde

Browsing **Strathprints** to locate an article using the "published" limiter and with the subject "philosophy" a record was retrieved for the article 'Children: animals or persons?' by Claire Cassidy, published in the journal *Thinking: The Journal of Philosophy for Children*, Volume 17, no3, pp. 13-16, 2004. As full-text access is not provided, users are then directed to an external site; rather than suggesting the library catalogue it point you to the Monclair State University host site for the journal, where the text is only available via subscription (<http://cehs.montclair.edu/academic/iapc/thinking.shtml>).

Thinking: the Journal of Philosophy for Children is, in fact, recorded in the OPAC where the full-text is available via the SFX journal search, with Wilson Web making the item available electronically. As an aside, the journal is not findable using "journal of philosophy for children" but only using "Thinking" as the title, as this is how it is referred to in Wilson Web's Journal Directory.

Children: animals or persons?

Cassidy, C. (2004) *Children: animals or persons?* Thinking: The Journal of Philosophy for Children, 17 (3). pp. 13-16. ISSN 0190-3330

Full text not available from this repository.

Official URL: <http://cehs.montclair.edu/academic/iapc/thinking.s...>

Item Type:	Article
ID Code:	5400
Keywords:	children; cognition; philosophy;
Subjects:	B Philosophy, Psychology, Religion > B Philosophy (General) B Philosophy, Psychology, Religion > BF Psychology
Department:	Faculty of Education > Educational and Professional Studies
Deposited By:	Strathprints Administrator
Deposited On:	15 Feb 2008
Last Modified:	21 Oct 2008 23:42

Repository staff only: [item control page](#)

Figure 15- Strathprint record retrieved for the article "Children: animals or persons?" by Claire Cassidy

3.6 National Marine Biological Library

The **Marine & Ocean Science EPrints Archive @ Plymouth** and the **National Marine Biological Library** both place, as noted previously, something of an emphasis on historical records and publications. In this area, it does not take long to find examples of overlap; so too in the area of more recent staff publications.

The **LibWeb OPAC** contains a bibliographic record for the following journal article:

<218793>

SE.I/GB.1

SIMKANIN, C.; POWER, A.-M.; MYERS, A.; MCGRATH, D.; SOUTHWARD, A.; MIESZKOWSKA, N.; LEAPER, R.; O'RIORDAN, R.

Using historical data to detect temporal changes in the abundances of intertidal species on Irish shores. *Journal of the Marine Biological Association of the United Kingdom*, 85(6), 1329-1340, 2005
[63 shores surveyed in both 1958 and 2003]

Figure 16 - Bibliographic OPAC record for a journal article by a member of MBL staff.

In the repository, the full-text of this article is available, yet no links are provided from the OPAC to the IR. The following screenshot shows the repository record for this publication:

Using historical data to detect temporal changes in the abundances of intertidal species on Irish shores

Simkanin, C. and Power, A.-M. and Myers, A. and McGrath, D. and Southward, A. and Mieszkowska, N. and Leaper, R. and O'Riordan, R. (2005) *Using historical data to detect temporal changes in the abundances of intertidal species on Irish shores*. Journal of the Marine Biological Association of the United Kingdom, 85 (6). pp. 1329-1340. ISSN 0025-3154

Full text available as:



PDF - Requires [Adobe Acrobat Reader](#) or other PDF viewer.
312 Kb

DOI: [10.1017/S0025315405012506](https://doi.org/10.1017/S0025315405012506)

Official URL: <http://journals.cambridge.org/action/displayJournal?jid=MBA>

Abstract

An historical data set, collected in 1958 by Southward and Crisp, was used as a baseline for detecting change in the abundances of species in the rocky intertidal of Ireland. In 2003, the abundances of each of 27 species was assessed using the same methodologies (ACFOR [which stands for the categories: abundant, common, frequent, occasional and rare] abundance scales) at 63 shores examined in the historical study. Comparison of the ACFOR data over a 45-year period, between the historical survey and re-survey, showed statistically significant changes in the abundances of 12 of the 27 species examined. Two species (one classed as northern and one introduced) increased significantly in abundance while ten species (five classed as northern, one classed as southern and four broadly distributed) decreased in abundance. The possible reasons for the changes in species abundances were assessed not only in the context of anthropogenic effects, such as climate change and commercial exploitation, but also of operator error. The error or differences recorded among operators (i.e. research scientists) when assessing species abundance using ACFOR categories was quantified on four shores. Significant change detected in three of the 12 species fell within the margin of operator error. This effect of operator may have also contributed to the results of no change in the other 15 species between the two census periods. It was not possible to determine the effect of operator on our results, which can increase the occurrence of a false positive (Type 1) or of a false negative (Type 2) outcome

Item Type: Article
Keywords: Abundance; community composition; intertidal environment; long-term changes; species diversity
Status: Published
Subjects: [Q Science > QL Zoology](#)
[Q Science > QH Natural history > QH301 Biology](#)
[Q Science > QH Natural history](#)
ID Code: 2219
Deposited On: 27 November 2007

Figure 17 - Full-text version of the article held in the Marine & Ocean Science Repository.

However, links from the OPAC to the IR are being instantiated in some cases, particularly in the case of historic records. Bibliographic data may be available in the OPAC while items themselves are held in the IR. In one example, the MBA website offers access to the repository copy of the 1968 publication *'Torrey Canyon' pollution and marine life. A report by the Plymouth Laboratory of the Marine Biological Association of the United Kingdom*, edited by J.E. Smith. Sections and chapters are available as separate PDFs while the complete volume is also available for download.

Torrey Canyon pollution and marine life. A report by the Plymouth Laboratory of the Marine Biological Association of the United Kingdom

Smith, J.E. , ed. (1968) 'Torrey Canyon' pollution and marine life. A report by the Plymouth Laboratory of the Marine Biological Association of the United Kingdom. Cambridge University Press , Cambridge (UK) . ISBN 0 521 07144 5

[Added March 2007]

[Download complete volume](#)
Download individual sections

Contents	p. iv-vi
List of plates	p. vii-ix
Preface	p. xi-xiv
Chapter 1. Introduction	p. 1-10
Chapter 2. Oil & detergents	p.11-22
Chapter 3. Sea surveys	p. 23-35
Chapter 4. Shore surveys – rocky shores	p. 36-74

Figure 18 - Hyperlinks visible in this webpage link the NMBA's bibliographic data to full-text chapters and other sections held in the Repository.

Although the LibWeb Serials Catalogue (which is separate from the main OPAC) contains listings of holdings information, item level records are not provided. This keeps the scopes of the IR and OPAC separate in terms of journal articles. Nevertheless, links to historic articles published in the *Journal of the Marine Biological Association of the United Kingdom* - records of which are held in the Institutional Repository - are provided on the Special Collections pages of the library website.



Edgar Johnson Allen F.R.S.
1866 -1942

MBA Director 1895-1936, then Governor. Elected to the Royal Society 1914.

Obituary: *Journal of the Marine Biological Association of the United Kingdom*, 25(4), pp671-684, 1943 by G.P. Bidder

The Library holds:

- his [published works](#) in the main collection
- his [personal and scientific papers](#) in the Archives

► [View full-text papers published in JMBA](#)

Figure 19 - Links provided on the OPAC pages to historical works by former members of the Marine Biological Association.

Results for Simple Search

Refine search	New search
-------------------------------	----------------------------

Authors/Creators/Editors matches all of "allen,e". Results ordered by year (most recent first).

Displaying results **1** to **35** of **35**. Search time: 1s.

[Index last built at: Wed Sep 9 12:19:34 BST 2009]

- Allen, E. J. (1941) [Obituary Arthur James Smith 1874-1941 A personal tribute by E.J. Allen](#). The Journal of the Marine Biological Association of the United Kingdom, 25 (2). pp. 223-224.
- Allen, E. J. (1940) [Miss A.R. Clark: Obituary](#). Journal of the Marine Biological Association of the United Kingdom, 24 (2). 459.
- Allen, E. J. (1938) [Obituary. Edward Thomas Browne. 1866-1937](#). Journal of the Marine Biological Association of the United Kingdom, 22 (2). pp. 405-410.
- Allen, E. J. and Harvey, H.W. (1928) [The laboratory of The Marine Biological Association at Plymouth](#). Journal of the Marine Biological Association of the United Kingdom, 15 (3). pp. 735-752.
- Allen, E. J. (1928) [Arthur Everett Shipley](#). Journal of the Marine Biological Association of the United Kingdom, 15 (1). 361.
- Allen, E. J. (1928) [Walter Campbell De Morgan](#). Journal of the Marine Biological Association of the United Kingdom, 15 (1). pp. 362-364.
- Allen, E. J. (1927) [Fragmentation in the genus Autolytus and in other syllids](#). Journal of the Marine Biological Association of the United Kingdom, 14 (4). pp. 869-876.
- Allen, E. J. (1919) [A contribution to the quantitative study of plankton](#). Journal of the Marine Biological Association of the United Kingdom, 12 (1). pp. 1-8.

Figure 20 - Some of the records retrieved from Marine and Ocean Sciences@Plymouth for Edgar Johnson Allen.

4 Theses

4.1 University of Glasgow

The first item listed when searching for the term "heart" on the "Department of Cardiac Surgery – Ph D. Theses (Cardiac Surgery) Collection" within the **Glasgow DSpace Service** is the PhD thesis of Michael Leon James Rose:

Date of Issue 1998
Title Development of a muscle powered blood pump : fluid mechanic considerations
Authors Rose, Michael Leon James PhD
Item Type Thesis

Figure 21 - DSpace record for a Cardiac Surgery PhD Thesis.

A record for the thesis is also retrievable from the **Enlighten** repository, as illustrated below:

Development of a muscle powered blood pump: fluid mechanic considerations

Rose, Michael Leon James (1998) *Development of a muscle powered blood pump: fluid mechanic considerations*. PhD thesis, University of Glasgow.

Item Type: Thesis (PhD)

Qualification Level: Doctoral

Project Acronym: OCRIS
 Document: Appendix 3 to OCRIS Final Report
 Version: 1
 Contact: Kathleen Menzies or Gordon Dunsire
 Date: September 2009

Subjects: [R Medicine > R Medicine \(General\)](#)
[T Technology > TJ Mechanical engineering and machinery](#)
Faculty and Department: [Medicine > Division of Cardiovascular and Medical Sciences](#)
 Supervisor: Theses, Redirect Id
Date of Award: 1998
Unique ID: glathesis: 1998-84
Copyright: Copyright of this thesis is held by the author.
Deposited On: 30 January 2008

Figure 22 - Record for the same Thesis item, retrieved from the Enlighten IR at Glasgow.

Again, both the **classic catalogue** and the **Encore RDP** turn up records for this item, where links to the electronic version within Enlighten are prominent.

KEYWORDS (1-2 of 2)

Most relevant titles entries 1-2

1 [Development of a muscle powered blood pump : fluid mechanic considerations / Michael Leon James Rose](#) 

Rose, Michael Leon James.
1998.

Status	Location	Shelf-mark
REFERENCE ONLY	Level 12 Spec Coll	Thesis 11494
REFERENCE ONLY	Library Research Annexe	Thesis 11494 c.2

2 [Development of a muscle powered blood pump \[electronic resource\] : fluid mechanic considerations / M](#) 

Rose, Michael Leon James.
1998.
[Connect to electronic resource](#)

Figure 23 - How the RDP presents results for different manifestations of a thesis.

There is a clear duplication of scope across IRs, OPACs and RDPs with regards to the theses generated by the University - but as results are interfiled or links between the two are established, there is no "silo" effect.

