JISC Information Environment Portal Activity: supporting the Needs of e-Research.
– **Web-based Library and Information Services** –

Rob Allan
CCLRC e-Science Centre, Daresbury Laboratory,
Daresbury, Warrington WA4 4AD

Rob Crouchley
Centre for e-Science, C Floor Bowland Annexe, Lancaster University, Lancaster LA1 4YT

Caroline Ingram
CSI Consultancy Ltd., 42 Coquet Terrace, Newcastle upon Tyne NE6 5LE

Contact e-Mail: `r.j.allan@dl.ac.uk`, `r.crouchley@lancs.ac.uk`,
`caroline@csiconsultancy.co.uk`

July 17, 2006

Abstract

This is a short information report on the kind of Web-based library and information services frequently used by researchers. It will be used as part of our user requirements and gap anaylsis and compared to services provided by the JISC IE.

Contents

1   Background

This document was produced from brainstorming and several "interviews" carried out in an unstructured way via e-mail and face to face with the people acknowledged.

The research process is about interpreting data, but also about going from resource discovery to knowledge.

The kind of resources accessed depends on the activity of the researcher. Experts are likely to use journal articles or conference proceedings to look for current results new to their field. On the other hand, students or established researchers learning new methodology or techniques are likely to use books.

We asked, *why do researchers use library information services?* The answers supported other input indicating that there are two types of search widely carried out:

- A research is looking for new information, perhaps because he/ she is new to a field or seeking to find out about prior work;

- A researcher is searching for a specific piece of information already known to exist, perhaps a publication which contains details of work by an author they know to be working in the field.

These two "modes" have been likened to the use, in the first case, of a keyword-based search engine such as Google, and in the second to the use of one or more in-depth subject or institution specific engines.

This immediately illustrates the problem with cross search portals which attempt to do everything – they quickly become the "lowest common denominator" and more like a Google than an expert research tool. The reason is that trying to do deep searching across multiple repositories is a very hard task.

The report on user requirements from the EVIE project [2] has some interesting comments on this. In Computer Science, researchers rely heavily on Google, which gives them excellent coverage. Within Geography, however, researchers turn first to the Web-of-Science, a subscription service provide by ISI. Within Medicine, researchers rely on Medline and PubMed. The Web-of-Science appears to have poor coverage on computer science topics. Researchers in Geography and Medicine make use of Google, but the terms that they use in their searches have wide applicability and appear in many unrelated documents.

An unexpected comment from one source was that institutions might not fully re-use portal services provided by JISC. Indeed at present most research libraries have home-grown Web interfaces to their chosen range of IT services, often dictated both by user requirements and cost. It is clear that portal services provided by JISC must be in a form that can operate alongside other institutional tools and be configured to have a common "look and feel".

2   Introduction

Rob A. and Rob C. had a short phone call at about 17:00 hrs and we want to see what your reaction is to what we think might be the most likely finding of our ITT, namely *the need for new standards and tools inside portals to support the cross searching of (institutional, commercial and JISC funded) digital repositories by researchers*. This kind of facility would mean that researchers would not need to login separately to each system, the whole lot (of selected repositories) would be available at one go. This is our first thought, though there may be others, but it would be good to find out what level of support there is out there for this kind of thing. What do you think researchers will want?

If you agree we would need to find out if there is any supporting evidence for facilities that simplify the research process. There is a shed load of reasons why this is difficult to do, but is the bullet Jisc will need to bite, if it wants to avoid fragmentation. There may be another way of doing it not quite sure yet.

Supporting evidence would come from the list of DR tools and the standards they use, e.g. Fedora, D-Space be it:

- Institutional, e.g. Lancaster and Heriot Watt university libraries, CCLRC library as an example of a Research Council.

- Commercial, e.g. JSTOR, Elsevier etc.

- JISC funded, e.g. SPP.

We also discussed recent trends and how libraries will need to change. There is a trade off between purchasing copies of books and journals for libaries and subscribing to an on-line service. The latter can be expensive and restrictive owing to emerging business models, but the former has overheads in storage and effort required. We mentioned Safari [6] in this context.

We thought that it might be better is we go out with some thoughts about what things researchers want, rather than just try and make sense of many different offerings we are likely to get if ask them what they want, as they probably will not have thought about it before. They may all be of the mind/ attitude, *this is what I have done in the past and that's all I need.*

The interfaces all the repositories are web based, there is only one I know of that comes as a web service, so there is another problem here for combining them all in a portal,

I think that the further we get with this kind of discussion before we meet anyone the better, we need them to feel that we know what the issues are (regarding the big picture) the moment we start talking to them, other wise they will feel that they are wasting their time. We would need to check our vision with ID. We would also have to come up with a strategy for getting there.

Perhaps there is a good reason why we would not want to change anything, besides it being difficult and expensive, perhaps there is a compromise. There may also be other lower changes that people want.

Roddy MacLeod (Heriot Watt) sent the following comments: *Why didn't the SPP develop into subject portal services? JISC are wary of funding expensive services. Also, it was possibly thought that*

commercial portal products such as MetaLib and ZD Portal solved the problem for researchers, or even CSAs Multisearch.

Of course, if you're a wealthy university you can possibly afford to buy such products, and then often spend months configuring them.

For less wealthy institutions (and Heriot Watt is a good example) we can't afford such products and have to struggle on with things like this list of databases as being the access point `http://www1.hw.ac.uk/library/electresa2z.php` for databases. Please - scroll down that list to see how user unfriendly it is! This is similar to the situation at CCLRC, see Figure 3.

You don't ask this question, but I will offer what I think researchers "really" want.

If you offered UK academic researchers something like the Irish are developing, and the Spanish are developing, they'd jump at it. Information about the IReL is at `http://www.sfi.ie/content/content.asp?section_id=467&language_id=1&publication_id=1252` and I understand that CSA may be working on providing subject perspectives on this. In Spain the the Biblioteca Electrónica de Ciencia y Tecnología `http://www.biblioteca.secyt.gov.ar/` has just been announced. In Germany there's the the vascoda initiative `http://www.vascoda.de/`. In Canada there's the Scholar's Portal `http://www.scholarsportal.info/`.

I think you have to involve the commercial world for such initiatives to be successful.

Services like that would be much more useful to researchers than anything else.

