

# JISC Information Environment Portal Activity: supporting the Needs of e-Research.

## Comparison of surveys

### **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](mailto:r.j.allan@dl.ac.uk), [r.crouchley@lancs.ac.uk](mailto:r.crouchley@lancs.ac.uk),  
[caroline@csiconsultancy.co.uk](mailto:caroline@csiconsultancy.co.uk)

August 10, 2006

## Introduction

The surveys chosen for this review were felt to have asked some questions relevant to the question of whether a portal could be developed which would suit researcher/user requirements, and more specifically in terms of e-research. Unfortunately no previous studies into e-research user requirements were found in the UK. Related studies from Australia and the US are outlined and their conclusions recorded.

e-Science is a new paradigm in research, often characterised by a "deluge" of data analysed by massive distributed computing power. Physicists and astronomers may seek to leverage the large investments in data production and collaborate on the analysis of massive data outputs, whilst educationalists and ethnographers may hope to share large multimedia research datasets. Grid technology, emerging in response to these challenges, is enabling exciting possibilities for better research, even creating new disciplines like astro-informatics. In this context, a wide range of national and international initiatives are under way.

The concept of e-science is now broadening and evolving into e-research generally, to encompass the social sciences and the arts and humanities. Increasingly research is becoming more multi-disciplinary and more collaborative. However, different communities are at very different stages in their awareness of the new technologies: thus the current needs of a large international scientific collaboration are likely to be much more complex than those of the lone humanities researcher, wishing to collaborate more effectively with a handful of colleagues world-wide in the same field of interest. There will be a whole range of requirements to suit research in different areas.

With regard to a portal for e-research, we take the following definition by Andy Powell, reported in Miller 2003<sup>1</sup>: "In the context of the JISC Information Environment (IE), portals typically focus on supporting the end user in their learning and/ or research activities by providing personalised discovery services across multiple, heterogeneous content providers."

## Background literature review

The Superjournal (1996-1998) study<sup>2</sup> had already identified different types of user which could partly be mapped to discipline:

---

<sup>1</sup> Miller, P. (2003) Towards a typology for portals. *Ariadne*, 37 <http://www.ariadne.ac.uk/issue37/miller/intro.html>

<sup>2</sup> <http://www.superjournal.ac.uk/sj/baserept.htm> (1998)

- journal-focused; these people focused on 4-5 journal titles and 50% of the resources used were full-text (mainly scientists)
- topic-focused; these people searched by subject rather than journal, used many articles (mainly social scientists)
- article-focused; these people searched only one journal (mainly scientists)

A further study by DLF/CLIR/Outsell<sup>3</sup> (2002) into use of the scholarly information environment found that there were differences between researchers in the use of electronic resources according to subject discipline. Asked whether they used electronic sources “All/most/some or none” of the time, 65% of law respondents used e-resources all the time, dropping to 56% in business, 48% in the biological sciences and engineering, 46% in the physical sciences and maths, 37% in the social sciences and 25% in the arts and humanities.

In terms of unmet information and content needs, 76% of law respondents reported no unmet needs, compared with 54% of business, 54% of arts and humanities, 51% of physical sciences and maths, 48% of social sciences, 41% of engineering, and 42% of biological sciences. “More online journals” was the major unmet need across all groups, though substantial groups of arts and humanities, physical science, maths, biological science and social science wanted more print journals. The biggest problem identified across all groups by the DLF/CLIR/Outsell study is ‘having enough time’, followed by ‘knowing what’s available’ and ‘having access to all information from one place’. This all points to a lack of training into the use of online resources, and a lack of adequate marketing and ‘pushing’ of resources to the appropriate users. [The EVIE UR doc agreed with this.](#)

The PREST study<sup>4</sup> (2000) found that 83% of university staff were collaborating with researchers from another discipline or intended to do so. Work by Evaluation Associates<sup>5</sup> (commissioned by the funding bodies to examine whether assessment discriminated against interdisciplinary research) in 1999 estimated that 46% of research time by UK academics was spent on interdisciplinary research and 10% were fully interdisciplinary researchers. It is expected that for some discipline areas (particularly **particularly those involved in the e-Science programme**) these percentages will have increased.

A report by Key Perspectives<sup>6</sup> (2005) has a number of elements which are relevant in the context of the current research:

- The study found that there was no major difference between disciplines when it came to ease of access to resources: 54% of respondents said they had easy access to most of the articles they need for their work: the lowest percentage of respondents reporting this level of ease were in Law and Politics (41%); Humanities (44%), Business & Management (47%) and Psychology (49%).
- The age of needed articles (i.e. how long ago most articles needed were published) varied considerably by discipline group, with the peak age in humanities being about 20 years ago, in chemistry, engineering and medicine 10 years ago, and computer science, life sciences and information science, less than 5 years ago.

The study also explored how researchers look for research information – that is, in their role now as *users* of information. Bibliographic databases are quite heavily used. Nineteen percent of respondents search them daily and 42% claimed to do so at least once a week. A further 21% search them at least once per month; only 16% do so less frequently than once per month. In all, 98% of respondents use this type of database as a search tool. This is in

<sup>3</sup> DLF/CLIR/Outsell study (2002). *Dimensions and use of the scholarly information environment*. <http://www.clir.org/pubs/reports/pub110/contents.html>

<sup>4</sup> PREST, *Impact of the RAE and the future of quality assurance in the light of changes in the research landscape*. April 2000, Final report prepared for HEFCE.

<sup>5</sup> Higher Education Funding Council for England. *Interdisciplinary Research and the Research Assessment Exercise*, RAE 1/99, April 1999, Report for HEFCE by Evaluation Associates, p.5.

<sup>6</sup> Swan A & Brown S. (2005). Open access self-archiving; an author study. Key Perspectives Ltd

contrast to the proportion of people who use other means to locate research information they need for their work, where a variety of methods are used, dependent on subject area.