4.2 University of Strathclyde

As an example of records duplication for theses at the University of Strathclyde, records for the PhD thesis of Kevin D. O’Gorman published by the University of Strathclyde in 2008 and entitled, "The Essence of Hospitality from the Texts of Classical Antiquity: The development of a hermeneutical helix to identify the origins and philosophy of the phenomenon of hospitality" were selected.

Both the University’s IR, **Strathprints** and the **Strathclyde Library Catalogue** were found to hold duplicate records for this thesis. However, links between bibliographic OPAC records and PDFs in Strathprints are being accomplished through the use of the **Primo Resource Discovery Platform** (known as **SUPrimo**) which is discussed further in Section 8.



Subjects matches "B Philosophy (General)" AND Item Type matches any of "Thesis"

Displaying results 1 to 1 of 1.
[Refine search](#) | [New search](#)

Order the results: by year (most recent first)

Export 1 results as

O'Gorman, Kevin D. (2008) *The Essence of Hospitality from the Texts of Classical Antiquity: The development of a hermeneutical helix to identify the origins and philosophy of the phenomenon of hospitality*. PhD thesis, University of Strathclyde. 

1. [The development of a hermeneutical helix to identify the origins and philosophy of the phenomenon of hospitality](#). PhD thesis, University of Strathclyde.

Displaying results 1 to 1 of 1.
[Refine search](#) | [New search](#)

Order the results: by year (most recent first)

Figure 24- Strathprints record for the PhD thesis of Kevin D. O'Gorman.

University of Strathclyde Library Catalogue

[New Search](#) | [Titles List](#) | [Headings List](#) | [Your Account](#) | [Login](#) | [Make Request](#) | [Search History](#) | [Other Libraries](#) | [Exit](#)

[Display Detailed Record](#)

Author: O'Gorman, Kevin D.
Title: The essence of hospitality from the texts of classical antiquity : the development of a hermeneutical helix to identify the philosophy of the phenomenon of hospitality / Kevin D O'Gorman

Publication Details: 2008.
Description: 323 leaves : col. ill. ; 30 cm.
Notes: Strathclyde theses - ask staff. Thesis no. : T12003
One copy only. Not for lending.

Location: Main Library Theses Reference - Make request
Shelved at: T12003

Number of items: 1
Status: check shelves

print/save/email selected records in Brief, Detailed, or RIS format
select required format [RIS?] Brief record
enter your email address:

Figure 25 - Strathclyde's Library Catalogue record for the PhD thesis of Kevin D. O'Gorman.

Despite theses duplication, none of the 12 patents held in the Strathprints IR are discoverable in the OPAC. For example, whilst the thesis for the author Stuart Mackay of the University of Strathclyde's Business School is recorded in the OPAC, his patent (WO/2004/086771) for a desktop display which is held in Strathprints, it appears, is not. This suggests that despite the OPAC claiming to collect metadata on patents, some recent ones generated by institutional staff may be going instead into the IR and evading the OPAC altogether.

Staff working with Strathprints informed us that the University Library intends to begin making theses available via its Digitool service rather than through the IR, despite the fact that theses are in scope for Strathprints and some are already deposited there. This will potentially create duplication and fragmentation for the end user.

Indeed, there is already duplication of content from the "Strathclyde Papers" Collection, recorded both in **Digitool**

(<http://oleg.lib.strath.ac.uk:8881/R/SHC37BGJYVVDJ4BQHNFTMLNBT33EYMS3LLHQVM7D>)

L89AMRNTY-00723?func=collections-result&collection_id=1119) and **Strathprints**, which holds many of the same papers (although, having been withheld when copyright clearance was being sought, the full-text is not yet available in Strathprints). Staff are now considering whether it is worth adding links to the full-text versions or removing the records altogether.

Indeed, there is much debate at present about whether Digitool or Strathprints would be a better repository for theses – for example, Digitool cannot provide a unique URL for each item whereas EPrints can. The library perceives cataloguing e-theses in Digitool as a natural extension of their existing work with print theses. Yet Strathprints is seeking to encourage the deposit of institutional outputs and would like theses included in *its* collection.

Perhaps such complications concerning the rationalisation of collection policy will undermine the role of Strathprints as it begins to become more central to planning for the REF. Records from the existing "2Research Metrics" database, currently administered by the Department of Engineering are already being transferred into Strathprints with a view to replacing the old system.

5 Name authorities and publications by author

5.1 Cambridge University

Taking the author Arandjelović, Ognjen, a member of staff at the Department of Engineering, a search of the **Universal Catalogue** retrieves the following record:

Main Author: Arandjelović, Ognjen . Title: Automatic face recognition from video / Ognjen Arandjelović. Other Entries: University of Cambridge . Department of Engineering . Description: 323 leaves : ill. ; 30cm. Notes: Date approved: 20 November 2007. Thesis (Ph.D) -- University of Cambridge, 2007. Format: Archival/Manuscript Material
Location: UL: Order in Manuscripts Room (Not borrowable) Classmark: PhD.30754. Number of Items: 0

Figure 26 - edited transcript of the catalogue entry for a member of the Department of Engineering.

Ognjen has not submitted anything to **DSpace@Cambridge**. However, in the **CUED Publications Database** there are 29 records for "Arandjelović" (<http://publications.eng.cam.ac.uk/view/people/oa214.html>) including his Masters Thesis but notably not the Ph.D Thesis catalogued in the OPAC.

Interestingly, from the CUED record, it can be seen that he is identified as being "person: oa214" which is his institutional (or at least his departmental ID). His website uses the same reference within its URL (<http://mi.eng.cam.ac.uk/~oa214/academic/>).

There are problems encountered, however, with regards to naming conventions across the IR and OPAC for both authors and the Departments to which they belong.

In CUED Arandjelović is referred to as "Arandjelovic, Dr O." but in the OPAC his Main Author entry is "Arandjelović, Ognjen". Here, he is also associated with the Corporate Name "University of Cambridge. Department of Engineering."

As indicated earlier, the Department's own IR is referred to as "Cambridge University Engineering Department", whilst in DSpace@Cambridge, the Department is simply called "Department of Engineering". This has implications which are discussed further in Section 5.1.

5.2 University of Glasgow

Using the example of author "Rogatchevski, Andrei", a member of staff at Glasgow in the Slavonic Studies Department (School of Modern Languages and Cultures) a number of records for published

research output were retrieved across the OPAC, RDP and both the Enlighten, DSpace and Thesis Service IRs. The **OPAC** retrieved this record:

Eduard Limonov : a critical study : thesis presented for the degree of Ph.D. in the Department of Slavonic Languages and Literatures of the University of Glasgow / by Andrei Rogatchevski.

This is the only entry. Simply using the name "Rogatchevski" to search, entries are returned for "Rogatchevski Andrei" and "Rogatchevski Andrei Course Convenor" (here, the name appears in the "Alt Author" field).

The screenshot shows a search interface with the following elements:

- Buttons: New Search, Limit/Sort, Another Search, Other Resources.
- Search criteria: AUTHOR (dropdown), rogatchevski (text input), All Formats (dropdown).
- Sort options: System Sorted (dropdown), Sort, Search.
- Checkbox: Limit search to available items.
- Buttons: Save Marked Records, Save All On Page.
- Table of results:

Num	Mark	AUTHORS (1-2 of 2)	Entries Found
1	<input type="checkbox"/>	Rogatchevski Andrei	1
2	<input type="checkbox"/>	Rogatchevski Andrei Course Convenor	1

Buttons: Save Marked Records, Save All On Page.

Buttons: New Search, Limit/Sort, Another Search.

Figure 27 - Searching the OPAC using just the surname returns 2 sets of results.

Records for the author were also retrieved from **Encore, Enlighten** and **DSpace**:

1. Eduard Limonov : a critical study : thesis presented for the degree of Ph.D. in the Department of Slavonic Languages and Literatures of the University of Glasgow / by Andrei Rogatchevski Rogatchevski, Andrei.

REFERENCE ONLY - Level 12 Spec Coll - Thesis 11232

Show all 2 copies/volumes

Location	Call No.	Status
Level 12 Spec Coll	Thesis 11232	REFERENCE ONLY
Slavonic Lang Dept Lib	Thesis 11232 c.2	NOT KNOWN 1998

2. Slavonic Studies 1B:Post-War cinema articles

Figure 28 - extract of records retrieved from Encore for Rogatchevski.

Searching **Enlighten** using the 'Author Field' returns:

A biographical and critical study of Russian writer Eduard Limonov

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

Rogatchevski, A. (2003) *A biographical and critical study of Russian writer Eduard Limonov*. Studies in Slavic language and literature, 20 . Edwin Mellen Press. ISBN 0773468471
Full text not available from this repository
Item Type:Book
Status:Published
Authors:Rogatchevski, A.
Subjects: [P Language and Literature > PN Literature \(General\)](#)
Faculty/Department:[Faculty of Arts > Department of Slavonic Studies](#)
Publisher:Edwin Mellen Press
Unique ID:glasEPrints:2003-5185
ID Code:5185
Deposited On:23 Mar 2009 10:56
Last Modified:23 Mar 2009 10:56

Figure 29 – Results from Enlighten for a member of the Slavonic Studies Department.

In **DSpace** the thesis and course materials are not in scope, but a Conference Paper is:

Title: Aliens in New York: Slavonic émigré Authors as Taxi Drivers
Authors: Rogatchevski, Andrei Date of Issue: 2003
Abstract: In order to ascertain whether recent Russian fiction sheds new light on the topos of taxi driving, the author would like to give a detailed analysis of Vladimir Lobas's autobiographical novel *Zheltye koroli* (*Novyi mir*, 1-3, 1991; written in early to mid-1980s, it was translated into English as *Taxi from Hell*, New York, Soho Press, 1991), dealing with the experiences of a Jewish émigré taxi driver in New York. Furthermore, to determine whether there are any a) specifically Russian and b) specifically male components in Lobas's take on things, the author would like to compare his book to Iva Pekárková's autobiographical novel *Dej mi ty prachy* (1996; translated into English as *Gimme the Money*, London, Serpent's Tail, 2000), written on the same subject some ten years after Lobas's book (Pekárková also earned a living as a cab driver in New York while she was living there during her period of emigration between 1985 and 1996).
URI: <http://hdl.handle.net/1905/23>
Item Type: Conference Paper
Appears in Collections: [13th International Congress of Slavists](#)

Figure 30 - DSpace returns one result for Rogatchevski.

Clearly there *is* scope duplication across Glasgow's systems; although the thesis and book are findable in both the relevant (in terms of scope) IR and the OPAC, it is notable that Rogatchevski's thesis is not recorded in the Glasgow Theses Service IR.

6 Name authorities and research departments and groups

6.1 University of Southampton

With 10 IRs (most but not all arranged on a discipline/subject basis), the University of Southampton presents a clear opportunity to compare and contrast a variegated suite of systems. Here, overlap between **e-Prints** and the **ECS EPrints** repositories is explored.

Although ECS e-Prints records are clearly in scope for e-Prints Soton, there are significantly more publications in the ECS's own "Nanoscale Systems Integration Group" folder (1746) than there are in the corresponding section of e-Prints Soton (25).