Such initiatives mean that the majority of researchers use the same interface and have access to the same sources. It facilitates research pooling, which our present setup doesn't. Free OA content (of the kind often cross-searched by TechXtra), can also fit into this.

Finally, we at PerX have written about the need for subject-based approaches: `http://www.icbl.hw.ac.uk/perx/analysis.htm`

3   Typology of Services

We probably need a typology of digital repository services, (Wikipedia has the following definition: *The word typology literally means the study of types.*) possibly defined by the standards/ interfaces they use for searching their content.

Such a typology was developed by Rachel Heery and Sheila Anderson, see `http://www.ukoln.ac.uk/repositories/digirep/index/Typology` [13].

The following notes are taken from their Web site.

Digital repositories can be classified by Content type:

• Raw research data
• Derived research data
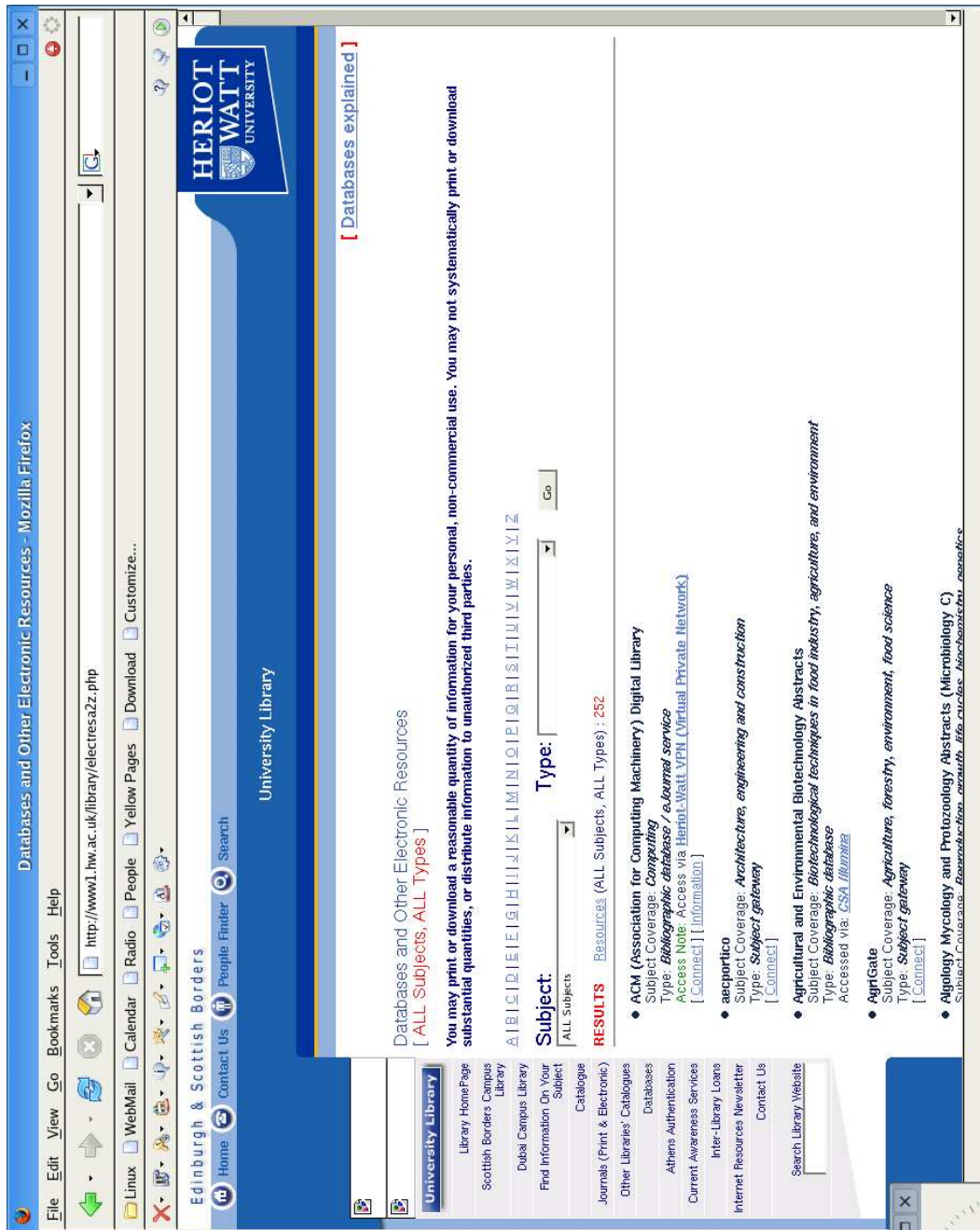• Full text pre-print scholarly papers

Figure 1: Heriot Watt Library on-line Services

- Full text peer-reviewed final drafts of journal/conference proceedings papers
- e-theses
- Full text original publications (institutional or departmental technical reports)
- Learning objects
- Corporate records (staff and student records, licences etc)

There is some content that currently appears to be largely missing within deployed repositories. For example there is little evidence of awareness within repository deployment of connections with archival management of courses as opposed to learning objects, what Lynch refers to as "composite structures (such as entire courses – in various sense, including both course "frameworks" and actual populated "instances" of courses within such frameworks – exported from learning management systems) ... " (Lynch, 2003). We think this would come under the heading of learning object repositories which is outside the scope of the current study.

A significant number of entries in institutional repositories are "metadata only" – with no link to the full text. This appears to be due to caution regarding copyright and IPR. Repository administrators and authors are reluctant to come into conflict with publishers regarding copyright issues so will not include "full-text" when there is doubt about copyright. In addition some repositories will only include links to full text for those entries published and/ or authored whilst the author was employed by the institution. So for example only a percentage of entries within the Southampton ECS e-Prints repository link to full text, and the CCLRC repository has a significant percentage of metadata only records. We will consider this in out "vision" under IPR and Watermarking.

They can also be classified by Coverage:

- Personal (author's personal archive)
- Journal (output of a single journal or group of journals)
- Departmental
- Institutional
- Inter-institutional (regional)
- National
- International

By primary Functionality of repository:

- Enhanced access to resources (resource discovery and location)
- Subject access to resources (resource discovery and location)
- Preservation of digital resources
- New modes of dissemination (new modes of publication)
- Institutional asset management
- Sharing and re-use of resources

By target user group:

- Learners
- Teachers
- Researchers

A variety of Content is hosted including:

- Multimedia, images
- Museum, library, archive objects

This typology of repositories has been used as the basis for the following deliverable from PerX [14].