When searching closed (or 'pay-to-access') archives, that is collections where the user has access only to the full-text of articles in journals to which they or their institution have a subscription, most respondents who could name a service use ScienceDirect (54%), but obviously these will be mostly scientists. Researchers in the humanities and other non-science subject disciplines were found to predominantly use other full-text services. In total, 86% of people use this type of service.

*When you search for articles online in CLOSED ARCHIVES, (i.e. where you are only allowed access to the full-text of articles in journals you or your library subscribe to), which services do you use regularly?*

<b>Services used regularly to search closed archives</b>	<b>Percentage of total respondents</b>
ScienceDirect	54
Subject-specific full-text services	25
CrossRef Search	7
Individual publishers' websites	38
Other cross-subject full-text services	22

Open access archives are reached by different means. Respondents were provided with a list of search engines that specifically harvest from these archives around the world and were asked to indicate which they use. Respondents who stated that they search open access archives do so fairly infrequently compared to the findings for bibliographic databases. Fifty percent search open access archives less frequently than once per month, 17% do so at least once per month, 12% at least once per week and 9% several times per week.

*Which services(s) do you use to search the content of OPEN ACCESS ARCHIVES (i.e. where you can access the full-text of any research article)?*

<b>Services used regularly to search open access archives</b>	<b>Percentage of total respondents</b>
Scirus	14
Citebase Search	10
Open Archives Initiative Information in Engineering, Computer Science & Physics(OAII ECSP)	3
OAlster	3
Public Knowledge Project Open Archives harvester (PKP)	2
Perseus	1.6
Arc	1.5
CYCLADES	0.5
Callima	0.4
SAIL-eprints	0.2
TORII	0.1

When it came to searching the World Wide Web for research articles on other websites, 72% use Google<sup>7</sup>, and 23% use another search engine. 11% don't search websites for this kind of information since this would not be fruitful for their research area.

---

<sup>7</sup> this survey was carried out just before Google Scholar was launched late in 2004

A paper by Stephen and Harrison<sup>8</sup> (2002) puts forward the view that it is essential to have a disciplinary focus when building electronic systems and resources for scholarly use. They suggest that existing e-resources fail to allow scholars within a field to search and retrieve material on the basis of discipline, both because of the way they are constructed and because they have insufficient historical depth: “*When an electronic service for scholars is expansive, multidisciplinary and ahistorical, it is not possible to pursue questions within the disciplinary frameworks that may originally have given them sensibility*”. Because scholars cannot guide their searches through the discipline-specific meaning of a concept, a keyword search approach yields a great deal of tangential material. Anecdotal evidence<sup>9</sup> in the UK points to researchers welcoming a certain degree of serendipity in their searching, hence tangential material could be useful. However, this should not be to the detriment of finding the result required. Stephen and Harrison appeal for a discipline-focused approach to designing e-resources:

“...central to each discipline is its own periodical literature, and disciplinary members are usually thoroughly aware of the prestige hierarchy within that set of journals and annuals... Similarly, e-mail discussion lists may play a vital role in one field but are never used in another, and article archives that emphasize recent materials suffice in particular fields while including older materials may be the only appropriate model for other fields.” *JISCMail archives...*

Fry (2004<sup>10</sup>) investigated the relationship between a degree of ‘mutual dependence’ and ‘task uncertainty’ in the production and use of digital resources across three distinct fields of research. He assessed this through interviews with researchers in these areas. With reference to the current study the following results table is most relevant since it clearly shows some differences between three disciplines:

<b>Field</b>	<b>High energy physics</b>	<b>Corpus based linguistics</b>	<b>Social/cultural geography</b>
<b>Culture of relationships between researchers</b>	High degree of mutual dependence, with low degree of task uncertainty	Moderate degree of mutual dependence with moderate degree of task uncertainty	Low degree of mutual dependence with high degree of task uncertainty
<b>Differential role of informal and formal communication</b>	Speedy establishment of knowledge claims via informal communication system of conference papers and pre-prints. Publication mainly serves citation criteria	Need to communicate a high concentration of technical information supported in conference proceedings, reports and manuals	Formal communication system important due to lower levels of interpersonal recognition (e.g. low people to problem ratio) and need to justify goals, approaches and techniques in literature. Informal communication system determined by individual groups and specific social networks

<sup>8</sup> Stephen, T. & Harrison, T. (2002) Building Systems Responsive to Intellectual Tradition and Scholarly Culture. *The Journal of Electronic Publishing*, University of Michigan Press, Vol 8, Issue 1, Aug 2002  
<http://www.press.umich.edu/jep/08-01/stephen.html>

<sup>9</sup> Discussions with Nicky Ferguson (Clax) and Hugh Look (Rightscom)

<sup>10</sup> Fry, J. (2004) Scholarly Research and Information Practices, A Domain Analytic Approach, Information Processing and Management

<b>Role of ICTs in communication system</b>	Tightly coordinated system for the informal dissemination of research results via integrated digital networks (GRID); production of centralised field-based digital resources	Quest for the development of a coordinated system for the informal dissemination of research results hindered by local ICT infrastructures; decentralised locally produced field-based digital resources	Non-production of field-based digital resources; reliance on commercially produced generalist digital resources; ICT infrastructure determined at the level of the employing institution [rather] than the field or discipline
---	---	--	--

Fry (2004) concluded that 'relative' was a key word in the results, as there could be no absolute measure of the factors given, only a positioning of one field relative to another in terms of disciplinary differences.

### Recommendations from recent studies

A number of user requirements analyses, or studies encompassing a degree of user needs questioning have been published in the last three years. Outcomes relevant to this study arising from other surveys and questionnaires are outlined below.

#### e-Science Gap Analysis<sup>11</sup> (2003)

This report, carried out at the University of Indiana, presented the results of personal interviews with 80 scientists in the UK, US and Europe concerning the current state of Grid and Cyberinfrastructure technology with respect to their use in e-Science. The guiding theme of each interview was the identification of issues affecting the development by 2006 of robust e-Science infrastructure that would be generally usable by both large and small groups wishing to set up or join virtual organizations.