Taking the year 2007 as a test sample, 3 records are found to be common to the two repositories:

In e-Prints Soton:

de Planque, M.R.R., Raussens, V., Antoranz Contera, S., Rijkers, D.T.S., Liskamp, R.M.J., Ruyschaert, J-M., Ryan, J.F., Separovic, F. and Watts, A. (2007) [β-sheet structured β-amyloid\(1-40\) perturbs phosphatidylcholine model membranes.](#) *Journal of Molecular Biology*, 368, (4), 982-997. ([doi:10.1016/j.jmb.2007.02.063](#))

Ramanujan, C.S., Sumitomo, K., de Planque, M.R.R., Hibino, H., Torimitsu, K. and Ryan, J.F. (2007) [Self-assembly of vesicle nanoarrays on Si: a potential new route to high-density functional protein arrays.](#) *Applied Physics Letters*, 90, (3), 033901-033903. ([doi:10.1063/1.2431774](#))

Gonzalez, D.C., Kiziroglou, M.E., Li, X., Zhukov, A.A., Fangohr, H., de Groot, P.A.J., Barttlet, P.N. and de Groot, C.H. (2007) [Long range ordering in self-assembled Ni arrays on patterned Si.](#) *Journal of Magnetism and Magnetic Materials*, 316, (2), e78-e81. ([doi:10.1016/j.jmmm.2007.02.059](#))

Figure 31 – an edited version of results retrieved from e-Prints Soton, showing duplication of bibliographic records between e-Prints and ECS e-Prints (fig.25).

In ECS EPrints:

de Planque, M. R. R., Raussens, V., Antoranz Contera, S., Rijkers, D. T. S., Liskamp, R. M. J., Ruyschaert, J. M., Ryan, J. F., Separovic, F. and Watts, A. (2007) [β-Sheet Structured β-Amyloid\(1-40\) Perturbs Phosphatidylcholine Model Membranes.](#) *Journal of Molecular Biology*, 368 (4). pp. 982-997.

Ramanujan, C. S., Sumitomo, K., de Planque, M. R. R., Hibino, H., Torimitsu, K. and Ryan, J. F. (2007) [Self-assembly of vesicle nanoarrays on Si: A potential route to high-density functional protein arrays.](#) *Applied Physics Letters*, 90 (3). 033901.

Gonzalez, D. C., Kiziroglou, M. E., Li, X., Zhukov, A. A., Fangohr, H., de Groot, P. A. J., Barttlet, P. N. and de Groot, C. H. (2007) [Long range ordering in self-assembled Ni arrays on patterned Si.](#) *Journal of Magnetism and Magnetic Materials*, 316 . e78-e81. (In Press).

Figure 32 – the corresponding records in ECS e-Prints.

Neither system makes the full-text available and both provide links to the official location of the full-text article at Science Direct. However, while the citation details presented are nearly identical, it is notable that the research group is referred to across both IRs by different naming conventions.

The "Nanoscale Systems Integration Group" is referred to on its webpages as "Nano" and "Nano Research Group" – and it is unclear what the definitive name for this group within the University is. ePrints however, reflects accurately the structure of the University, as shown in Figure 2.

The screenshot shows the e-Prints Soton website interface. At the top, the logo for e-Prints Soton, University of Southampton, is displayed. Below the logo is a navigation menu with buttons for 'home', 'my eprints', 'browse', and 'search'. The main content area shows a header for a view titled '["viewname_eprint_structure" not defined]: Nano-Scale Integration Group'. Underneath, a hierarchical list is shown:

- [University Structure \(42442\)](#)
 - [Faculty of Engineering, Science and Mathematics \(17581\)](#)
 - [School of Electronics and Computer Science \(452\)](#)
 - Nano-Scale Integration Group (25)

Below the list, a horizontal line is followed by the text 'Number of items: 25.'

Figure 33 - Extract from the 'University Structure' hierarchy used in e-Prints Soton.

ECS EPrints, with its scope tied to a specific school (Electronics and Computer Science), organises items by Research Group, an example of which is shown below:

The screenshot shows a hierarchical list of research groups from the Cogprints at Southampton website. The list is as follows:

- * ECS Groups (13989)
 - Pervasive System Centre (43)

Figure 34 - extract from Cogprints at Southampton showing browse-able Research Group hierarchy.

Although these Research Groups may match the way they are referred to internally, the structure of the University and the departments of which they are a part is not referred to in the records. Only the general "ECS Groups" label is applied. However, it is worth noting that on the main webpage of the ECS Department, the term "Pervasive Systems Centre" (complete with extra "s") is used.

6.2 University of Strathclyde

The list of "Departments" used by **Strathprints** reflects and matches the institutional hierarchy at the University of Strathclyde and uses the University's in-house coding scheme, which matches the names of the Departments across the spectrum of the University.

A publication by Senior Lecturer Mark Sheridan from The Department of Creative and Aesthetic Studies, for example, can be found by navigating through the following browseable hierarchy:

University Departments (6910)

- Faculty of Education (391)
 - **Creative and Aesthetic Studies**

Clearly this is extremely useful at an institutional level and this structure might be reflected well, for instance, within a shared XML record, provided that "University of Strathclyde" sits at the top of the tree.

Abstract

The conference will bring together some of the world's most foremost authorities on the teaching of composition and improvisation in elementary and secondary schools. Three days of workshops and presentations on the latest teaching ideas and scholarship on musical creativity.

Item type:	Conference or Workshop Item (Paper)
ID code:	3323
Keywords:	education; Scottish music culture; music; composition;
Subjects:	M Music and Books on Music > MT Musical instruction and study L Education > L Education (General)
Department:	Faculty of Education > Creative and Aesthetic Studies
Deposited by:	Strathprints Administrator
Deposited on:	12 May 2007
Last modified:	20 Oct 2008 22:46

Figure 35 – Partial record for a publication by Senior Lecturer Mark Sheridan at Strathclyde (<http://strathprints.strath.ac.uk/3323/>).

At record level, author names are constructed in a variety of ways, depending on the format type a user wishes to export the record in. In the OPAC, there is a record for the author Stuart McKay (a member of the University's Business School) where he is referred to under the personal name "McKay, Stuart".

McKay, S - if you look at the Dublin Core record and EndNote.
McKay, Stuart - in the EP3 XML.
S. McKay - in BibTeX.
Stuart McKay - in the ASCII citation and HTML.
McKay, Stuart - in JSON
S. McKay - in METS
McKay, S. - in Open URL Context Object and Reference
S. McKay - in Reference Manager.

Figure 36 – the many ways Stuart McKay's name is rendered by various format standards.

It might be noted that the author's name is formatted according to the same rules in both the OPAC and the EP3 XML encoding. Could these formats be potential bridges between the systems?

In **Strathprints**, McKay's thesis is findable via the following Departmental headings - "StrathclydeBusiness School > Hospitality and Tourism Management" - whilst in the OPAC the following term is associated with the record: "University of Strathclyde. Dept. of Hospitality and Tourism Management".

Such relative inconsistency has clear implications for current cross-searching, the exchange of metadata and the interoperability of both systems. Further, it could conceivably have a detrimental impact upon the compilation of accurate quantitative data on the research outputs of both individual authors and university departments for the purposes of research evaluation.

7 Subject Headings

7.1 Cambridge University

The CUL **Newton** OPAC does not provide information on subject coverage. Users must click through into individual MARC records to reveal the LCSH in use. This is typical of HEI OPACs because of course, their scopes are extremely general and at the same time, comprehensive. Subject fields can be used for keyword searching by end users but they do not tend to adopt the approach taken by IRs of offering browseable subject menus; these would likely become unwieldy (certainly in the case of

Cambridge's OPAC, which includes holdings information for 89 libraries). Nor is there a standard classification scheme in use across Cambridge's libraries; records may use different terms to describe the same subject.

It is interesting to compare some of Newton's LCSH terms with those used by the Cambridge University Engineering Department Repository (**CUED**), which applies two levels of classification to its subject headings: "Library Classification" and "Research Areas" (as illustrated below) with an approximately equivalent number of items represented for each. Again, there are an approximately equal amount of **CUED** library classifications as there are areas of research, and it is not clear whether these are the same items discoverable through different Library/Research subject searches.

Where Newton will use "Mechanical Engineering", or "Mechanics – Applied", CUED will use "Mechanics and Mechanics of Machines". The CUED system of classification appears fairly structured, and may well be based to some extent on LCSH or a conflation of its terms (some terms are indeed the same, such as "Fluid Mechanics" which is an LCSH authority heading). In another example, where CUED uses "Computers (Programming)", Newton uses the Subject heading "Programming (Electronic computers)".

There is no single recognised standard being applied to the subject headings in use across *both* systems. They therefore cannot efficiently be mapped or cross-searched using a single standard. Whilst interoperability of the two systems is not currently being planned, the lack of common standards could prove prohibitive if it were to be sought in future.

Subject: Subjects (CUED Library Classification and Research Areas)

- **Subjects (CUED Library Classification and Research Areas) (13632)**
 - CUED Library Classification (13446)
 - CUED Research Areas (13550)

Subject: CUED Library Classification

- Subjects (CUED Library Classification and Research Areas) (13632)
 - **CUED Library Classification (13446)**
 - A - History, Biography, Design, Useful Arts (1557)
 - B - Management, Behavioural Sciences, Psychology, Law, Economics, Education (905)
 - D - General Engineering (General to groups) (1043)
 - E - Mathematics (165)
 - F - Structures (1060)
 - G - Civil Engineering, Soil Mechanics, Surveying (1003)
 - J - Materials, Workshop Processes including Welding (1189)
 - L - Physics and Chemistry (where not placed with Groups) (739)
 - N - Electricity (3754)
 - Q - Automatic Control (1179)
 - R - Mechanics and Mechanics of Machines (452)
 - T - Fluid Mechanics, Aeronautics, Hydraulics (1297)
 - V - Thermodynamics, Heat Transfer (776)
 - X - Nuclear Engineering, Chemical Engineering (65)

Subject: CUED Research Areas

- Subjects (CUED Library Classification and Research Areas) (13632)
 - **CUED Research Areas (13550)**
 - Civil, Structural, Environmental and Sustainable Development (1860)
 - Electrical Engineering (2103)
 - Energy, Fluid Mechanics and Turbomachinery (1981)
 - Information Engineering (2673)
 - Manufacturing and Management (1716)
 - Mechanics, Materials and Design (3197)

- Other Research Themes (316)

Figure 37 - CUED library classifications and areas of research.

7.2 University of Hull

The University of Hull **OPAC** can be searched using keyword searches of top-level "Subjects" while the **eDocs Institutional Repository** operates using un-controlled or informal subject "tags", generated (one assumes) by those depositing. Thus, while the IR lets you browse by "Subject", no controlled subject vocabulary is in place.

Compare this extract of a subject list from the OPAC:

- Accounting and Finance
- American Studies
- Art History
- Biological Sciences
- Biomedical Sciences
- Business and Management
- Chemistry
- Coastal Marine Biology
- Computer Science incl. Internet Computing
- Creative Media Technology

With an extract from one given in the IR:

- Isotopically substituted compounds
- Ja-sig, jisc, repomann, portal
- Jisc
- Johnson, magnus
- Kent
- Law
- Law and regulation
- Ledauniece, edith
- Legal material
- Legal matters
- Leicestershire
- Leucaspius delineatus
- Licence
- Lincolnshire
- Liquid crystal
- Little domesday
- Lloyd, Christopher
- Managing change

It is not immediately clear what guidelines are available to end users on how to submit materials, what the policies and procedures are in place, or how metadata is generated. This is due to the fact that authors currently submit content through an authenticated interface (using Muradora) which is not visible to the public and which Hull has developed via various JISC projects.¹⁴ Thus we cannot readily discern what options are presented to those making use of the repository.

We can observe however that Subject terms are often inconsistently assigned and that varying amounts of granularity are used (some very high-level, others very specific).

¹⁴ See: <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2005/repomman.aspx> [Accessed 26th August 2009].

For example, one of the most popular Collections relates to the data and structure of the Domesday Book Online project, yet some items relating to this may be classified as being about "Kent" or about "Statistics" or "domesday", "domesday book", "domesday boroughs", "domesday statistics" and even erroneously, "domeday".