The CDLOR project is developing a typology of learning communities that use repositories of teaching and learning materials. Their first deliverable is a report based on desk research and initial stakeholder consultation, which outlines the problems of developing such a typology, as well as giving an early typology of dimensions of communities rather than a typology of discrete community types, similar to the Cosmic Wheel of Fortune's breakdown of dimensions or characteristics of repositories noted in the Ecology section [16].

In the following sections we list some existing information services which are of interest to researchers by "Provider", which is similar to the "Coverage" typology above. These are only examples, and not a complete list.

If we list these, and say that the main gap is that there is no common standard, then what JISC need to do is fund the development of tools that map between the standards, so that they bring it all together in a simpler tool. The Open Archive Initiative community has made good progress in developing tools for the OAI harvesting standard. We also think that JISC will need to broker links with commercial and Institutional repositories, in order to get them to work together.

The on-line resources already accessible in some way via the JISC programmes are noted in the Collections Catalogue [1]. These are classified primarily by subject and then separated into those which are free of charge, and those which are subscription based. Many of them are noted below.

Just as one example, the User Requirements report from the EVEI VRE project [2] listed some of the key resources which their respondents used across a broad range of disciplines: Web of Science 39; Google 29; Medline 14; pubmed 6; Cinahl 4; Embase 4; Cochrane Library 3; Google Scholar 3; SOSIG 3; Amed 2; arXiv 2; Biosis 2; JStor 2; Psychinfo 2; Psychlit 2; Science Direct 2.

3.1 JISC supported Services

The JISC Information Environment home Web site is at `http://www.jisc.ac.uk/index.cfm?name=ie_home`. There are two major aspects to such an information environment and to the tools and mechanisms that it will enable institutions to use and exploit it. The first relates to users of the services. Being able to cross-search and use customised, value added and other services will considerably simplify users' interactions with online resources. This should encourage take-up and greatly improve means of accessing these resources. In turn, institutions will be able to incorporate a range of these services within their own institutional online environments, presenting local content alongside nationally provided resources. Portals are an important aspect of this as many institutions have their own partals, e.g. for administration or teaching. Facilities will also be provided for institutions and individuals to create and to share content, and in many cases to adapt them to local use.

The second aspect relates to making the information environment actually work. This requires the implementation of a range of commonly-agreed technical standards and protocols. This is a complex undertaking, involving influencing national resource providers and services, commercial suppliers and vendors and content creators, and requires considerable ongoing investment.

AHDS: The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities. Currently, they cover five subject areas: archaeology, history, visual arts, literature languages and linguistics and performing arts. It is organised via an Executive at King's College London and five AHDS Centres in these domains, hosted by various Higher Education Institutions. The AHDS is funded by the Joint Information Systems Committee and the Arts and Humanities Research Council, see `http://ahds.ac.uk`.

ArchivesHub: A national gateway to currently 19,723 descriptions of archives in UK 145 universities and colleges. These are primarily at collection-level, although complete catalogue descriptions are provided where they are available. The Archives Hub forms one part of the UK's National Archives Network, alongside a number of related networking projects. The hub also has a BLOG and provides training and advice on issues relating to electronic cataloguing of archives. `http://www.mimas.ac.uk/archiveshub/`

COPAC: COPAC, Combined University Catalogue, is a union catalogue. It provides FREE access to the merged online catalogues of 24 major university research libraries in the UK and Ireland plus the British Library, the National Library of Scotland, and the National Library of Wales/ Llyfrgell Genedlaethol Cymru `http://copac.ac.uk/`. COPAC is a MIMAS service funded by JISC with records supploed by CURL.

EDINA: Provide a range of on-line services, some listed below. Based in Edinburgh. `http://edina.ac.uk`

IESR: Information Environment Service Registry – a machine readable registry of electronic resources. The aim of the IESR is to make it easier for other applications to discover and use materials which will help their users' learning, teaching and research. The IESR currently holds this information for a selected set of electronic resources within the JISC Information Environment, provided by: Arts and Humanities Data Service (AHDS), Edina, MIMAS, Resource Discovery Network, UK Data Archive, UK Mirror Service (now JISC National Mirror Service). The project partners are MIMAS, UKOLN and The University of Liverpool. See `http://iesr.ac.uk/about/index.html`

JISCMail: JISC mail list management service has archiving facilities. Archived E-mails can be quoted as references in publications. JISCMail can also host discussions and surveys and is looking into a Wiki service. `http://www.jiscmail.ac.uk`

JORUM: is a free online repository service for teaching and support staff in UK Further and Higher Education Institutions, helping to build a community for the sharing, reuse and repurposing of learning and teaching materials. See `http://www.jorum.ac.uk`

LandMap: A MIMAs service for satellite image data `http://landmap.mimas.ac.uk`.

MIMAS: Provide a range of on-line services, some listed below. Based in Manchester. `http://www.mimas.ac.uk`

RDN: Resource Discovery Network. The RDN is a free Internet service dedicated to providing effective access to high quality internet resources for the learning, teaching and research `http://www.rdn.ac.uk`. Collection includes: Artefact (arts and humanities); EEVL (engineering, mathematics and computing); Geosource (geography and environment); Biome (health, medicine

and life sciences); ALTIS (hospitality, leisure, sports and tourism); HUMBUL (humanities); PSI-Gate (physical sciences); SOSIG (sociel science, business and law). The RDN selects, catalogues and delivers high-quality internet resources for further and higher education *the best of the Web*. SPP, the Subject Portals Project, provides a common interfaces to RDN resources.

RSC: JISC Regional Support Centres

UKDA: UK Data Archive, including Census data `http://census.data-archive.ac.uk/`.

ZETOC: Zetoc provides access to the British Library's Electronic Table of Contents of around 20,000 current journals and around 16,000 conference proceedings published per year. The database covers 1993 to date, and is updated on a daily basis. It includes an email alerting service, to enable you to keep up-to-date with relevant new articles and papers. Zetoc is free to use for members of JISC-sponsored UK higher and further education institutions. It is also available to English NHS Regions, NHS Scotland and Northern Ireland. A small number of other institutions are eligible to subscribe to Zetoc. We note that Research Council institutions are currently not eligible but have to use the British Library Citation Service instead. Run from MIMAS `http://zetoc.mimas.ac.uk/zetoc/` onbehalf of the British Library and JISC. ZETOC also provides and RSS news feed service.