This analysis focussed on "portal services", which are the Grid components which control the marshalling of information from a variety of Web or Grid services and allow the user to view and interact with them. This problem was viewed not just in terms of building a Web page, since the views of multiple Grid services need to be presented in a way that is easy for users and administrators to customise. The service architecture proposed resulted in the portal integrating, or more precisely aggregating, the individual fragments from content providers (or web services) into a web page. At the time of the study it was reported that there were some emerging technologies that should allow greater sharing of portal and problem solving environment components between different e-Science projects. A Problem-Solving Environment (PSE) is an integrated software environment for the computational solution of a particular problem, or class of related problems. PSEs are used in areas such as design optimization, parameter space studies, rapid prototyping, decision support, and industrial process control, as well as for support for large scale collaborative simulation and modelling in science and engineering, for which the ability to use and manage heterogeneous distributed high performance computing resources is a key requirement.

Further responses of relevance to the current study were focussed on access to data, and databases. The main recommendation for e-Research was that any user of data needs to know how it was produced. Although clearly needed, the methods for recording provenance were not deemed to be well codified and the Semantic Grid was thought to offer an important technology for specifying data lineage. Support for provenance is an essential requirement in an e-Science environment as data sharing is central to the basic concept of a virtual organization.

Further the study concluded that:

<sup>11</sup> <http://grids.ucs.indiana.edu/ptiupages/publications/GapAnalysis30June03v2.pdf>

- Provenance is key to establishing the quality, reliability, and value of data in the discovery process;
- The interface needs to present the views of multiple Grid services in a way that is easy to customize for users and administrators;
- There is a need to understand more fully the differences between e-Business and e-Science requirements.

The latter differences were identified during the study as security and workflow, but were thought to require further research.

### **Portals and Portlets workshop (2003)**

This workshop, and the ensuing report, focussed on the needs of the e-Science and Information Systems communities with respect to portals and portlets.

Conclusions relevant to this study included that:

- The easy delivery of access to both Grid resources and information services to end users involved in multi-disciplinary research and training is more important than ever;
- Research issues include ones related to data format diversity and others such as metadata, mapping and vocabulary;
- Developers should save effort by sharing services and methodologies and customising the existing Web-based presentation layers for delivery to all end-user projects;
- A range of toolkits (thin clients, portals, scripting languages, GUIs etc.) should be developed to extend and simplify access to Grid resources and information systems leading to the eventual emergence of one or more interfaces to a Virtual Research Environment.

With respect to e-Science and the needs of e-Researchers wanting to work collaboratively the following recommendation was made: “A portal for interactive collaborative working requires somewhat different services from a Grid application or resource discovery portal”, i.e. it requires:

- A mechanism to set up members (people, devices) in collaborative sessions;
- Generic tools: text chat, white boards - need shared updates to text message streams;
- Audio-video conferencing – to share events specifying changes in compressed streams;
- Applications such as web pages, presentations – all of which share key strokes to access master document;
- Visualisation – to share events corresponding to changes in pixels of a frame buffer, maybe using SVG;
- Shared maps, instruments (e.g. medical).

### **Loughborough University Library Portal study<sup>12</sup> (2004)**

The following couple of conclusions are drawn from an extensive final report arising from this study. It was of some concern to the Academic Librarians at Loughborough that the subscription to electronic resources were not being fully used and their belief, based on anecdotal evidence, was that students preferred to use Google and similar search methods rather than accessing the library web pages for their research.

One of the reasons for Loughborough investigating the concept of a library portal was this need to promote the use of their resources more widely, as well as creating an interface which presents resources attractively and is easy to use in a similar way to Google. A portal was defined in this case as an interface that provides an easily accessible facility which enables staff and students to search good quality sources more effectively.

---

<sup>12</sup> [http://www.lboro.ac.uk/departments/dis/lisu/downloads/LibPortal\\_final\\_report.pdf](http://www.lboro.ac.uk/departments/dis/lisu/downloads/LibPortal_final_report.pdf)

It was found that researchers often wish to conduct a quick search of all relevant databases and view the 'cross-searched' results, whereas academic librarians would prefer users to carry out detailed searches in the native interface of the database.

It was felt by many (researchers and librarians alike) that some researchers may need further training in the selection of databases and appropriate keywords.

### **EEVL survey (2004)**

This web-based questionnaire resulted in 349 responses from across the engineering and RDN user communities. The study report (summarised in Macleod, 2004<sup>13</sup>) concluded that user requirements for a portal for research in engineering were:

- access to full text resources
- training required in using online tools
- discovery tools needed
- more promotion of available services
- better communication of possibilities
- cross searching free databases
- cross searching subscription databases
- alerts (new publications, conferences, funding calls)
- more research (over teaching) resources
- links to text books
- different ways of using by different users (i.e. personalisation)
- more information related to industry for applied subjects
- filtering for certain subjects/ resources
- identify subscription based resources on results page
- identify whether user has access (i.e. local subscription) to resource
- intuitive interface
- access to departmental and local resources as well as external from one interface

Entrants were asked to rate six ideas for possible new EEVL services, and the results were not completely as expected. There was considerable support (80 ranking it a "Very good" idea, 113 "Good" and 107 "OK") for a service which allows searching a range of free databases from one place. There was slightly less support for the same type of service searching subscription databases, with 26 thinking this was actually a poor idea. Considering that they might be thought of as niche services, there was an unanticipated level of support for a journal Table of Contents (TOC) announcement service, a new book announcement service, and an events and conference announcing service, with only slightly less enthusiasm (165 OK, 75 Good, 40 Very Good) for a service which gave details of funding calls.

### **JISC Disciplinary Differences and Needs Report<sup>14</sup> (2005)**

This survey attracted 780 respondents, although it was noted in the report that the responses received represented less than 2% of the active research population (based on HESA 2004 statistics). The survey was carried out by Rightscom<sup>15</sup>.