A representative of the IR at Hull informed OCRIS:

"We haven't standardised on subject headings yet, but have focused around what is most useful in the cases in question. For some areas, like e-theses, we need to start using formal headings like LoC, but for others, like committee papers, the subject headings will be standard but remain local."¹⁵

The adoption of a "local" system of subject classification may do little to enhance future interoperability with the OPAC or other systems however the repository is undergoing constant development; no doubt such issues are being addressed. With its focus being primarily on administrative needs and digital rights management issues pertaining to content clearly needing curated by the institution.

At present **eDocs** focuses on gathering digital administrative materials which the University has a clear need to manage, rather than Open Access collections or research papers. As the repository begins to gather "published articles and conference papers, book chapters, and other forms of written research output from UoH academic staff" (<http://edocs.hull.ac.uk/muradora/aboutRepository.jsp>) as is planned, it will be interesting to see how its application of metadata evolves, in line with the aims of RepoMMan and other projects. As eDocs staff have stated: "the search for a tool that would fulfil [the requirement for descriptive metadata relating to textual objects] has been likened to the search for the Holy Grail".¹⁶

7.3 National Marine Biological Library

Usefully, indices are provided by the NMBL OPAC to assist with keyword searching, allowing narrower, broader or related terms to be selected. A typical indice offered in response to the keyword "zooplankton" can be seen below:

¹⁵ Personal correspondence via email.

¹⁶ JISC. (2007). *RepoMMan Final Project Report: September 2007*. Page 15. Available online at: <http://ie-repository.jisc.ac.uk/125/> [Accessed 26th August 2009].

National Marine Biological Library

Keywords anywhere in record **zooplankton**

[Home] [Search] [History] [About] [Help]

[Select] [PgDn] [PgUp] [GO TO] [Go back]

<input type="checkbox"/>	ZOOPLANKTON	1955
<input type="checkbox"/>	ZOOPLANKTONFAUNAN	1
<input type="checkbox"/>	ZOOPLANKTONFAUNANS	1
<input type="checkbox"/>	ZOOPLANKTONIC	16
<input type="checkbox"/>	ZOOPLANKTONIN	1
<input type="checkbox"/>	ZOOPLANKTONORGANISMEN	1
<input type="checkbox"/>	ZOOPLANKTONS	3
<input type="checkbox"/>	ZOOPLANKTONUNTERSUCHUNGEN	1
<input type="checkbox"/>	ZOOPLANKTONVERTEILUNG	1
<input type="checkbox"/>	ZOOPLANKTOVOROUS	1
<input type="checkbox"/>	ZOOPLANTONUNTERSUCHUNGEN	1
<input type="checkbox"/>	ZOOPLNAKTON	1
<input type="checkbox"/>	ZOOREMEDIATION	1
<input type="checkbox"/>	ZOOROB	1
<input type="checkbox"/>	ZOOROB, G.	1
<input type="checkbox"/>	ZOOS	2
<input type="checkbox"/>	ZOOSANIATAIRE	1
<input type="checkbox"/>	ZOOSPORE	4
<input type="checkbox"/>	ZOOSPORES	26
<input type="checkbox"/>	ZOOSPORIC	3

[Select] [PgDn] [PgUp] [Go back]

Figure 38 - a typical controlled subject term index within the NMBL's LibWeb.

Geographic and subject indexing terms have been included only "for publications reporting on British estuaries and sea areas or marine pollution."

Although complex, the OPAC is set up in a way designed to reflect the knowledge of experienced researchers familiar with the terminology being used. With access authenticated it may be unlikely that those outwith marine science and environmental sciences will gain access; yet while such users can be expected to be familiar with the terminology in use, they may nevertheless find the interface and arrangement a hindrance.

The IR is by contrast extremely general with little customisation, using default LCC indices supplied with EPrints and the following list of item types:

- Article
- Book Section
- Monograph
- Conference or Workshop Item
- Book
- Thesis
- Patent
- Other
- Multimedia

- Maps

The vast majority of LCC numbers will be inappropriate to **eprints@Plymouth** given its very specific scope. Searching between the OPAC and IR, users may find both the best aspects of user interface design (customisation; simplicity of search and browse) and the worst (complicated rules and terminology in the OPAC requiring much explanation; lack of focus and specificity) demonstrated by the NMBA systems.

It will be interesting to see how the repository develops when the upgrade is complete and as a greater number of current research outputs are added and whether further links with the OPAC will be established.

7.4 *University of Southampton*

There are clear differences among the various systems at Southampton in the case of both Subject and Name Authorities. The **OPAC** uses LCSH, whilst **e-Prints Soton** uses the Library of Congress Classification (erroneously referring to these as "Subject Areas"). These are not in fact subjects but broad category classifications. **Cogprints** encompasses items from various departments and disciplines and usefully provides a broad statement on scope and subject coverage on its homepage:



Welcome to **CogPrints**, an electronic archive for [self-archive](#) papers in any area of [Psychology](#), [Neuroscience](#), and [Linguistics](#), and many areas of [Computer Science](#) (e.g., [artificial intelligence](#), [robotics](#), [vision](#), [learning](#), [speech](#), [neural networks](#)), [Philosophy](#) (e.g., [mind](#), [language](#), [knowledge](#), [science](#), [logic](#)), [Biology](#) (e.g., [ethology](#), [behavioral ecology](#), [sociobiology](#), [behaviour genetics](#), [evolutionary theory](#)), [Medicine](#) (e.g., [Psychiatry](#), [Neurology](#), [human genetics](#), [Imaging](#)), [Anthropology](#) (e.g., [primatology](#), [cognitive ethnology](#), [archeology](#), [paleontology](#)), as well as any other portions of the [physical](#), [social](#) and [mathematical](#) sciences that are pertinent to the study of cognition.

Figure 39 - statement on scope and subjects on the Cogprints homepage.

Inside, Cogprints has adopted a mixed and more granular approach. LCSH, with some use of narrower terms or sub-divisions, alongside in-house headings deemed useful to users reflect the specificity of items in the collection. As illustrated below, journal titles are erroneously used within "Subject" menus alongside subject headings:

Browse by Subject

Please select a value to browse from the list below.

- [Subject Areas](#) (3358)
 - [Biology](#) (522)
 - [Animal Behavior](#) (132)
 - [Animal Cognition](#) (116)
 - [Behavioral Biology](#) (88)
 - [Cognitive Archeology](#) (1)
 - [Ecology](#) (15)
 - [Ethology](#) (75)
 - [Evolution](#) (173)
 - [Population Biology](#) (12)
 - [Primateology](#) (29)
 - [Sociobiology](#) (60)
 - [Theoretical Biology](#) (143)
 - [Computer Science](#) (1016)
 - [Artificial Intelligence](#) (504)
 - [Complexity Theory](#) (123)
 - [Dynamical Systems](#) (136)
 - [Human Computer Interaction](#) (49)
 - [Language](#) (148)
 - [Machine Learning](#) (200)
 - [Machine Vision](#) (62)
 - [Neural Nets](#) (242)
 - [Robotics](#) (206)
 - [Speech](#) (26)
 - [Statistical Models](#) (118)
 - [Electronic Publishing](#) (87)
 - [Archives](#) (61)
 - [Copyright](#) (14)
 - [Economics](#) (33)
 - [Peer Review](#) (19)
 - [JOURNALS](#) (406)
 - [Behavioral & Brain Sciences](#) (19)
 - [Calicut Medical Journal](#) (9)
 - [Indian Pacing and Electrophysiology Journal](#) (165)
 - [Internet Health](#) (2)

Figure 40 - extract from <http://Cogprints.org/view/subjects/>

Not all the subject terms used in Cogprints are present in the OPAC; perhaps they are too specific.

The figures below show the results of a subject search for "applied cognitive psychology" on the University of Southampton Libraries OPAC and a search for a more specialised Subject heading ("numismatics") across both e-Prints and the OPAC returned 69 hits yet in e-Prints only 1.

UNIVERSITY OF Southampton University of Southampton Libraries WebCat

My Account Search/Home Other Catalogues Reser

Go Back Help New Search Previous Next Logout

subject "applied cognitive psychology" found no matches in any library.
 You may change your search, or select a new search from the closest matches below.

Change your search words	
APPLIED COASTAL RESEARCH	1
Applied cost-benefit analysis	1
Applied ecology	5
Applied ethics.	7
Applied ethics--Encyclopedias.	1
Applied ethics--Study and teaching.	1
APPLIED GEOLOGY	1
Applied human geography.	2
Applied kinesiology.	2
Applied linguistics	42
Applied linguistics--Addresses, essays, lectures.	4
Applied linguistics--Africa.	1
Applied linguistics--Congresses.	3

Go Back Help New Search Previous Next Logout

Figure 41 - Example of the results of a subject search for "applied cognitive psychology".

The screenshot shows the University of Southampton Libraries WebCat interface. At the top, the university logo and 'University of Southampton Libraries WebCat' are visible. A navigation bar includes links for 'My Account', 'Search/Home', 'Other Catalogues', and 'Reserve Collection'. Below this is a search results page for the subject 'numismatics', which has found 69 titles. The results are displayed in a list format with five entries visible. Each entry includes a call number, title, author, and location information, along with 'Details' and 'Mark' buttons.

#	Call Number	Title	Author	Year	Location
#1	CD 5363 KRZ	Aegean seals : an introduction	Krzyszowska, O.	2005	1 copy at Hartley.
#2	quarto CD 5363 PAL	Aegean seals, sealings and administration : proceedings of the NEH-Dickson Conference of the Program in Aegean Scripts and Prehistory of the Department of Classics, University of Texas at Austin, January 11-13, 1989	Palaima, Thomas G.	1990	1 copy at Hartley.
#3	quarto DT 268.S4 HAS	Africa proconsularis : regional studies in the Segermes Valley of Northern Tunisia	Orsted, Peter.	2000	1 copy at Hartley.
#4	quarto DT 268.S4 HAS v.1	Africa proconsularis : regional studies in the Segermes Valley of Northern Tunisia	Dietz, Søren.	1995	2 copies at Hartley.
#5	DB 920 v.1	Alterthümer des frühen Mittelalters in Ungarn	Hampel, József, 1849-1913.	1971	3 copies at Hartley.

Figure 42 - Search results for the subject heading "numismatics" on Southampton OPAC

As part of a JISC project, **EdShare** has made sure to use standard subject schemes. It is as much a VLE (Virtual Learning Environment) as it is a conventional IR. Collections can be browsed using:

- JACS Subjects
- University School or Unit
- Course Modules
- Keywords

This seems extremely useful and well thought-out, catering clearly for a designated community. The following illustrations show some of the ways in which these menus appear to end users:

The screenshot shows the EdShare website interface. At the top left is the EdShare logo (University of Southampton). To the right is a search bar with the text 'Enter your search...' and a 'Search' button. Below the logo is a 'User Tools' sidebar with links: 'Create a share', 'Create a collection', 'Bulk upload', 'EdShare Organiser (beta)', 'Latest additions', and 'Tag cloud'. Below that is a 'Send us your feedback!' section with links: 'Fill-in our online survey' and 'Email us'. The main content area is titled 'Browse by JACS Subjects' and contains a list of subjects with their counts in parentheses: A - Medicine and Dentistry, B - Subjects allied to Medicine, C - Biological Sciences, D - Veterinary Sciences, Agriculture and related subjects, F - Physical Sciences, G - Mathematical and Computer Sciences (1), H - Engineering, J - Technologies (1), K - Architecture, Building and Planning, L - Social studies, M - Law, N - Business and Administrative studies, P - Mass Communications and Documentation, Q - Linguistics, Classics and related subjects, R - European Languages, Literature and related subjects, T - Eastern, Asiatic, African, American and Australasian Languages, Literature and related subjects, V - Historical and Philosophical studies, W - Creative Arts and Design (2), X - Education (1). Below the list, it says 'Number of items at this level: 0.' At the bottom, there are links for 'Help - Take-down Policy - Terms and Conditions - Guide for adding content to EdShare', contact information 'edshare@soton.ac.uk', copyright notice '© 2008-2009 University of Southampton', and production credits 'Edshare was produced by the EdSpace Project and JISC'.