3.2   Institutional Services

There are many institutional services, the list below just shows some examples.

AIM25: AIM25 is a major project to provide electronic access to collection level descriptions of the archives of over fifty higher education institutions and learned societies within the greater London area (inside the M25, hence the name). Coordinated at King's College `http://www.aim25.ac.uk`

CURL: Consortium of University Research Libraries, provide a variety of institutional services and act as a discussion forum `http://www.curl.ac.uk`. CURL's mission is to increase the ability of research libraries to share resources for the benefit of the local, national and international research community.

ePubs: CCLRC's e-Pubs system for research publications relating to staff and facility users [5]. See Figure 2.

Royal Society Library:

CETL: Centres for Excellence in Teaching and Learning

Regional Libraries: e.g. Manchester City Library on-line services `http://www.manchester.gov.uk/libraries/online`

SCONUL: Society of College, National and University Libraries `http://www.sconul.ac.uk`. All universities in the United Kingdom and Ireland are SCONUL members: so too are many of the UK's colleges of higher education. Also members are the major national libraries both sides of the Irish Sea. Most of our activities are carried out by the heads of library services, often through SCONUL's range of expert groups or our Executive Board. Full-time support and coordination is provided from SCONUL's office in London. They have three staff.
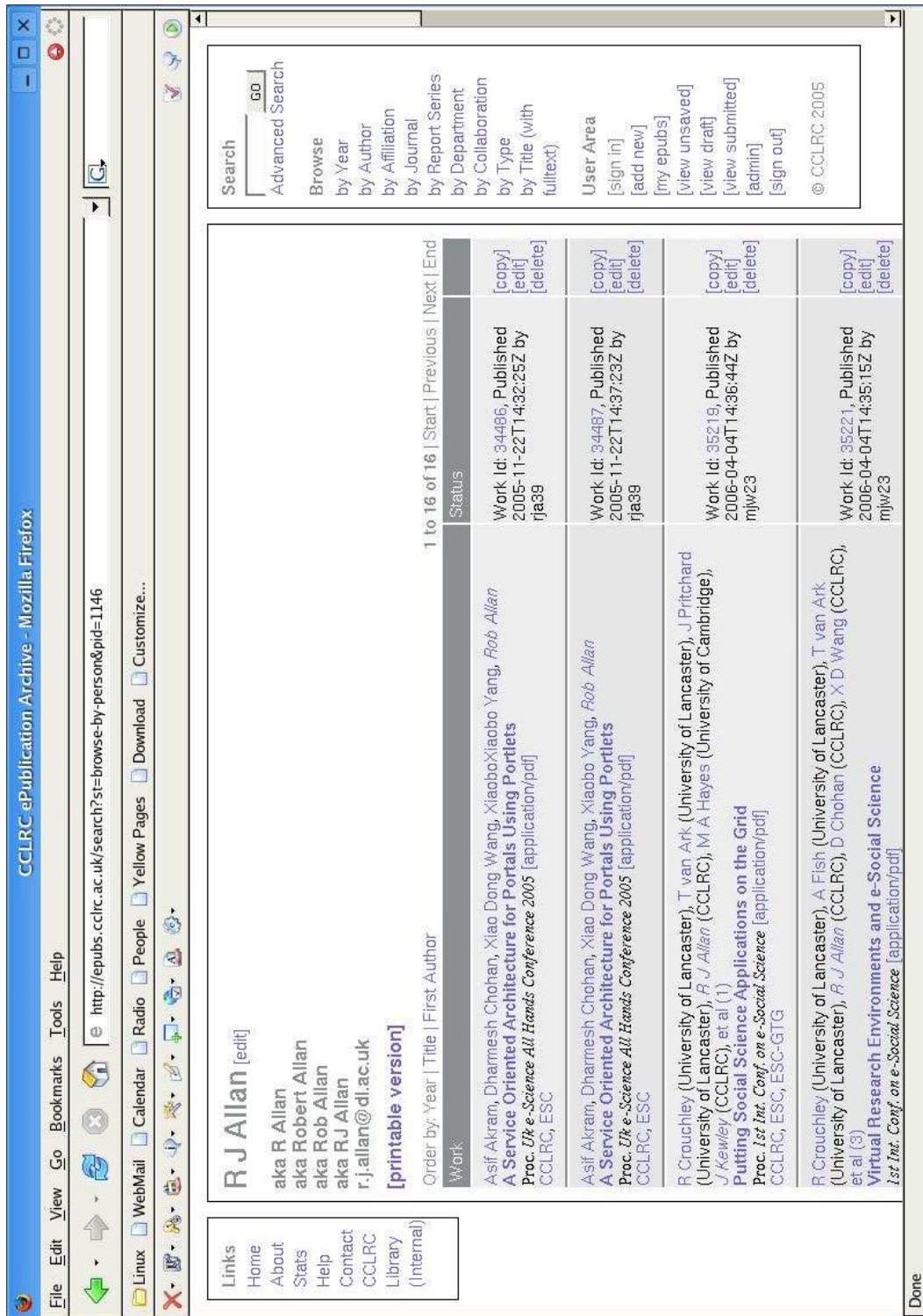
Figure 2: CCLRC ePubs Interface

3.3   Subject-specific Services

There are also many subject-specific services – here is a selection:

BADC: The 7 NERC Data Services

BMA: British Medical Association

BUBL: A large collection of links arranged by subject. Produced for the Higher Education community Free access `www.bubl.ac.uk`

CDS: Daresbury Chemical Database Service. Online access to a variety of chemical database services. UK academics can register free of charge for this service.

ChemWeb: Access to a library of chemical journals, a conference diary, and news service. Register free by visiting the site. `http://www.chemweb.com/`.