The survey identified several important areas in which preferences and behaviour vary between disciplines. These include both the access to different types of information resources and the channels for communicating research results

The survey was designed to test some of the more widespread assumptions about the ways in which subject discipline affects information-seeking behaviour and modes of communication and collaboration. In addition, it specifically sought to investigate whether the resources provided by higher education institutions were differentially satisfactory for the different disciplinary groups in the context of the debate about whether journal 'big deals' have tended to favour scientists by squeezing funds for the acquisition of other subject journals or other types of material e.g. books.

---

<sup>13</sup> <http://www.ariadne.ac.uk/issue40/eevl/>

<sup>14</sup> [http://www.jisc.ac.uk/uploaded\\_documents/Disciplinary%20Differences%20and%20Needs.doc](http://www.jisc.ac.uk/uploaded_documents/Disciplinary%20Differences%20and%20Needs.doc)

<sup>15</sup> [www.rightscom.com](http://www.rightscom.com)

*'Do you encounter problems gaining access to the resources you need to carry out your research?'*

		Yes	No
Umbrella Group	Medical and biological sciences	52.5%	47.5%
	Physical sciences and engineering	42.4%	57.6%
	Social sciences	46.7%	53.3%
	Languages and area studies	48.0%	52.0%
	Arts and humanities	53.4%	46.6%
Total		47.7%	52.3%

Overall, there were mixed results for this question, with just less than half of respondents indicating that they can gain access to all the resources they need in order to carry out their research.

*What search tool or reference source is most essential to you, the one you would be lost without?*

% within Umbrella Group

	Umbrella Group				
	Medical and biological sciences	Physical sciences and engineering	Social sciences	Languages and area studies	Arts and humanities
Other	13.0%	5.7%	6.7%	8.0%	3.9%
Subject-specific abstracts and indexes	18.5%	20.6%	22.4%	6.0%	13.6%
Subject-specific online gateways	22.8%	3.3%	6.7%	2.0%	2.9%
General bibliographic resources	9.9%	11.5%	15.2%	46.0%	29.1%
Citation databases	21.0%	21.5%	9.9%	4.0%	3.9%
Search engines	14.8%	36.4%	35.9%	24.0%	36.9%
Works of reference		1.0%	3.1%	10.0%	9.7%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

Detail in the responses showed that there was a broad spread of discovery tools in use. Search engines were regarded as the "most essential" reference source by the physical and social scientists, and also in the arts and humanities, all at around 36%. Looking at the detail within the umbrella group for physical sciences showed that computer scientists, earth scientists, applied mathematicians and electronic engineers are more likely to view search engines as 'most essential' than are physicists. Search engines were also chosen by about half of sociologists and education specialists. Search engines seemed to be least significant in the medical and biological sciences, with only 14.8% of respondents citing them.

I had a crazy thought that this might be a function of how publically accessible the discipline is. Google has lots of links to stuff from public Web sites, Getting info from Google on cats is likely to be more productive than getting info on locusts. Does this mean that increasing the reliance on Google is leading to the enhancement of popular science at the demise of hard science???



Subject-specific online gateways were chosen by 23% of the medical and biological scientists, while 21% of this group named citation databases and 18% subject specific abstracts and indexes as some of the 'most essential' online resources. These were also 'most essential' to 21% of physical sciences and engineering researchers and 22% of social scientists.

General bibliographic resources were seen as 'most essential' by nearly half of language and area studies researchers and nearly 30% of arts and humanities.

*Access to research resources: What are the main problems you encounter?*

	Medical and biological sciences	Physical sciences	Social sciences	Languages and area studies	Arts and humanities
Library does not take the journals I need	89.4	82.0	81.9	79.2	69.1
Library does not buy the books I need	18.8	31.5	38.1	62.5	61.8
Library does not subscribe to the databases I need	22.4	36.0	34.3	16.7	32.7
I cannot get access to the conference proceedings I need	18.8	44.9	25.7	12.5	23.6
Key information is proprietary	10.6	12.4	17.1	8.3	12.7
I need to travel to access resources and funding isn't available	9.4	14.6	24.8	58.3	58.2

Other points of note in the report which are useful to this study were that:

- Disciplinary differences did not appear to be significant in relation to overall ease of access to resources, nor access to journals
- Researchers continue to use a range of bibliographic tools and resources off and online
- There are clear disciplinary differences in the spread of resources used, though journal articles cross most disciplinary boundaries in their importance
- A majority of researchers in all disciplinary groups were supportive of the idea that research funding bodies should mandate the deposit of research results in open archives
- Only a small proportion of researchers reported being involved in using online collaborative spaces – the highest at 13% for the physical sciences - or writing blogs.

**Australian e-Research – Building Partnerships (2005)**

This report, carried out by research support at University of Melbourne, asked "what is required to position Australia in e-Research?" It was felt that as e-Research changed the focus from high performance computing toward the application of advanced information and communication technology for research across a broader range of disciplines, disciplines which are text and media based, new know-how was needed.

Interviews with researchers identified the need for:

- Technical infrastructure and services
  - Terabytes of disc storage, network upgrades for handling large data volumes and video streaming, access grid nodes, middleware
  - Friendly grid portal, technical support for integration and implementation of databases and software within a virtual organisation environment, access to grid programming expertise, help in extending the tools, access to expertise in big databases.
- Leadership and coordination
- Need for an e-Research champion to bring together a virtual team across the university and beyond, help researchers understand what's possible, help people work on their own and collaboratively.

Interestingly, the information professionals highlighted matters such as “data mining” and “data preservation,” but those in the research communities did not recognize these areas as being beneficial to them.

What was found was a close correlation to the list developed by Markus Buchhorn (2004<sup>16</sup>) of researcher’s e-Research wants:

- Access to storage and computational resources
- Access to computational software and services
- Videoconferencing and collaboration tools
- User friendly application specific web based portals
- Shared access to large data repositories for searching, replication and updating
- Assistance with organising and managing their own research data sets
- Collaborative steering of remote research experiments and an ability to collaborate in international projects.