Figure 43 – browsing EdShare by JACS subject groupings.

The screenshot shows the EdShare website interface. At the top left is the EdShare logo (University of Southampton). To the right is a search bar with the text 'Enter your search...' and a 'Search' button. Below the logo is a 'User Tools' sidebar with links: 'Create a share', 'Create a collection', 'Bulk upload', 'EdShare Organiser (beta)', 'Latest additions', and 'Tag cloud'. Below that is a 'Send us your feedback!' section with links: 'Fill-in our online survey' and 'Email us'. The main content area is titled 'Browse by University School or Unit' and contains the text 'Please select a value to browse from the list below.' followed by a list of university units with their counts in parentheses: University Structure (1321), Faculty of Engineering, Science and Mathematics (794), School of Chemistry (14), School of Electronics and Computer Science (149), School of Engineering Sciences (2), School of Geography (13), School of Mathematics (606), School of Ocean & Earth Science (SOC/SOES) (2), School of Physics and Astronomy (3), Faculty of Law, Arts and Social Sciences (319), School of Art (15), School of Education (2), School of Humanities (292), School of Management (2), School of Social Sciences (8), Faculty of Medicine, Health and Life Sciences (53), School of Biological Sciences (1), School of Health Professions and Rehabilitation Sciences (1), School of Health Sciences (1), School of Medicine (40), School of Psychology (10), Professional Services (155), iSolutions (17), Learning and Teaching Enhancement Unit (LATEU) (20), University Library (118). Below the list, there are links for 'Help - Take-down Policy - Terms and Conditions - Guide for adding content to EdShare', contact information 'edshare@soton.ac.uk', copyright notice '© 2008-2009 University of Southampton', and production credits 'Edshare was produced by the EdSpace Project and JISC'.

Figure 44 – browsing EdShare by University School or Unit.

7.5 University of Strathclyde

An interesting example here is provided by the SFX generated "Electronic Journals" subject list, (http://ensor.lib.strath.ac.uk:3210/sfxlcl3/azlist/default?¶m_perform_value=searchCategories) visible when making use of the Andersonian Library's SFX linking services. These broad headings and sub-headings relate to journal *titles* and correspond to journals available via certain subscription

database packages such as ABI Inform. Users have to search separately within the Library's Electronic Database Services to find *articles* on a specific subject. The "Ejournals by Subject" pages include journals recorded in the library catalogue (i.e. NOT those where access derives from database aggregation services). These "Ejournals by subject" lists use controlled library terms and direct users to SFX activated abstracting databases and various other relevant subject-based resources.



Ejournal lists (SFX generated) - by Title, Subject or Search [SFX?]

Use these broad SFX subject headings to find individual journal **titles**. To search for journal **articles** Use the CTRL key to choose more than one subheading and the Shift key to choose all subheadings

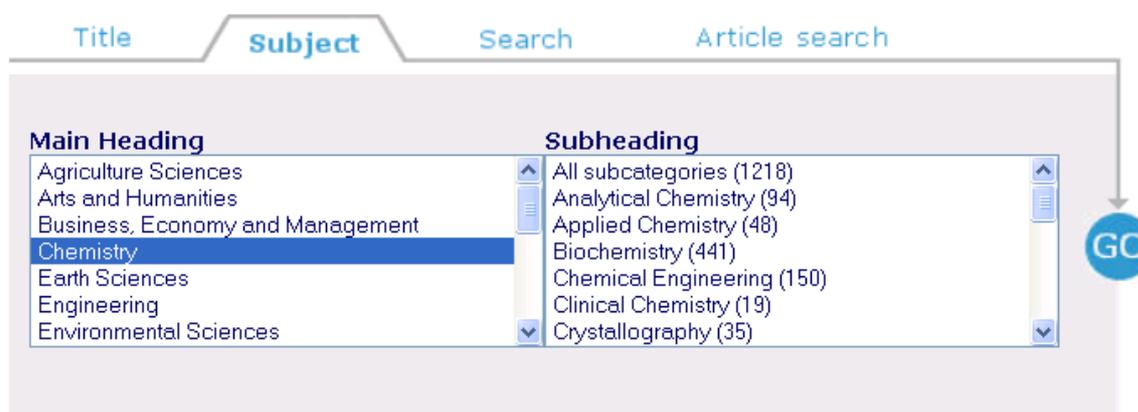


Figure 45 - partial screenshot of the SFX-generated subjects and sub-headings used for searching aggregated journal packages.

Unsurprisingly then, the SFX generated Electronic Journal headings do not match the Subject Headings being utilised by the OPAC or the IR. **Strathprints** uses LCC (the top level of which is supplied with EPrints 3) however, some more specific, lower level headings (around 8 -10) have been added by IR staff to reflect the granularity of the content deposited.

An SFX generated Subject Heading might be "Library and Information Sciences" while **Strathprints** would encode this subject as both "Library Science. Information Science" and "Libraries (General)". SFX records the following subheadings:

- All subcategories
- Abstracting and Indexing
- Acquisition
- Archives & Preservation
- Cataloging
- Classification
- Digital Libraries
- General and Others
- Information Sources, Services and Retrieval
- Librarianship and Libraries
- Library Automation
- Library Management
- Library Networks
- Serials Control

Figure 46 - Examples of SFX sub-headings for "Library and Information Sciences."

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

Returning to the example for the author Kevin D. O'Gorman (presented in Section 4.2) a subject is not given for the thesis in the OPAC. However, in the IR, a variety of subject headings are used, again, depending on the format you wish to export the record in. The main entry is shown below:

```
Item Type:Thesis (PhD)
ID Code:7645
Keywords:hospitality industry; history; classical antiquity; hospitality;
tourism;
Subjects:D History General and Old World > DE The Mediterranean Region.
The Greco-Roman World
H Social Sciences > HF Commerce
B Philosophy. Psychology. Religion > B Philosophy (General)
C Auxiliary Sciences of History > CB History of
civilizationDepartment:Strathclyde Business School > Hospitality and
Tourism Management
Deposited By:Dr Kevin O'Gorman
Deposited On:05 Mar 2009 12:26
Last Modified:05 Mar 2009 12:26
```

The keywords assigned to this record for the metadata formats supported are shown below:

The keywords assigned to this record in BibTeX format are:

hospitality industry; history; classical antiquity; hospitality; tourism

In Dublin Core the Subject is given as:

subject: DE The Mediterranean Region. The Greco-Roman World

subject: HF Commerce

subject: B Philosophy (General)

subject: CB History of civilization

In EP3 XML, where 'Family Name' and 'Given Name' tags are used for author, the Subjects are referred to only by their numeric encoding:

```
<subjects>
<item>DE</item>
<item>HF</item>
<item>B1</item>
<item>CB</item>
</subjects>
```

In EndNote format only keywords are given:

hospitality industry; history; classical antiquity; hospitality; tourism

In JSON, the numeric encoding is used again, for Subjects and for Department:

```
"subjects": [
  "DE",
  "HF",
  "B1",
  "CB"
],
"divisions": [
  "13314"
```

In METS the Subjects are:

```
</mods:abstract>
```

```
<mods:classification authority="lcc">DE The Mediterranean Region. The Greco-Roman
World</mods:classification>
```

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

```
<mods:classification authority="lcc">HF Commerce</mods:classification>  
<mods:classification authority="lcc">B Philosophy (General)</mods:classification>  
<mods:classification authority="lcc">CB History of civilization</mods:classification>
```

—
<mods:originInfo>

In Refer the keywords again are given instead:

%K hospitality industry; history; classical antiquity; hospitality; tourism

And finally, in Reference Manager:

KW - hospitality industry; history; classical antiquity; hospitality; tourism

Figure 47 – the different ways in which keywords are chosen and assigned in Strathclyde's EPrints IR.

8 Resource Discovery Platforms

Vertical/specialised search services are marketed more to the LMS community by vendors but are being equally considered for IR development.

Many respondents to both the IR and LMS/OPAC OCRIS questionnaires indicated that integrated RDP services were either in place or pending. The installation at Strathclyde provides a typical (and local) example of what these products might achieve and how they function.

8.1 University of Strathclyde

At the University of Strathclyde, the aforementioned thesis by Kevin D. O'Gorman's is now findable in both print and electronic form via the newly installed **Primo Resource Discovery Platform (SUPrimo)**.

Fortunately, the Strathprints cover sheet attached to deposits gives full citation details and alerts users to the source from which the item is being retrieved. SUPrimo's navigation menus list Strathprints items returned in search results as belonging to "Other Collections" (i.e. collections not entirely controlled by the library) rather than referring to the Strathprints repository by name.

Using the author's name, or the title of his thesis, as search terms, records for both the print and electronic versions of his work are returned, as shown below. The record for the electronic version links the end user through to the full text available from the **Strathprints IR** (<http://strathprints.ac.uk/7645/>). SUPrimo, however, doesn't search subscription electronic databases such as Web of Science.

The screenshot displays the SUPrimo search interface. At the top, there are tabs for "Library Collections" and "Exam Papers". Below these is a search bar containing the text "o'gorman kevin d." and a "GO" button. Underneath the search bar, there is a small text box with search tips: "For phrases use quotes, 'global warming' - truncated searches, use * eg, optic* Class codes: no spaces, eg. 30345 or MA110 - n.b. classes or names may change." Below this are links for "New Search" and "Advanced Search".

The results section is titled "Results for: SU all" and shows "1 Results, sorted by: relevance". A dropdown menu is set to "relevance" and a link "Did you mean: gorman kevin d?" is provided. The first result is a book:

- The essence of hospitality from the texts of classical antiquity : the development of a hermeneutical helix to identify the philosophy of the phenomenon of hospitality** (View details)
- Kevin D. O'Gorman University of Strathclyde. Dept. of Hospitality and Tourism Management
Thesis [Ph. D.] -- University of Strathclyde, 2008. 2008
- more There are 2 versions of this item
- Add to e-Shelf Available (find item)

Figure 48 – Strathclyde's new Primo implementation, "SUPrimo" retrieves the two version of O'Gorman's Ph.D thesis.