IOP: Institute of Physics

NPOR: National Pipe Organ Register and Database of British Organ Builders, maintained at Unviersity of Cambridge and Birmingham City Library `http://npor.emma.cam.ac.uk`

OMNI: Biomedical and healthcare, OMNI offers free access to a searchable catalogue of Internet sites covering health and medicine. Free access. `http://omni.ac.uk`

ChipCenter QuestLink: Information on electronic components. Information on semiconductor intellectual property. Register free by visiting the site. Many resources are free access `http://cmpmedia.globalspec.com/`

Projects: Project Web sites

EEVL: Links to quality engineering, mathematics or computing Web resources. Free access `http://www.eevl.ac.uk`. See also other RDN services provided by JISC, ESRC and AHRC.

PhysicsWeb: Physics News, press releases, events, jobs Free access, but some resources are restricted to Institute of Physics members `http://www.physicsweb.org`

Regard: Searchable database of UK research funded by the Economic and Social Research Council. In-depth information on projects, publications, and other research tools `http://www.esrc.ac.uk/ESRCInfoCentre/index.aspx`

3.4   Open Access and Web-based Services

In June 2005 the UK Research Funding Councils released their position statements supporting open access and the use of institutional and subject-based repositories. The councils' statements means that the eight publicly funded research councils in the UK join with the National Institutes for Health in the USA and the international Wellcome Trust in supporting Open Access and the results of research funding being made available to the widest audience.

The principle that public funding of research should result in public availability of research findings is of undoubted benefit to researchers, the research process and the public good. The guidance that the Research Councils have given in supporting open access to funded research will increase the availability, dissemination and use of UK research with benefits to both individual researchers and institutions. Different Research Councils have slightly different requirements for their researchers to fulfill: some recommend open access deposition, while others require it as a condition of grant. All of them support the RCUK position statement previously released and available from the RCUK Web site `http://www.rcuk.ac.uk/access/`.

For a description of tools developed by members of the Open Archives Initiative community, see `http://www.openarchives.org/tools/tools.html`. Tools are available for personal or public document managment and are written in a variey of languages, such as Perl, PHP and Java.

For references to Open Access see [10, 11]. A report commissioned by JISC gives further background on feasibility and requirements [12].

Registered service providers are listed here `http://www.openarchives.org/service/listproviders.html`. Some of these services are described below. Services on this list have been cited by people we interviewed or were mentioned in previous surveys as being used by researchers.

Archimede: Laval University Library – Archimede is an open-source software for institutional repositories. It features full text searching, multiplatform support, Web user interface, and more. Archimede fully supports OAI-PMH requests version 2.0.

Amed: Allied and Complementary Medical Database. a unique bibliographic database produced by the Health Care Information Service of the British Library. It covers a selection of journals in three separate subject areas: several professions allied to medicine; complementary medicine; palliative care. `http://www.bl.uk/collections/health/amed.html`

Arc: Old Dominion University – Arc is released under the NCSA Open Source License. Arc is a federated search service based on OAI-PMH. It includes a harvester, a search engine together with a simple search interface, and an OAI-PMH layer over harvested metadata. Arc can be configured for a specific community, and enhancements and customizations by the community are encouraged. Arc is based on Java Servlet technology and requires JDK1.4, Tomcat 4.0x, and a RDBMS server (tested with Oracle and MySQL).

ARL: Association of Research Libraries, Washington DC. Office of Scholarly Communication `http://www.arl.org/osc/index.html`.

arXiv: arXiv is a free e-print service for physics, mathematics, non-linear science, computer science, and quantitative biology `http://uk.arxiv.org`.

[Biosys:] IOSIS information solutions for the global life sciences community are now provided by Thomson Scientific. Their databases are the most complete resource for finding life sciences information quickly and efficiently. They select documents from thousands of sources worldwide, index and abstract them into citations which describe their content, and maintain databases for searching citations, adding more than 600,000 new entries each year. `http://www.biosis.org`

British Library: `http://www.bl.uk`

Callima: from infoball is a search engine for scientific articles from various subject areas and sources. It provides a single point of access to a significant number of Open Archives.

CERN: CERN preprint server `http://weblib.cern.ch/share/hepdoc/`. Where the Web came from! There are also various tools from CERN, such as CDSWare.

Citebase Search: from Southampton University provides users with the facility for searching across multiple archives with results ranked according to many criteria, such as creation date and citation impact.

Cochrane: The Cochrane collection of clinical trials results and associated library `http://www.cochrane.org`

Connotea: EPrints repositories can be part of the Semantic Web in a real and practical way thanks to Nature Publishing Group and its free Connotea online reference management service. This service enables users to publicly bookmark and tag articles from within EPrints repositories with remarkable potential to expand the visibility and findability of those articles.

COS: Community of Science provides a full range of Internet-based services for the world's researchers enabling universities, corporations, societies, private institutions, government agencies and individual researchers to find funding, promote their work, identify experts, manage resources, and collaborate with colleagues. It is run from the University of Maryland, Bethesda and is free to register `http://www.cos.com`.

CYCLADES: from the European Research Consortium for Informatics and Mathematics (ERCIM) is a system, designed to provide an open collaborative virtual archive environment, which (among others) supports users, communities (and their members) with functionality for (i) advanced search in large, heterogeneous, multidisciplinary OAI compliant digital archives (ii) collaborative work; and (iii) filtering and recommendation of records, users, communities and collections. CYCLADES allows users to register any OAI archive to CYCLADES, which then will be automatically harvested and indexed.

DSpace: HP Labs and MIT Libraries – DSpace is an open source digital asset managment software platform that enables institutions to capture and describe digital content. It runs on a variety of hardware platforms and supports OAI-PMH version 2.0.

EPrints: EPrints [9] worldwide community promoting open access for academic publications. Related services include CiteBase, ROAR, ROARMAP, ROMEO, OAA Bibliography, Open Citation Project. Managed from University of Southampton. As an example of the e-Prints news service, see `http://www.eprints.org/news/features/connotea.php`. See also the SHERPA project `http://www.sherpa.ac.uk`.

ESPcenet: European Patent Office gateway for online access to full-text patents worldwide, from 1920 to within the last few weeks. Access from your desktop. No password required. `http://gb.espacenet.com/search97cgi/s97_cgi.exe?Action=FormGen&Template=gb/EN/home.hts`

Fedora: Cornell University – An open source digital repository architecture that allows packaging of content and distributed services associated with that content. Fedora supports OAI-PMH requests on content in the repository.