Buchhorn also noted that researchers want to access resources transparently; make data readily and appropriately available to others; and make collaboration easier. And further, went on to suggest that federated access was the future, for technical and political reasons.

Resource discovery and access implies:

- harvesters or registries
- portals or hubs
- virtualisation

as well as a demand for multistage repositories. Buchhorn suggested that the most common types of queries and searches would be:

- Metadata-based
- Objective, scientific
- Subjective, annotation
- Geospatial
- Chronological

In conclusion:

- One size (infrastructure) does not fit all (yet/ever); But 3-4(0) sizes may fit most (for now)
- Some domains have very different definitions of sustainability, rights issues, data movement needs; but many don’t or should...
- User and developer education is still needed

### **CREE – Contextual Resource Evaluation Environment project, University of Hull (2005<sup>17</sup>)**

The CREE project was structured around two main goals in response to the needs users may have for systems providing search functionality:

- The investigation of user requirements for the presentation and delivery of search tools through a variety of institutional environments and contexts.
- The investigation of the JSR 168 and WSRP portlet standards to allow the presentation of existing search tools within conformant portal frameworks (e.g., uPortal<sup>18</sup>).

The user requirements strand, which included a survey, focus groups, and user testing, resulted in a validated set of user requirements for the presentation of search tools within portal and non-portal contexts.

The survey and focus groups found that Google is the most popular approach to searching. However, library resources are used and recognised for what they are. The user is looking for a quick return on a search and is working with limited time. Quick access can be followed by

---

<sup>16</sup> Buchhorn, M. (2004) Research Data Meets ‘The Real World’  
<http://www.aarnet.edu.au/engineering/middleware/archive/middle/2004/forum/Buchhorn.pps>

<sup>17</sup> <http://www.hull.ac.uk/esig/cree/downloads/CREEsummaryoverview.pdf>

<sup>18</sup> <http://www.uportal.org/>

more detailed use if a user perceives they have found results of value from a search tool. The user testing sessions reinforced this finding and also found users welcome the idea of using search tools within different institutional environments. The presentation of subject-based search tools within an online learning environment/VLE gathered the most enthusiastic response, associating searching with learning activities. The use of search tools in different environments was particularly welcomed in the case of search tools that the users were not previously familiar with: bringing the search tool to the user made them aware of it and provided them with a convenient local access point.

In terms of user requirements the recommendations were that:

- Users want to be able to access search tools more easily and directly. This mimics the style of Google (from where the idea mainly comes)
- Users will investigate the further potential of a search tool if they have had initial success when using it.
- Users are open to the idea of search services being delivered within different environments, and indeed welcome this. Careful selection of which tools are presented can help guide the user in their searching, with subject-based tools favoured by users.
- Access to communication tools through a portal was also welcomed, though there needs to be a clearly defined reason for accessing them via this route as opposed to direct access.
- User feedback is valuable to feed into technical development and essential as part of the development cycle. The use of a range of methodologies helps to validate findings and provide explanations for views held.

And in terms of technical development the project team found that:

- Portlets within an institutional portal offer potential, but require further work to fully develop services that users feel comfortable using.
- The portlet standards are better able to support asynchronous communication.
- There is recognition that it is possible to adapt search tools for use within portlets, but there are reservations about the efficiency of this that would have to be borne in mind in each individual case. The support of search services delivered via this route is equivalent to supporting the original search service via the native website.
- For search services WSRP appears to be the best standard to follow because of its ability to present remote services within the portal framework.

It was recommended that the HE/FE community consider alternative ways of delivering search services in different institutional environments to suit the needs of users. This recommendation was made in the light of the positive response gathered through CREE to such possibilities.

### **South African e-Research Portal (Page-Shipp *et al.*, 2005<sup>19</sup>)**

The University of Pretoria has undertaken the development of, first, an institutional portal and latterly a research portal, with a focus on the needs of e-research. Interviews with staff across the university were carried out between 2001 and 2003, during the development and testing of a prototype institutional portal. Development took the stance that staff needed:

- High levels of functionality and integration - a seamless interface.
- Advanced personalisation and customisation capability
- A portal which supports both the teaching and research roles of academics.

The portal was also developed to give access to information sources, such as e-journals, e-articles, e-reserves, e-archives, databases, e-books, e-dissertations, library catalogues, and local research databases; as well as personal sources of data, including subject experts and information specialists.

And the following functionality was indicated by the staff interviewed during the development:

---

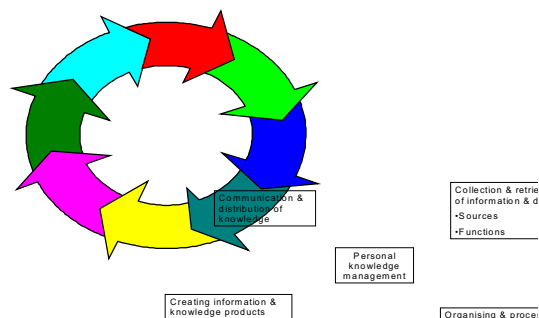
<sup>19</sup> Page-Shipp, R.J., Hammes, M., Pienaar, H., Reagon, F., Thomas, G., van Deventer, M.J., Veldsman, V. 2005. *eResearch support services: responding to a challenge facing the South African research and information communities*. *SA Journal of Information Management*, December 2005, 7(4) <http://www.sajim.co.za>

- Web search engines,
- Global search function,
- Listservers, chat rooms, e-mail,
- Adding of URLs,
- Interface with document delivery and inter library loan systems.
- Ability to evaluate and add information sources to the portal

The figure below illustrates an e-researcher's management of information

Has something gone wrong with this figure?