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

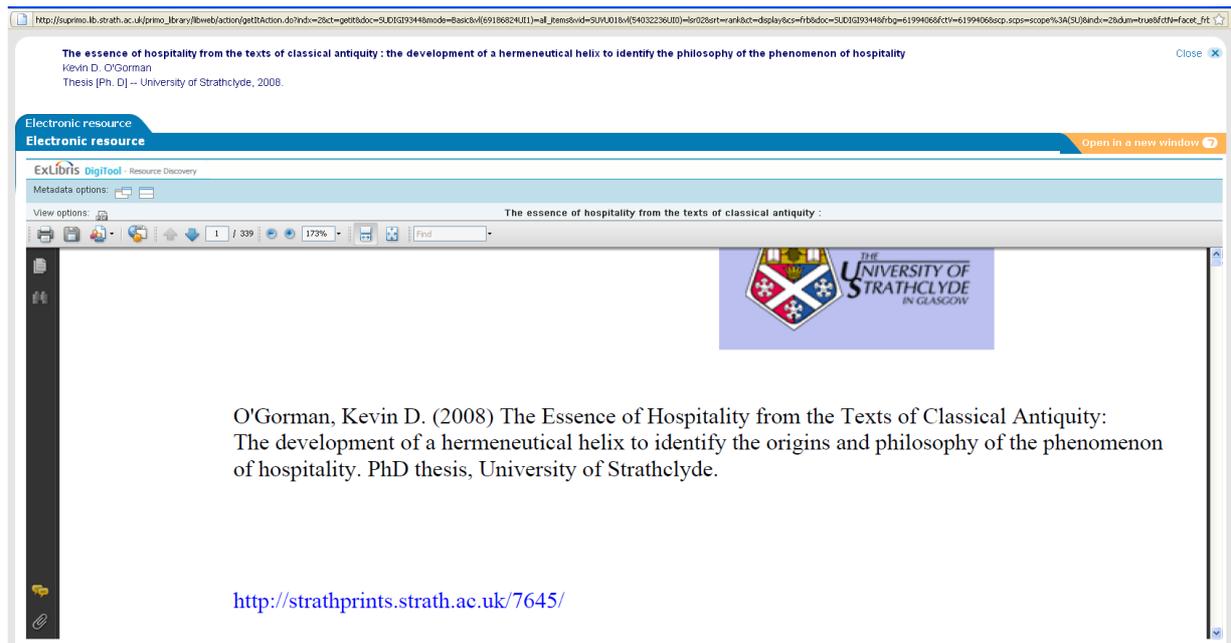


Figure 49 – SUPrino links to the Strathprints version of a PhD thesis.

9 Three examples from the Welsh Repository Network (WRN)

As part of its investigations into the more detailed aspects of scope duplication across IRs and OPACs, OCRIS sought to address the issue of a national repository network with requirements for interfaces and records to be made available to end users in a national language other than English and the implications this may hold for the future interoperability of these IRs with the library catalogues of the Universities they serve.

A brief survey was therefore undertaken of 3 HEIs within the Welsh Repository Network (WRN)¹⁷: the Universities of Aberystwyth, Cardiff and Glamorgan.

Objectives a) and e) of the WRN "Start-up Project"¹⁸ were to establish:

- a) An operational open access digital repository in every Welsh HEI that forms an integral part of the learning, teaching, research and administrative fabric of its host institution.
- and
- e) A national resource discovery and search service which provides integrated access to aggregated data of Welsh interest and which also makes provision for long-term preservation and archiving.

Clearly there is a requirement for institutions to seek the retrospective deposit in their IRs of recently produced articles and PhD theses if these goals are to be achieved.

Objective c) of the Start-up Project announced that a principal component, to aid end users to access the resources of the WRN would be "A bilingual interface which is fully compliant with the requirements of the Welsh Language Act" (WRN, Final Report p. 4).

Of specific interest to OCRIS, therefore, was the necessity for WRN IRs to support both bilingual interfaces *and* metadata (for the provision of integrated access) and the impact this may have upon the comprehensive harvesting of metadata by national/international aggregation services and the compilation of accurate publications lists of research outputs; raising the issue of what may be lost or gained if catalogue records remain 'un-translated' between systems and services.

The Welsh Repository Network, *Final Report*¹⁹ specifies the technology and standards that the WRN will use:

"All repositories will be installed using standard-compliant repository software. At a minimum this is to expose records expressed in simple Dublin Core via OAI-PMH. Other standards such as RSS or ATOM feeds will be provided where supported by the software. If possible, new JISC-funded standards such as the Scholarly Works Application Profile (eprints application profile) and Simple Web-service Offering Repository Deposit (SWORD) will be included" (Welsh Repository Network, *Final Report*, p. 2).

9.1 University of Aberystwyth

In considering repositories which fulfil a socio-political and even Nationalist purpose, certain provisions may be made regarding scope. CADAIR (the IR at Aberystwyth) states that in addition to collecting "cutting-edge peer-reviewed" research outputs:

¹⁷ <http://www.jisc.ac.uk/whatwedo/programmes/reppres/sue/welshrepositorynetwork.aspx> [Accessed 28th August 2009].

¹⁸ <http://www.whelf.ac.uk/wrn/> [Accessed 28th August 2009].

¹⁹ <http://www.jisc.ac.uk/whatwedo/programmes/reppres/sue/welshrepositorynetwork.aspx> [Accessed 28th August 2009].

The following categories of successful theses and dissertations may be deposited in CADAIR in addition to the two paper copies in the library:

- Successful research theses (PhD and MPhil) in their final corrected form
- Postgraduate taught masters' dissertations (MA and MSc)
 - which receive distinctions (whether or not a student gets a distinction overall)
 - on the subjects of Welsh, Celtic Studies and Welsh History

Figure 50 - extract from scoping statement on the CADAIR IR.

This would seem to suggest that the quality criteria for dissertations covering subjects of particular relevance to aspects of Welsh national identity are less stringent (as it does not state that PG dissertations on these topics must receive distinctions).

The homepage allows you to move between Welsh and English and, as illustrated in the figure below, content can be browsed according to department, the names of which are provided in *both* English and Welsh within the same interface. Regardless of which interface you are using, you can use English or Welsh words as search terms, as translations to Welsh are applied to Dublin Core descriptors such as "title/teitl " "author/awduron" etc. This is illustrated below:

Teitl: Painting the Rock: Images of Uluru by Michael Andrews and Lloyd Rees
Awduron: [Pierse, Simon](#)
Dyddiad Cyhoeddi: 2004
Cyhoeddwr: Australian Studies, The Journal of the British Australian Studies Association
Dyfyniad: Pierse, Simon, 'Painting the Rock: Images of Uluru by Michael Andrews and Lloyd Rees', Australian Studies, The Journal of the British Australian Studies Association (2004) 17(1) pp.167-188
Cyfeirnodwyd gan: <http://www.kcl.ac.uk/schools/humanities/depts/menzies/basa/journal.html>
DCU: <http://hdl.handle.net/2160/904>
ISBN: 0954-0954
Ymddengys yng Nghasgliadau: [School of Art](#)

Ffeiliau yn yr Eitem Hon:
Nid oes ffeiliau yn gysylltiedig â'r eitem hon.

[Dangos cofnod eitem llawn](#)
[Allforio i EndNote](#)

Figure 51 - Metadata fields translated into Welsh, with the text contained in them un-translated.

Shown below is a list of Departments. Click on a name to view the departmental home page.

- ❖ **[Addysg a Dysgu Gydol Oes / Education & Lifelong Learning \[2\]](#)**
 - ❖ [PhD theses from the School of Education and Lifelong Learning \[2\]](#)

- ❖ **[Astudiaethau Gwybodaeth / Information Studies \[109\]](#)**
 - ❖ [Archives and Records Management \[6\]](#)
 - ❖ [Health and Information Systems \[28\]](#)
 - ❖ [Information Behaviour and e-learning \[30\]](#)
 - ❖ [Information Literacy \[8\]](#)
 - ❖ [MSc dissertations from Department of Information Studies \[2\]](#)
 - ❖ [Management and Knowledge Management \(for Library and Information Services\) \[28\]](#)
 - ❖ [PhD theses from Department of Information Studies \[2\]](#)
 - ❖ [Publishing and e-Publishing \[5\]](#)

- ❖ **[Astudiaethau Theatr, Ffilm a Theledu / Theatre, Film & Television Studies \[136\]](#)**
 - ❖ [Drama and Theatre Research \[25\]](#)
 - ❖ [Film and Television Research \[36\]](#)
 - ❖ [Lord Of The Rings research \[6\]](#)
 - ❖ [Performance Research \[31\]](#)
 - ❖ [PhD theses from the Department of Theatre, Film & Television Studies \[2\]](#)
 - ❖ [Production Photography \[2\]](#)
 - ❖ [Welsh Theatre and Media Research \[26\]](#)
 - ❖ **[Centre for Performance Research \[8\]](#)**
 - ❖ [CPR Interviews \[3\]](#)
 - ❖ [Items relating to the Centre for Performance Research \[5\]](#)

- ❖ **[Canolfan Iaith a Dysgu / Language and Learning Centre \[34\]](#)**
 - ❖ [Journal Articles / Papers \[34\]](#)

- ❖ **[Celf / Art \[35\]](#)**
 - ❖ [School of Art \[35\]](#)

- ❖ **[Cyfrifiadureg / Computer Science \[387\]](#)**
 - ❖ [Advanced Reasoning Group \[139\]](#)
 - ❖ [Computational Biology Group and Wolfson Bioinformatics Unit \[71\]](#)
 - ❖ [Intelligent Robotics Group \[122\]](#)
 - ❖ [PhD theses from Department of Computer Science \[3\]](#)
 - ❖ [Software Engineering \[15\]](#)
 - ❖ [Vision, Graphics and Visualisation Group \[35\]](#)
 - ❖ **[ART Project \[2\]](#)**

Figure 52 - Consistent use of department names in both Welsh and English.

Neither the OPAC nor the IR explicitly list "Theses" as being in scope of their catalogues, (CADAIR opting to include them within "Other major forms of academic research content").

One example of theses duplication can be seen below:

The screenshot shows the Aberystwyth University OPAC interface. At the top, there is a navigation bar with links for Search, My Searches, My Bibliography, My Account, Reading Lists, New Books, eBooks, and eLibrary. Below this is a search bar with a 'Go' button and a 'Search History' link. The main content area displays the following information for a thesis record:

Anoraciaeth yr acen : agweddau ar gerdd dafod gyfoes /

Title: Anoraciaeth yr acen : agweddau ar gerdd dafod gyfoes / Iwan Rhys.
Author: [Rhys, Iwan](#)
Publisher: Aberystwyth, 2008.
Format: Book
Note: Traethawd (M.Phil.) - Prifysgol Cymru.

Holdings Information

Location: Hugh Owen Library: Issue Desk Store
Classmark: AC1.W1.2762
Number of Items: 1
Status: Available

Availability:	Loan type	Location	Classmark	Volume	Status	Due date
	Thesis	Hugh Owen Library: Issue Desk Store	AC1.W1.2762	—	Available	—

Additional links: [Send by SMS](#) [QR code](#)

Figure 53 - a record for a thesis in the Aberystwyth OPAC.

The screenshot shows the Aberystwyth University CADAIR Online Research Repository record for the same thesis. The header includes the Aberystwyth University logo and the CADAIR logo. The main content area displays the following information:

Title: Anoraciaeth yr Acen: Agweddau ar Gerdd Dafod Gyfoes
Authors: [Rhys, Iwan](#)
Supervisor: [Edwards, Huw Meirion](#)
Issue Date: 2009
Publisher: Aberystwyth University

Abstract: Astudiaeth feintiol ar rai o agweddau technegol Gerdd Dafod yw hon. Yn seiliedig ar erthygl gan Gwynfor ab Ifor yn Barddas, mae'r bennod gyntaf yn cymharu arddull gynganeddol awdlau arobryn tri chyfnod gwahanol yn yr ugeinfed ganrif a dechrau'r unfed ganrif ar hugain ac yn dangos y gellir rhannu'r awdlau yn ddau ddsbarth penodol. Gan fynd â'r syniad o wahaniaethau mewn arddull gynganeddol ymhellach, mae'r ail bennod yn cymharu sampl o gywyddau gan bedwar cywyddwr cyfoes ac yn llunio 'peiriant' i geisio dyfalu awduraeth cywydd gan un o'r pedwar bardd ar sail arddull gynganeddol yn unig. Llunias 'Holiadur y Cynganeddwyr' i'w lenwi gan ymarferwyr y grefft, a chanlyniadau'r holiadur yw'r sail ar gyfer penodau tri a phedwar. Mae'r gyntaf o'r ddwy bennod yn disgrifio'r broses o lunio'r holiadur ac yn dadansoddi'r cwestynau cefndirol ac mae'r bennod olaf yn dadansoddi'r ymatebion i brif gorff yr holiadur, sy'n holi barn cynganeddwyr am amryw bynciau technegol a fu'n destun trafod dros y blynyddoedd diweddar, ac yn cymharu'r canlyniadau gyda chanllawiau rhai o'r llawlyfrau mwyaf poblogaidd ar gynganeddol.