Google: Google [4]

Google Scholar: Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. Google Scholar helps you identify the most relevant research across the world of scholarly research.

Google Maps: a mapping facility which can be used to present other information, such as the location of related services. An example is the worldwide lication of Sakai servers: `http://www.dr-chuck.com/sakai-map`

Google Earth: a repository of aerial images across the whole of the planet which can be navigated and linked for projects

INSPEC: References to literature in physics, computing and engineering research. For more information see Description of INSPEC. Apply for an Athens password, for access from your desktop. There is a CHEST deal. `http://edina.ac.uk/inspec/`

JSTOR: basic idea was to convert the back issues of paper journals into electronic formats that would allow savings in space (and in capital costs associated with that space) while simultaneously improving access to the journal content. Originally a Mellon Foundatiaon project. Now a not-for-profit organisation `http://www.jstor.org`.

NMSI: National Museums of Science and Industry `http://www.nmsi.ac.uk/index.asp?flash=yes`

OAIIECSP: Open Archives Initiative Information in Engineering, Computer Science and Physics

OAIster: A worldwide search service based on OAI data harvesting which is hosted at the University of Michigan Digital Library Production Service `http://oaister.umdl.umich.edu/o/oaister/`. OAIster can currently search 8,857,208 records from 680 academic institutions (as of 3/8/06), including CCLRC's e-Pubs and Lancaster's e-Prints. OAIster allows searching of mang digital resources, including text, image, audio, video and datasets. There if a Firefox toolbar plugin for OAIster so it will get more widely used. We note that OAIster handles repositories from multiple languages.

Perseus: from Perseus. The system harvests registered OAI repositories and incorporates the information into its search interface.

PKP: Public Knowledge Project open archives harvester

PsychLit: PsychLit or as it is sometimes called Psycho/Info is a database and search engine that provides access to publications in psychology, psychiatry, neuropsychology, and neuroscience. It primarily contains journal articles and is fairly easy to use. The IUS library has the system on site, although it is not available online in its complete form. I would suggest that PsychLit be your primary source of information for research and literature reviews. An alternate database that is also of use is Medline that contains medical as well as psychological/ neuropsychological citations. `http://homepages.ius.edu/DBURTON/psychlit.htm`

RGO: Researcher's Guide Online

SAIL ePrints: from CNR - Area della Ricerca di Bologna. SAIL-eprints (Search, Alert, Impact and Link) is an electronic open access service provider for finding scientific or technical documents,

published or unpublished, in Chemistry, Physics, Engineering, Materials Sciences, Nanotechnologies, Microelectronics, Computer Sciences, Astronomy, Astrophysics, Earth Sciences, Meteorology, Oceanography, Agriculture, and related application activities. SAIL-eprints has been designed primarily to collect information on scientific documents (metadata) authored by CNR researchers and deposited as preprints or postprints in CNR institutional open access archives. Also, SAIL-eprints collects metadata from other data repositories all over the world publishing materials in the same scientific fields.

Scientific news: Science news bulletins from a variety of UK and US sources Free access `http://www.scicentral.com/`

Scirus: from Scirus, distinguishes itself from existing search engines by concentrating on scientific content only and by searching both web and (often proprietary) databases. Scirus' aim is to provide scientists with one comprehensive search platform covering both the web and the normally "invisible" databases. Scirus now harvests Open Archives data.

TORII: International School for Advanced Studies, Trieste, Italy. TORRII provides unified access to various open archives (Physics and Computer Science). Filtering and advanced searching. Personalization. For more information, see the TIPS consortium web pages at `http://tips.sissa.it`

Web of Knowledge: The ISI (Institue for Scientific Information) Web of Knowledge Service for UK Education provides a single route to all the Thomson Scientific products subscribed to by your institution. Connect to the ISI Web of Knowledge Service and select individual products for searching from the list on the home page. `http://wok.mimas.ac.uk/`

Xreferplus: Online access to many reference books including dictionaries, thesauri, encyclopedias, books of quotations and atlases. Can also do topic maps. For more information see Xreferplus guide. Access from your desktop for internal use. No password required. For offsite access using Athens authentication. There is a JISC deal for academics. `http://www.xreferplus.com/search.jsp`.

The OpenDOAR project `http://www.opendoar.org` provides a gateway to a very large number of OA repositories. OpenDOAR is part of the SHERPA project.

3.5 Commercial Services

There are a number of important commercial sources of information, particularly from academic publishers of books and journals. Digitisation of services rather than sales of books and journals is leading to a huge change in the way access to content is managed. We have to be aware that new business models are still emerging and are likely to be in a state of flux for some time.

British Standards: catalogue available on line. Full text British and ISO Standards. Apply for an Athens password, for access from your desktop. These are expensive, but there is a JISC academic deal. `http://www.bsonline.bsi-global.com/server/index.jsp`

Cinahl: On-line health care information `www.cinahl.com`

Elsevier: `http://www.elsevier.com` offer a portal style interfaces to their publication lists which can browse by subjector product type or search individual listings.

Embase: more than 17 million validated biomedical and pharmacological records from Embase and Medline offered via Elsevier `http://www.embase.com`

info4education: Engineering and electronics product data, occupational health and safety, construction. Apply for an Athens password, for access from your desktop `http://www.info4education.com`.

Knovel: Online access to 10 engineering and science reference works. Access from your desktop. No password required for the free trial, but main service is commercial. `http://www.info.knovel.com/essentials/`

LexisNexis: LexisNexis provides authoritative legal, news, public records and business information; including tax and regulatory publications in online, print or CD-ROM formats `http://www.lexisnexis.com/`.

Optical Society of America: Conferences, publications and industry news. Many resources are free access, but some areas restricted to OSA members `http://www.osa.org`

PubMed: PubMed Central (PMC) is the U.S. National Institutes of Health (NIH) free digital archive of biomedical and life sciences journal literature. A service of the National Library of Medicine that includes over 15 million citations from Medline and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources. `http://www.pubmedcentral.nih.gov/`

Safari IT ebooks: Technical IT books on-line from O'Reilly, Addison Wesley and others. Institutions can subscribe to a "bookshelf", CCLRC currently accesses 150 books for a payment of £3.5k p.a., others available on additional payment. URL illustrates CCLRC's interface `http://proquest.safaribooksonline.com/?uicode=CCLRC`. Access from your desktop. You are strongly advised to familiarise yourself with the terms and conditions of use.