### Personal academic knowledge management: needs



Finally, the interviews also demonstrated a need for tools and services regarding the processing, creating & communication of data, information & knowledge products, such as:

- Indexing tools
- MS tools
- Support e-publishing
- Virtual conferences
- Web teaching: virtual classrooms, demonstrations and lectures
- Virtual work space for research projects

With regard to a portal for research (and a focus on e-research), further interviews found that the portal must offer the researcher access to:

- A single, user-friendly, access point to a family of repositories for data, digital objects and publications that would act, inter alia, as a record of research outputs.
- Easy communication with colleagues and other interested researchers
- Online content, including a standard suite of online commercial information resources, paid for nationally; other resources that are available without payment; commercially published resources, in addition to the above, to which the researcher's institution subscribes; a pay-per-view facility for any other resources
- Assistance in finding, accessing, contributing to and using databases that are needed for research.

In terms of functionality, the research portal would include:

- A global search engine that can search simultaneously all those sources that are accessible to a particular researcher
- 'Push' of references to published information – Commercial and Open Access according to individual profiles
- Virtual Communities of Practice/Curiosity
- Assistance in submitting large research databases/data streams to SA National Research Network for transmission to colleagues and co-workers anywhere in the world, and in receiving such databases/streams
- Online research support tools, for data and information/reference management

The portal will be developed to have the following characteristics, some of which are generic:

- Built in authentication
  - Customisation
  - Access to scholarly communication results (e.g. articles, theses, blogs)
  - Research data / data curation /metadata – national digital repository(ies)
  - Support of research communication and collaboration (e.g. virtual groups)
  - Research administration functions including funding opportunities;
- and others specific to the portal developed for South African e-research:
- Micro-payment (Pay per view) option
  - eTraining/eLearning to obtain full functionality
  - Global search engine (e.g. federated search engine and link resolver)
  - Online research tools e.g. models, databases and personal information management
  - Personal research sources (human)

In conclusion Pienaar states that in his experience, user requirements were found to be similar at the University of Pretoria to those reported in the UK<sup>20</sup>. He also comments that open source products and open standards should be used wherever possible in developments of this type.

### **EVIE project, Leed University (2005<sup>21</sup>)**

EVIE is a JISC-funded project which is aiming to develop a virtual research environment for staff based around portal technology. The methodology for the user needs analysis phase included structured one-to-one interviews with research staff, a series of focus groups with staff from the Schools of Geography and Medicine, and an online survey; 121 staff members completed the survey.

*“Researchers in all disciplines are increasingly expecting to be able to undertake research-associated tasks online. These range from collaborative activities with colleagues around the globe through to information-seeking in an electronic library environment. Many of the tools which enable these activities are already available within the local IT infrastructure. However, in many cases, the tools are provided through discrete, bespoke interfaces with little inter-linkage. Researchers face a number of challenges in this environment. The EVIE Project will address these issues by testing the integration and deployment of key software components within a portal framework.”* (EVIE project website: <http://www.leeds.ac.uk/evie/>)

The main aim of EVIE was to establish a prototype Virtual Research Environment (VRE) to support the needs of researchers. As such, the user needs analysis aimed to elicit views from the research community as to which areas of research activity and support the EVIE project should develop, including key areas for integration within the University’s portal framework and the content and applications that should be included.

Researchers were keen to see a unified interface, flagging that this needs to be easy to use and available to them from off-campus. Different aspects of research activity were prioritised as follows:

- Finding and acquiring published information such as articles, proceedings, monographs, etc.
- Finding out about funding opportunities; applying for funding.
- Collaborating with partners within the University or elsewhere.
- Sharing or archiving research results; improving permanence of outputs.
- Other activities.

Within these activities, respondents indicated the following issues and needs:

<sup>20</sup> He references the EVIE project in particular

<sup>21</sup> Sergeant, D. M., Andrews, S., Farquhar, A. [2005]. Embedding a VRE in an Institutional Environment (EVIE) & Workpackage 2: User Requirements Analysis User Requirements Analysis Report. [http://www.leeds.ac.uk/evie/workpackages/wp2/evieWP2\\_UserRequirementsAnalysis\\_v1\\_0.pdf](http://www.leeds.ac.uk/evie/workpackages/wp2/evieWP2_UserRequirementsAnalysis_v1_0.pdf)

- *Finding information:* Respondents want a simple, easy tool to enable them to search across several datasets; they want advanced search interfaces too; and want to build up their own search strategies.
- *Funding:* Respondents would like to see various aspects of the grants management process simplified; they want to be able to search and view previously submitted proposals; and they would like the portal environment to alert them about new funding opportunities.
- *Collaboration:* Respondents want access to their own email within the portal, alongside the ability to share diaries and meeting organisers; they want tools which enable them to work collaboratively on documents and large files; and they would like to be able to find out who has what expertise.
- *Research outputs:* Respondents want to be able to find the full-text of an output from the University publications listing, and processes in place to allow upload of their full-text to the system.

Other activities flagged as important include monitoring of financial expenditure; and facilities for booking meeting rooms and other resources.

The outcomes of the user requirements analysis suggested that the initial priorities for incorporation of systems, tools, and information into the EVIE portal would include: a single search interface to find information; tools to support grants submission; facilities for collaboratively managing documents; access to email and diaries; and facilities for uploading full-text documents to a repository.

#### **e-Research Needs Assessment University of Iowa<sup>22</sup> (2006)**

The 'E-Research Project' was a recent exploration of the research computing environment at the University of Iowa and was undertaken as part of a campus-wide information technology review.

Sometimes referred to as cyberinfrastructure, the environment investigated included the hardware, software, staffing, and facilities required to support research, scholarship, and creative activities.

Several factors led to conducting this assessment, which were given as national trends, faculty retention, broadening scope of demand for research computing, requirements from national funding agencies and keeping the university competitive. IT staff at Iowa felt that the development and management of the University's research computing infrastructure is a strategic issue that called for a coordinated, university-wide plan.

The E-Research team collected views and opinions from across the campus, and from staff in a variety of research areas, including all those individuals who were identified as carrying out e-research. In this case, this type of cutting-edge research has been defined to increasingly require distributed access to very large data archives, sophisticated information mining, and flexible data presentation methods.