URI: <http://hdl.handle.net/2160/1872>
Sponsor: Aberystwyth University
Appears in Collections: [Traethodau MPhil, Cymraeg / MPhil theses, Welsh](#)
[MPhil theses from Aberystwyth University](#)

Files in This Item:

File	Description	Size	Format	
f.doc		548 kB	Microsoft Word	View/Open
s.doc		100.5 kB	Microsoft Word	View/Open
d.doc		424.5 kB	Microsoft Word	View/Open
c.doc		277.5 kB	Microsoft Word	View/Open
b.doc		41.5 kB	Microsoft Word	View/Open
a.doc		23.5 kB	Microsoft Word	View/Open

[Show full item record](#)
[Export to EndNote](#)

Figure 54 - a record for the thesis - available as full text - in the Aberystwyth IR.

Yet links are not provided from IR to OPAC or vice versa. This would be extremely useful as many items in CADAIR are access restricted due to the assertion of rights by publishers or other copyright issues.

With specific reference to bilingual issues, at the time of writing there are 81 theses held or detailed in the institutional repository at Aberystwyth. Of these, 3 are MPhil theses, 52 are PhD theses, and 25 Taught Masters' dissertations. There are 3 Welsh language entries listed in the MPhil theses collection in the IR. Selecting the "Full metadata record" for one of these reveals that the item is

catalogued in its language of origin (Welsh), yet translation into English in the **dc.description.abstract** metadata field is not provided.

Title: Dau o weithiau Robert Holland (c. 1556/7-1622): Ymddiddan Tudur a Gronw a'r Basilikon Doron

Authors: [Morgan, Adrian](#)

Full metadata record

DC Field	Value	Language
dc.contributor.advisor	Williams, Gruffydd Aled	-
dc.contributor.author	Morgan, Adrian	-
dc.date.accessioned	2008-02-26T14:44:12Z	-
dc.date.available	2008-02-26T14:44:12Z	-
dc.date.issued	2008	-
dc.identifier.uri	http://hdl.handle.net/2160/512	-
dc.description.abstract	Yn y traethawd hwn ymdrinnir â dau o brif weithiau Robert Holland (c. 1556/7-1622), sef ei Ymddiddan Tudur a Gronw a gyhoeddwyd yn Rhydychen yn 1600 a'i gyfieithiad Cymraeg o Basilikon Doron y brenin Iago a gyhoeddwyd yn 1604. Yn y bennod gyntaf trafodir y manylion bywgraffyddol sy'n hysbys am Robert Holland ac ymdrinnir â'i gyfnod fel myfyriwr ym Mhrifysgol Caer-grawnt gan nodi nifer o unigolion a allai fod wedi dylanwadu ar yr efrydydd ifanc. Yna, trafodir ei safbwynt diwinyddol a cheisir olrhain ei yrfa fel offeiriad yn Eglwys Loegr. Ceisir crynhoi ei gysylltiadau diddorol yn sir Benfro a thu hwnt a chroniclir ei yrfa lenyddol hefyd. Yn yr ail bennod rhoddir sylw i'w draethawd yn erbyn dewiniaeth—sef Ymddiddan Tudur a Gronw—sydd ar ffurf ymddiddan rhwng dau gymeriad. Ymdrinnir â'r testun ei hun a cheisir crynhoi manylion ei argraffu. Ar sail tystiolaeth a ganfuwyd yn y Y Ffydd Ddiffuant gan Charles Edwards, dadleuir i'r testun gael ei gyhoeddi yn ystod y flwyddyn 1600. Creffir ar ddiddordeb mynych y cyfnod hwn mewn dewindabaeth a cheisir crynhoi ymateb y Cymry i'r mania hwn. Yna, ymdrinnir â chynnwys y testun. Mae'r atodiad sydd ynghlwm wrth y traethawd hwn yn gymhariaeth o'r fersiwn llawysgrif (sydd i'w gael yn Llsgr. Cwrtmawr 114B) a'r fersiwn printiedig a olygwyd gan Stephen Hughes yn 1681. Yn y drydedd bennod ymdrinnir â chefnidir y testun	cy

```

                                gwreiddiol a rhoddir sylw i
                                amgylchiadau ei gyhoeddi a'i
                                argraffu. Trafodir cynnwys y testun
                                yn fanwl a chreffir ar amgylchiadau
                                cyhoeddi'r fersiwn Cymraeg. Cynigir
                                nifer o resymau dros ei gyfieithu ac
                                ymdrinnir yn fras ag ansawdd y
                                cyfieithiad.

dc.format.extent 28672 bytes -
dc.format.extent 39424 bytes -
dc.format.extent 200192 bytes -
dc.format.extent 279040 bytes -
dc.format.extent 426496 bytes -
dc.format.extent 64512 bytes -
dc.format.extent 53248 bytes -
dc.format.extent 94720 bytes -
dc.format.mimetype application/msword -
dc.language.iso cy en
  dc.publisher University of Wales, Aberystwyth en
  dc.publisher Prifysgol Cymru Aberystwyth cy
    dc.title Dau o weithiau Robert Holland (c. en
           1556/7-1622): Ymddiddan Tudur a Gronw
           a'r Basilikon Doron
      dc.type Text en
dc.publisher.department Welsh en
dc.type.qualificationlevel research masters en
dc.type.qualificationname MPhil en
dc.type.publicationtype thesis or dissertation en
  Appears in Collections: Traethodau MPhil, Cymraeg / MPhil
                           theses, Welsh
                           MPhil theses from Aberystwyth
                           University

```

Figure 55 - a record from CADAIR with the abstract in Welsh but not English.

The approach of CADAIR here is helpful for Welsh language users, with catalogue records also being made available in Welsh. However, translations for end users are provided at the level of individual search/encoding fields *within* the metadata, not between records. The example above is from the Welsh language “Cymraeg” interface, but the English language interface carries the same Welsh language metadata for the “dc.description.abstract field”. Although integrated, the catalogue record (for this item at least) is not accessible in a distinct translated versions, presumably down to the resources available for such labour intensive cataloguing workflows.

CADAIR does integrate links to Social Networking and Social Bookmarking sites, such as “Facebook”, “del.icio.us”, “StumbleUpon”, “Digg” and “reddit” showing an awareness that its records are being accessed and re-used by wider linguistic communities.

9.2 University of Cardiff

The first point that might be made regarding **ORCA (Online Research @Cardiff)** and the **Voyager Library Catalogue** as Cardiff is that while the IR is presented, described, and catalogued in English only, the OPAC offers a Welsh interface offering translations of significant search fields.

Enw'r gronfa ddata: Cardiff University and Wales NHS Trust Libraries

The screenshot shows a search interface with a dark red header. Below the header are four tabs: 'Chwiliad Syml', 'Chwiliad Manwl', 'Rhestrau darllen', and 'Eitemau Newydd'. The main search area has a search bar containing 'celf' and a 'Terfynau:' dropdown menu set to 'Dim'. To the right is a 'Chwilio yn ôl:' dropdown menu with options: 'Keyword Anywhere', 'Author Browse', 'Title Begins With', 'Classmark Browse', 'Subject Browse', 'Journal Title Begins With', and 'Journal Title Keyword'. At the bottom, there is a '50 cofnod y dudalen' dropdown menu and two buttons: 'Chwilio' and 'Allosod'. Below the search area is a red banner with the text 'Opsiyau Ychwanegol ar gyfer Chwiliad Allweddair:' and a dropdown menu for 'Chwilio trwy'r catalog am gofnodion sydd yn cynnwys POB UN o'r termau hyn'.

Figure 56 - The bi-lingual (English-Welsh) Voyager Library Catalogue search interface at the University of Cardiff. Only certain fields can be viewed in either language.

Again though, only certain fields (usually at "top level") are translated into Welsh; many others used to structure searches (such as "keyword anywhere" or "title") are presented only in English. It would be interesting to investigate this further, as it may be that there are problems with making certain fields Welsh, which has implications for mapping and semantic interoperability and data exchange.

The translation of the University *name* into/from Welsh appears on the website but this process of translation does not apply to Cardiff's own IR or to *the records of any of its scholarly resources therein*, although (of course) users can search for Welsh language artefacts.

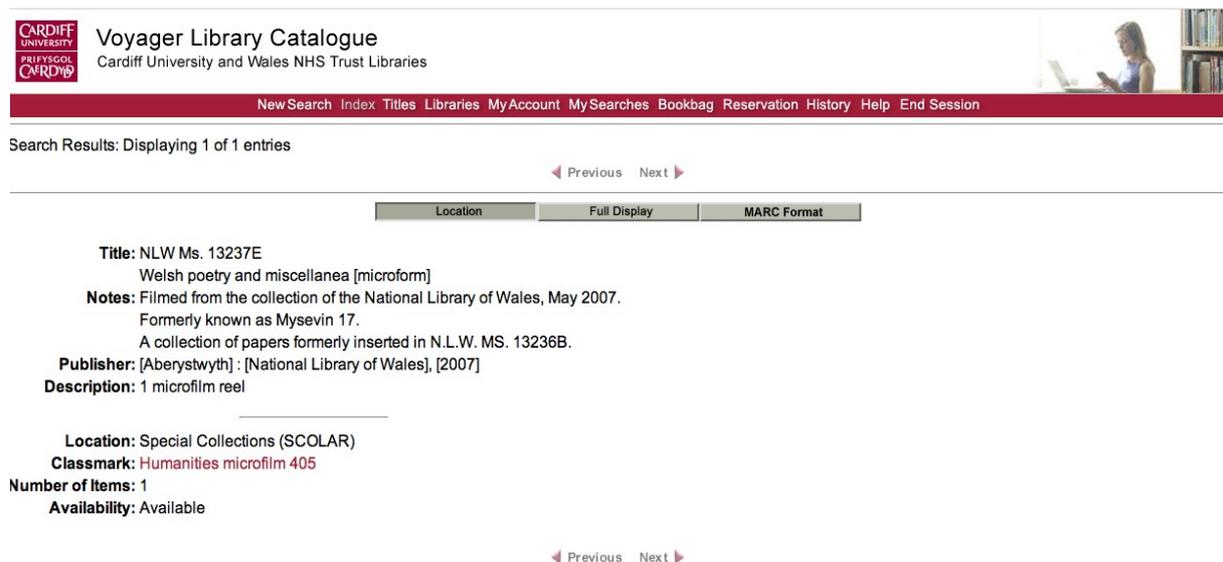
Given the importance of the Welsh national language in the nation's education system, the question remains as to whether such a discrepancy between the IR and OPAC suggested an assumption of intrinsically different user communities for - or at least, a lack of shared agendas and workflows between those administering - each system.

The only information on item scope in **OCRA** is that it is "a digital repository of research publications" and that it "is not intended to be a substitute for peer-reviewed journals, but has been developed to host work by Cardiff University researchers that has already been, or will be, published elsewhere." Although the advanced search menu lists item types these are standard EPrints/Dublin Core fields (<http://orca.cf.ac.uk/cgi/search/advanced>) - a *specific* statement on OCRA's scope is not provided.

The OPAC lists 8 item types where the IR lists 18. Other than "Book" there is no other clear, unambiguous "Item" search that could be repeated across both IR and OPAC – and "book" itself is not altogether without debate.