Science Direct: Science Direct

SFX: Library OpenURL resolver cross search service. URL illustrates CCLRC's interface to the journals for which we subscribe. `http://sfx4.exlibrisgroup.com:3210/cclrc/a-z/default`

Talis: Talis Information Limited `www.talis.com`, the leading provider of library management solutions for the U.K. and Ireland, announced on 27/4/06 that Endeavor Information Systems `www.endinfosys.com`, one of the leading worldwide providers of library management software, is the first company to join its new partnership programme, Talis Connexions.

Talis Connexions is a flexible framework that facilitates collaboration between like-minded organisations providing library management solutions to academic, research, corporate and government institutions. Members are encouraged to share knowledge and technical assets, support open standards and develop complementary product strategies.

Both Talis and Endeavor are dedicated to supporting access, interoperability and standards-based integration where appropriate, for the benefit of the wider communities served. Talis Connexions also provides the foundation for managing specific technology initiatives, such as interfacing Talis Alto, the company's integrated library management system, with Endeavor

Meridian, as well as fostering interconnectivity between Endeavor's integrated library system, Voyager, and Talis' reading list solution, Talis List.

We believe that Talis is also doing research linking to Google and Amazon and using Google Maps to locate institutions having copies of publications via OpenURL and RDF. This is potentially a far-reaching project.

Special deals exist for making commercial services available to academic researchers. Some of these are brokered by JISC and some by CHEST. As noted above, Research Councils such as CCLRC which are actively engaged in research support rather than teaching are not eligible for all of these, which can be a problem for researchers activly engaged in work on the large-scale facilities.

4   Putting it All together

What is required to make all these services accessible in a single portal? What are the gaps and impediments?

Some issues/ questions:

1. We need to list/ review all the deep searching standards, OAI-PMH `http://www.oaforum.org/tutorial/`, X39.50 and propose a tool so that a query in SQL (say) can be converted to each of them and then sent off;

2. SPP used Athens, Athens may not be up to the task in the big picture of cross searching all DRs, our solution would need to be designed with Shibboleth migration in mind. We need single sign on to all the DRs;

3. The SPP did a JSR-168 portlet for their cross searcher, but there was some caveat about it, perhaps they did not get single sign on to work. This originally worked inside Jetspeed, but if done today would probably use Pluto. Also JSR-168 was not an appropriate standard for the kind of functionality, is JSR-286 up to the job?

4. If would be good if we could propose a facility for researchers to create their own DR from Web based content on the fly from a seed taxonomy, like we do in ReDReSS. This would facilitate linking private and public information systems.

Many of the services listed above have their own diverse form-based Web interfaces for searching by author, title, subject etc. There are many issues to be faced in making all these available from a single interface. Some are noted below.

Authentication and single sign-on Athens, Shibboleth. The current Athens password service supported by Eduserve is being phased out and replaced by a Shibboleth service over the next year or so. This is likely to have a huge impact on the services provided. Charging for registrations has been suggested. Institutions may also wish to link to lodap LDAP services in the interim.

Authorisation is a key area.

Attribution Providers wish to retain IPR. There are also issues of liability in hosting content. Watermarking has been considered, e.g. in the Pixus project.

Cross search services e.g. SPP, CREE, Xgrain. Discussion of cross-search services vs. union catalogues, e.g. Powell [3]. MetaLib is a package from exLibris for federated searching which is widely used `http://www.exlibrisgroup.com/metalib_faq.htm`. It sits above SFX. It was possibe that, though expensive, commercial portal products such as MetaLib and ZD Portal solve the problem of cross search for researchers, or even CSAs Multisearch.

Location Service OpenURL is used as a basis for full text location from an appropriate source. This resolves a URL found from a cross search based on a variety of factors, such as subscription, institution type, location of document, etc. In some implementations it can automatically create forms for ordering copies via inter-library loan or other delivery service.

Using OpenURL resolvers means that copies of content can be kept in the original location rather than lodged on multiple host servers. This can avoid copyright/ IPR problems. Services such as Web of Science and IOP use OpenURL. Such OpenURL-enabled services can publish papers who's references are also OpenURL linked rather than using fixed URLs as in hypertext.

Publication Tools e.g. Google. Andy Powell wrote a useful paper comparing JISC and Google services and suggesting how Google could be used to find JISC services.

Search protocols e.g. Z39.50, SRU/ SRW.

Service Registries e.g. IESR

4.1   What works Now?

We illustrate what can be done today. As an example, an interface to on-line services provided by the Heriot Watt library was shown in Figure 1. Another for the CCLRC Library is shown in Figure 3. This includes links to the following: data sources: arXiv, BS online, CCLRC ePubs, Daresbury Chemical Database Service, Espcenet, info4education, INSPEC, Knovel and K-Essentials, Library electronic journal service, PubMed, Safari IT ebooks, Web of Knowledge, Xreferplus. It also provides access to a number of subject-specific gateways and other reference sources.

An interesting approach is the information portal provided as an interface to the Australian Bond University Information Repository. This has a Blogger on its home page `http://www.wordpress.com` and links to the e-Publication repository `http://wpublications.bond.edu.au/` and also the university library services `http://www.bond.edu.au/library/`.