While there was no typical research computing user or single research computing profile at Iowa, many of the researchers interviewed expressed a desire to interact to a greater extent with their research peers around campus.

Many of those interviewed were confused about where they might go for assistance in dealing with specific research computing issues. A strong and consistent message was the desire for better communication, coordination, identity, and leadership.

Researchers from every area of campus expressed the need for help in managing data. The amount of unique and irreplaceable data generated is massive and continues to grow at an ever-increasing rate. The need for robust data storage is substantial and immediate. However,

---

<sup>22</sup> IT Support of Research Subcommittee Report: *E-Research Needs Assessment* version Draft .931 January 30, 2006 For The University of Iowa

effective data management was also required, and demands an integrated technical architecture supporting the complete life cycle of digital information. Researchers wanted a total solution for data management including data backup, data cataloguing, data preservation, and access to staff with a variety of data management skills.

Researchers reported that their world is becoming increasingly collaborative and they often participate in projects that involve scholars at a number of different sites. They also voiced the need to share research data and a desire to use collaborative environments such as the Access Grid<sup>23</sup>, Elluminate<sup>24</sup>, and other technologies to enhance their collaborative efforts.

Based on this input, the E-Research team identified the following needs that researchers thought were critical to a vibrant research computing environment at the University:

- A campus research computing community
- Data management resources
- Collaboration resources
- More security for research computers
- A research computing infrastructure

Which of these should be in scope of IE?

### **ESRC RIESS (2006)**

This recent review of the information environment for social science researchers was commissioned by ESRC in response to widespread concerns and uncertainties about the rate of change in the environment within which social and economic researchers find and access information crucial to their work. It was written between September 2005 and January 2006, by Nicky Ferguson and Seb Schmoller, with guidance from an ESRC Advisory Group chaired by Lynne Brindley. The review has not been publicly released, and this project is grateful to the authors for sight of the report for reference in this study.

The report records results from:

- a web survey of the economic and social research community, including users, to which 342 people responded;
- 47 interviews with figures within that community, from a wide range of roles and disciplines, academia, the private sector, government and the voluntary sector;
- a focus group at the British Library involving 25 government researchers;
- a Future Look email exchange with 22 research, information and technology experts;

Extrapolating from the recommendations from the report, there were indications that:

- access to 'hidden' resources/ records
- indexing of library catalogues for text mining
- information skills training
- access to all aspects of information and data about funded research, including PhD theses
- improved access to subscription based resources; includes short term access/ access for all individuals funded by (e.g.) research councils/ subsidies/ government funded public access to key journals (c.f. Iceland)
- catalogue/ database for grey/ un-digitised literature
- digitise more resources
- printed resources form a central part of the information environment, particularly for social science researchers
- maintenance of existing services (IBSS/SOSIG)
- consideration of social book-marking systems

---

<sup>23</sup> Elluminate - <http://www.illuminate.com/>

<sup>24</sup> Access Grid - <http://www.accessgrid.org/>

would be useful in the context of carrying out social science research using online resources.

Also, although users want access to data, they need incentives to participate in open access publishing. There were repeated references to inequities of information access within sectors and between sectors to the extent that the study concluded there are “a multiplicity of information environments for social and economic researchers, rather than just one”. The access problems primarily involve subscription journals and databases, but can also include slow and restricted internet connections. Researchers at well-resourced institutions with a good infrastructure found fewer problems than those at small or less research-focussed institutions.

Study found that research staff are the most satisfied with the adequacy of the information environment. 90% of respondents (all social scientists) regularly used Google, 33% used Web of Science, and 26% used Google Scholar, 25%, IBSS. Other resources (e.g. SOSIG/ Census data/ COPAC) were used regularly by less than 20% of the respondents. When expanded to occasional use, Web of Science, Google Scholar, IBSS and SOSIG are used by c. 45% of respondents. In research support the recognition and use pattern is different, with over 70% making use of the resources named above and COPAC.

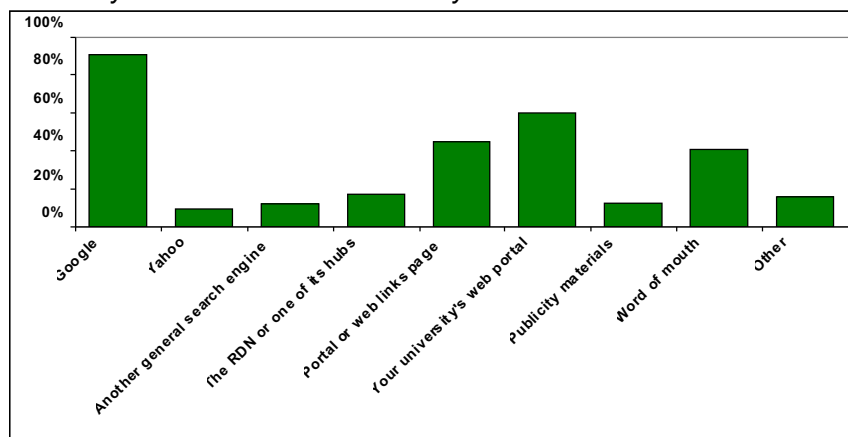
The report noted that paper (books, journals, printer output, grey literature) remains a preferred medium for many researchers. However, access to these hard copies is increasingly mediated by a digital search, hence the creation of digitised catalogues will be important for resources which are not likely to be digitised themselves.

#### **RDN (Intute) Researchers Survey (2006)**

Since the Resource Discovery Network service is rebranding in 2006, a web based survey was released aimed at establishing what kinds of research respondents engaged in, their current resource discovery techniques, and what services they would like to be able to access. The survey resulted in 355 respondents.