As objective e) of the WRN Start Up Project advocated "A national resource discovery and search service which provides integrated access to aggregated data of Welsh interest", a simple search was undertaken for books of, or about, poetry in Welsh.

Selecting the "set limits" facility and setting the language limiter to "Welsh", the item type limiter to "Book", and placing "Welsh poetry" in the "Title Begins With" search field, 10 records are returned, 1 of which is for a microfilm reel of "Welsh poetry and miscellanea" published by the National Library of Wales in 2007:



The screenshot shows the Voyager Library Catalogue interface. At the top, there is a header with the Cardiff University logo and the text 'Voyager Library Catalogue Cardiff University and Wales NHS Trust Libraries'. Below this is a navigation bar with links for 'New Search', 'Index', 'Titles', 'Libraries', 'My Account', 'My Searches', 'Bookbag', 'Reservation', 'History', 'Help', and 'End Session'. The main content area displays search results for 'Displaying 1 of 1 entries'. The record details include: Title: NLW Ms. 13237E, Welsh poetry and miscellanea [microform]; Notes: Filmed from the collection of the National Library of Wales, May 2007. Formerly known as Mysevin 17. A collection of papers formerly inserted in N.L.W. MS. 13236B.; Publisher: [Aberystwyth] : [National Library of Wales], [2007]; Description: 1 microfilm reel; Location: Special Collections (SCOLAR); Classmark: Humanities microfilm 405; Number of Items: 1; Availability: Available. Navigation arrows for 'Previous' and 'Next' are visible above and below the record details.

Figure 57 - OPAC record for microfilm reel of "Welsh poetry and miscellanea".

It is revealing that the MARC record contains bi-lingual descriptors, indicating one subject in Welsh (MARC field 650 0 |a Englyns) and another in English (MARC field 650 0 |a Welsh poetry).

This is illustrated below:

```
000 00797cam a22002415i 450
001 822583
005 20080225130645.0
008 071203s2007 wlk 000 m wel d
040 __ |a WlCaUW |c WlCaUW
098 00 |a Humanities microfilm 405
130 0_ |a NLW Ms. 13237E
245 10 |a Welsh poetry and miscellanea |h [microform]
246 33 |a LLGC 13237E
246 33 |a NLW 13237E
260 __ |a [Aberystwyth] : |b [National Library of Wales], |c [2007]
300 __ |a 1 microfilm reel
509 __ |a Filmed from the collection of the National Library of Wales, May
2007.
509 __ |a Formerly known as Mysevin 17.
509 __ |a A collection of papers formerly inserted in N.L.W. MS. 13236B.
650 _0 |a Welsh poetry
650 _0 |a Englyns
710 2_ |a National Library of Wales. |k Manuscript. |n 13237E
991 10 |d WELSH
```

Figure 58 - MARC record returned when searching for items in Welsh.

From this it appears that the locally defined 991 field is being used to allow "language" to be searched, rather than the language encoding in the 008/040 fields. However exploring other records presented in the search results list it is evident that in some cases it is the 509 or 549 fields which contain information on language, with 991 not used. It would be worth investigating further to discern whether this variation is because of different practices among the 18 different libraries contributing records or the fact that records have been imported from elsewhere and do not conform to Cardiff's internal standard (for instance). Whatever the case, there is a clue here about potential problems with offering interfaces based on language or attempting to text mine, when consistent standards or rules are not being applied.

9.3 University of Glamorgan

A search for "theses" limited by "PhD theses from the University of Glamorgan" from the search menu, retrieves 1 record within **Glamorgan DSpace**, for *Design and investigation of scalable multicast recursive protocols for wired and wireless ad hoc networks* by Firas Al-Balas:

12-Jun-2009	Design and investigation of scalable multicast recursive protocols for wired and wireless ad hoc networks	Al-Balas, Firas
-------------	---	---------------------------------

Figure 59 - partial record of a thesis held in Glamorgan's DSpace repository.

A search within the **OPAC** for the same author retrieves the record for the thesis there also.

The screenshot shows the OPAC interface with the following search criteria:

- Source: University of Glamorgan
- Collection: Whole Catalogue
- LRC Site: All
- Year: any year

Search results show 1 title matching the search:

Author Al-Balas, Firas.
Other Author(s) / Title(s) University of Glamorgan.
Title Design and investigation of scalable multicast recursive protocols for wired and wireless ad hoc networks.
Bib Id 552536
Dissertation Note Thesis (Ph.D.) - University of Glamorgan, Pontypridd, 2009.
Subject Computer networks.
Classification 004.6

There is 1 title matching your search. Number 1 of 1

Location	Shelfmark	Loan Type	Availability	Copies	On Loan
Treforest	Theses 004.6/ALB	Not for Loan	Available	1	0

Figure 13. Record for *Design and investigation of scalable multicast recursive protocols for wired and wireless ad hoc networks* by Al-Balas, Firas.

The full catalogue record is given below:

Author Al-Balas, Firas.

Project Acronym: OCRIS
 Document: Appendix 3 to OCRIS Final Report
 Version: 1
 Contact: Kathleen Menzies or Gordon Dunsire
 Date: September 2009

**Other Author(s)
 / Title(s)** University of Glamorgan.

Title Design and investigation of scalable multicast recursive protocols for wired and wireless ad hoc networks.

Bib Id 552536

Dissertation Note Thesis (Ph.D.) - University of Glamorgan, Pontypridd, 2009.

Subject Computer networks.

Classification 004.6

In this example, where there are two entry points to retrieve the same resource, there does appear to be clear scope duplication between selected IRs and OPACs within the WRN.

Browsing the repository by subject and selecting "artificial neural networks" retrieves a record for the following thesis, also listed in DSpace under "PhD theses from the University of Glamorgan."

Issue Date	Title	Author(s)
12-Jun-2009	Using artificial neural networks to forecast changes in national and regional price indices for the UK residential property market	Paris, Stuart David

Showing results 1 to 1 of 1

Figure 60 - record retrieved for a PhD thesis in Glamorgan's DSpace IR.

However an advanced search under the thesis title of Glamorgan's thesis collection for the current year 2009 in the OPAC retrieves no records, nor do any results match the author Stuart David Paris.

There is a question mark, therefore, on whether thesis deposits in the IR at Glamorgan are also being systematically recorded in the OPAC. The question over thesis deposit, however, also works in the opposite direction. For the year 2008, there were 19 theses catalogued in the OPAC, yet there are only 3 theses in total listed in the Glamorgan IR under "thesis collection".

9.4 Some conclusions

If the Welsh Repository Network is to achieve its stated aim of making its network of IRs integral to the fabric of learning, teaching and research at their host institutions, it must continue to address the issues of scoping, the visibility of theses and the provision of national language metadata. It will be interesting to see how the network develops.

The question also remains as to whether the Welsh language materials in IRs, which may be catalogued only partly in Welsh, can be adequately served by aggregator services and other national data banks. End users may retrieve records by browsing item types listed within their local repositories, or by knowing in advance what they are searching for (i.e. works by a given author, on a given subject or produced by a specific department such as the Welsh Department at Aberystwyth).

Questions remain as to whether the same records would be retrieved when undertaking similar searches where less is known in advance, or within the OPAC; and to whether the Abstract field will adequately support keyword extraction. For example, data mining tools might theoretically be used

within aggregator services (or virtual research environments) because there is no field provided to denote language.

If aggregator services assume that items are in English, they are unlikely to support searching by language or implement a "dc.language" field or LANG qualifiers for all records; this would, however, be advisable and could point toward the future incorporation of items in other non-English languages. Welsh uses the Latin alphabet so as a starting point, this should be easier to implement because support for foreign character sets would not be required.

The point is made, not to disparage a practice which advances the representation of national languages within the scholarly community, but to raise the issue of whether MPhil theses and dissertations output is being made sufficiently visible at the level of the national repository network through adequate record translation.

The consistent use of language fields would also increase the likelihood of LMS records and IR records being retrieved via advanced searching in ways most beneficial to bi-lingual users.

URLs for IRs and OPACs considered in the in-depth study

1. *University of Aberystwyth*

Institutional Repositories

<http://cadair.aber.ac.uk/dspace/>

OPACS

<http://voyager.aber.ac.uk/vwebv/searchBasic?sk=en>

2. *University of Cardiff*

Institutional Repositories

<http://eprints.cf.ac.uk/> / <http://orca.cardiff.ac.uk/> ORCA - Online Research @Cardiff

OPACS

<http://library.cf.ac.uk/>

3. *University of Cambridge*

Institutional Repositories

Computer Laboratory Technical Reports: <http://www.cl.cam.ac.uk/techreports/>

DSpace at Cambridge: <http://www.DSpace.cam.ac.uk/>

CUED Publications Database: <http://publications.eng.cam.ac.uk/>

OPACS

Newton: <http://www.lib.cam.ac.uk/newton/> and http://ul-newton.lib.cam.ac.uk/vwebv/searchBasic?sk=en_US

Janus: <http://janus.lib.cam.ac.uk/>

4. *University College for the Creative Arts*

Institutional Repositories

<http://vads.ac.uk>

OPACS

<http://librarycatalogue.ucreative.ac.uk/uhtbin/cgiirsi.exe/x/CANTERBURY/0/49>

5. *University of Glasgow*

Institutional Repositories

Enlighten <http://www.lib.gla.ac.uk/enlighten/>

Glasgow Theses Service: <http://theses.gla.ac.uk/>

Glasgow DSpace: <https://DSpace.gla.ac.uk/index.jsp>

OPACS

Encore (Resource Discovery Platform): <http://encore.lib.gla.ac.uk/iii/encore/app>

Classic catalogue: <http://eleanor.lib.gla.ac.uk/search/>

6. *University of Glamorgan*

Institutional Repositories

<http://dspace1.isd.glam.ac.uk/dspace>

OPACS

<http://opac.glam.ac.uk/TalisPrism/>

7. *University of Hull*

Institutional Repositories

<http://eDocs.hull.ac.uk/muradora>

OPACS

<http://library.hull.ac.uk/>

8. *National Marine Biological Library*

Institutional Repositories

<http://sabella.mba.ac.uk>

OPACS

http://www.mba.ac.uk/NMBL/collection/access_to_libweb.htm

<http://www.mba.ac.uk/isis/serlis/index.html>

9. *University of Southampton*

Institutional Repositories

<http://tprints.ecs.soton.ac.uk>

<http://Cogprints.org/>

<http://EPrints.soton.ac.uk/>

<http://www.edshare.soton.ac.uk/>

<http://ecrystals.chem.soton.ac.uk/>

<http://portfolio.ecs.soton.ac.uk/>

<http://EPrints.ecs.soton.ac.uk>

<http://www.6journal.org/>

<http://archive.serpentproject.com/>

<http://languagebox.EPrints.org>

Narrowing this list down to the ones most relevant to OCRIS (i.e. ones which are not so specialised that there may be overlap we get:

<http://EPrints.soton.ac.uk/>

<http://Cogprints.org/>

<http://EPrints.ecs.soton.ac.uk>

<http://www-lib.soton.ac.uk/>

<http://www.tdnet.com/soton>

OPACS

<http://www-lib.soton.ac.uk/>

10. *University of Strathclyde*

Project Acronym: OCRIS
Document: Appendix 3 to OCRIS Final Report
Version: 1
Contact: Kathleen Menzies or Gordon Dunsire
Date: September 2009

Institutional Repositories

<http://strathprints.strath.ac.uk>

OPACS and RDPs

<http://zen.lib.strath.ac.uk/> http://suprimo.lib.strath.ac.uk/primo_library/libweb/action/search.do?vid=SUVU01&fromLogin=true&fromLogin=true