The Australian Universities have a shared discovery service called ARROW: Australian Research Repositories Online to the World. This currently includes Australian Digital Theses, Library and Information Association, Policy Online and universities of Bond, Curtin, Flinders, Monash, Queensland Technology, Melbourne, Queensland, Suouthern Queensland, Tasmania and Wollongong.
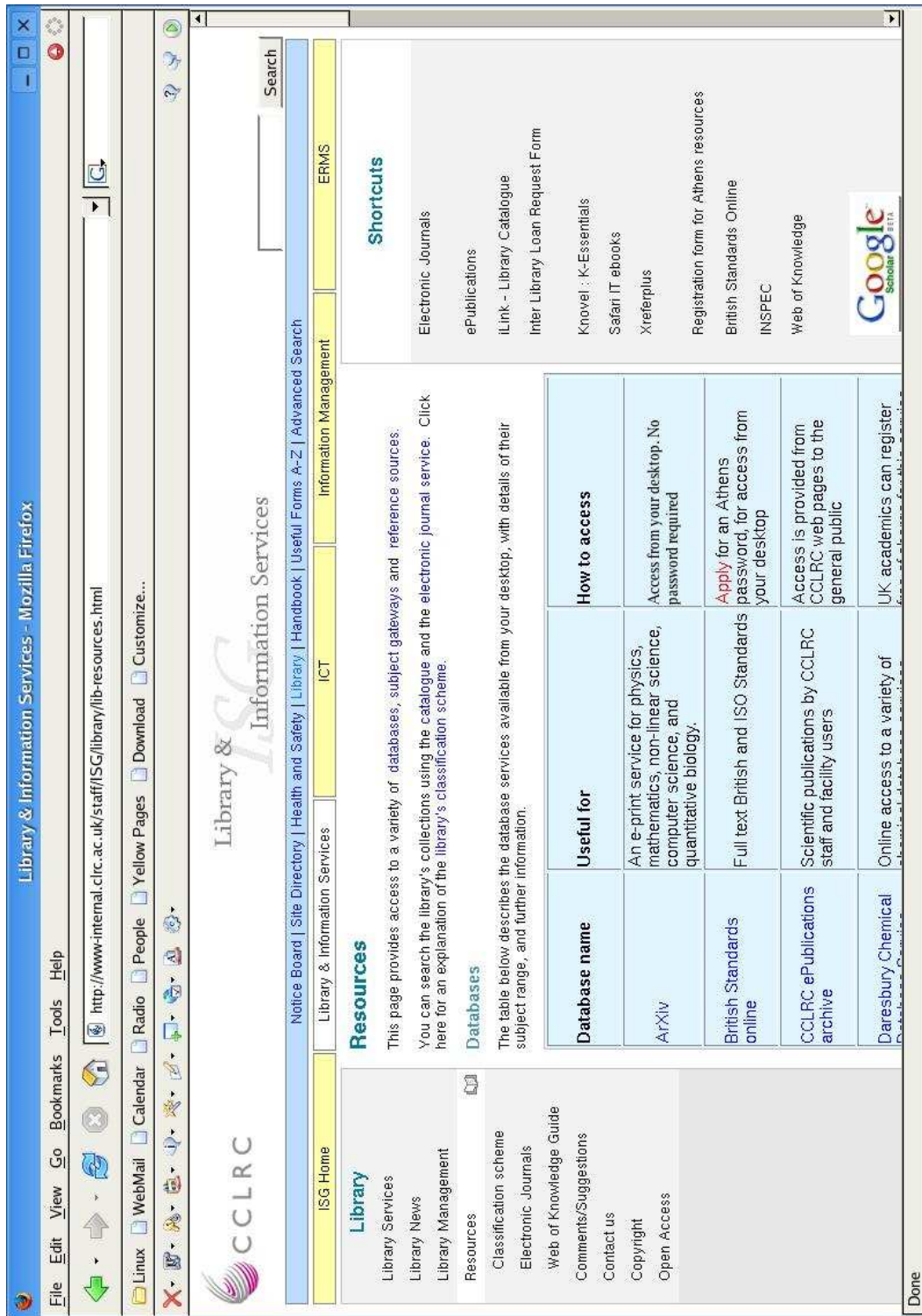
Figure 3: CCLRC Library on-line Services

4.2   What doesn't work?

[Summary of gaps and requirements for future funding.]

5   Acknowledgments

References

[1] e-Resources *Collections Catalogue of Online Resources* (JISC 2005) `http://www.jisc.ac.uk/collections`

[2] D.M. Sergeant, S. Andrews and A. Farquhar *Embedding a VRE in an Institutional Environment (EVIE). Workpackage 2: User Requirements Analysis* User Requirements Analysis Report (University of Leeds, 2006)

[3] A. Powell *JISC Information Environment (IE) Inventory with Diagrams* (UKOLN, 2005)

[4] Andy Powell *The JISC Information Environment and Google* Discussion paper for the JISC (UKOLN, November 2004)

[5] CCLRC e-Pubs `http://epubs.cclrc.ac.uk`

[6] `http://safari.oreilly.com/`

[7]

[8]

[9] ePrints `http://www.eprints.org`

[10] John Willinsky *Open Access, The Access Principle* (MIT Press, 2005) is now available Open Access.

[11] Okerson and O'Donnell *Open Access: Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing* (ARL, 1995) `http://www.arl.org/scomm/subversive/toc.html`

[12] H. James, R. Ruusalepp, S. Anderson and S. Pinfield *Feasibility and Requirements Study on Preservation of E-Prints* report to JISC (29/10/03) `http://www.jisc.ac.uk/uploaded_documents/e-prints_report_final.pdf`

[13] R. Heery and S. Anderson *Digital Repositories Review* `http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf` (UKOLN and AHDS, 19/2/05)

[14] Roddy MacLeod and Malcolm Moffat *Engineering Digital Repositories Landscape Analysis, and Implications for PerX* Version 1.0 (10/11/05) `http://www.icbl.hw.ac.uk/perx/analysis.htm`

[15] Neil McLean *The Ecology of Repository Services: A Cosmic View* Presentation given at ECDL 2004, University of Bath. `http://www.ecdl2004.org/presentations/mclean/`

[16] Animated presentation of the 'wheel' in action: `http://www.rubric.edu.au/extrafiles/wheel/main.swf`

[17] Implementation Project, Bond University `http://epublications.wordpress.com`

[18] R.J. Allan, R. Crouchley and C. Ingram *Scenarios, Use Cases and Reference Models* (CCLRC, June 2006)

[19] R.J. Allan, R. Crouchley and C. Ingram *Comparison of Surveys* (CCLRC, June 2006)

[20] R.J. Allan, R. Crouchley and C. Ingram *Web-based Library and Information Services* (CCLRC, June 2006)

[21] R.J. Allan, R. Crouchley and C. Ingram *The Information Environment and e-Research Portals* (CCLRC, June 2006)

[22] R.J. Allan, R. Crouchley and C. Ingram *Interim Report* (CCLRC, June 2006)

[23] R.J. Allan, R. Crouchley and C. Ingram *A Vision for Portal access to Global Information* (CCLRC, June 2006)

[24] R.J. Allan, R. Crouchley and C. Ingram *Final Report* (CCLRC, June 2006)