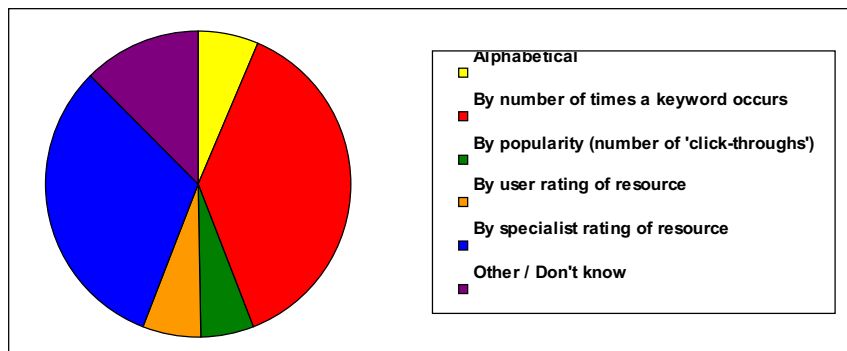
The survey was intended for people actively engaged in research at postgraduate level or above. The first few questions would have made this clear and should have dissuaded any accidental participants from continuing.

*“How do you find online resources for your research?”*



*“Which of the following methods of ranking search results would you prefer?”*

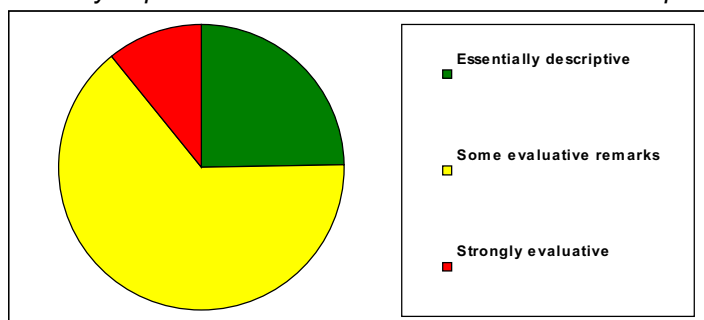




A typical quote: “depends on the type of resource! Provide all the above options” (Post-doctoral physics researcher, University of Oxford)

The RDN uses description for each resource, and this was felt to be an important factor in the value of the service: “the presence of a brief description (less than 100 words) I regard as particularly useful” (Researcher)

“Would you prefer our reviews of online resources to be primarily descriptive, or evaluative?”



“How useful do you think each of the following features would be?” – all the features given were rated as ‘useful’ apart from a list of resources specifically relating to project management and administration. Bibliographic database directory and journal directory (evaluative) were felt to be ‘Very Useful’.

Rank	Service
1	A directory of bibliographic databases in each subject
2	A directory of journals, with indicators of subscription numbers and prestige
3	A database of sources of funding and scholarships
4	A Directory of regular conferences, with indicators of size and prestige
=5	Searching and browsing filtered by the time period which the resource concerns
=5	Ability to filter out resources not directly applicable to research
7	Online events calendars - for conferences and other events in each subject area
8	National directory of researchers and research
9	Data that can be exported for use on other Web sites, or for other uses elsewhere
10	List of resources for graduate skills training - publishing, networking, thesis writing, etc.

A further oft repeated comment on the way in which resources are presented was the request to “flag up whether web resources are free or subscription-based” (*English Literature Researcher, Edinburgh University*).

## Conclusions

Linking research practice, resource discovery and information retrieval needs an environment into which they are all integrated. Previous surveys have taken too narrow a view of this, since they have mostly been discipline specific or have focussed on one aspect of this activity. The

joint space requirements still need further investigation, i.e. computing and collaboration, or personal information management and admin functions.

However, with regard to e-research, some key conclusions can be drawn out from previous studies, including that:

- Researchers need access to data storage and computational resources, as well as software and services.
- Provenance is key to establishing the quality, reliability, and value of data in the discovery process;
- Any interface needs to present the views of multiple Grid services in a way that is easy for users and administrators to access and customise;
- There is a need to understand more fully disciplinary differences in user requirements; e-research issues which have been raised in previous studies are related to data format diversity as well as metadata, mapping and vocabulary;
- Existing services and methodologies could be shared and Web-based presentation layers customised for delivery to users;
- A range of toolkits (thin clients, portals, scripting languages, GUIs etc.) should be developed to extend and simplify access to Grid resources and information systems leading to the eventual emergence of one or more interfaces to a Virtual Research Environment.

With the development of e-research groups new needs appear to have emerged. It is likely that the needs that will be important for a given institution will vary by the:

- Areas of research strength
- Extent of infrastructural development
- Strength of collaborative networks

A portal for e-research is likely to require:

- A mechanism to set up members (people, devices) in collaborative sessions; to include shared access to data repositories for searching, replication and updating
- Generic tools: text chat, white boards - need shared updates to text message streams;
- Audio-video conferencing and collaboration tools – to share events specifying changes in compressed streams;
- Applications such as web pages, presentations – all of which share key strokes to access master document;
- Visualisation – to share events corresponding to changes in pixels of a frame buffer, maybe using SVG;
- Shared maps, instruments (e.g. medical)
- Access to full text resources as well as tools to cross search both free and subscription based services; the results pages should identify subscription based resources and whether the user can gain access to them
- Alert services, promotion and training opportunities on how to use services accessed from the portal.
- Access to departmental and local resources and repositories as well as external resources from one interface.

Though we note that this would need further testing with users. Any provision also needs to be supported by assistance with organising and managing research data sets, as well as a training programme.

It is worth remembering that “a portal is not a repository, and a repository is not a portal” (Burnhill, 2006<sup>25</sup>), even though all repository facilities are likely to have websites, these are not necessarily portals. Also, what is functionally useful for users is different to a portal’s functionality; or, in other words, a portal could do without a repository if users are depositing in

---

<sup>25</sup> Peter Burnhill, EDINA, personal observation, May 2006

institutional and other digital repositories. In terms of e-research, the value of a portal could be as an interface to create and support a community of users, a space that people use for their collaborative research. However, the portal must point at some repository or other, or it loses some of its usefulness.

Overall, researchers appear to need more support for learning, adapting, and writing software specific to their research problems. Also, researchers who are generating and using large data sets need help managing their data. This need will become more pressing as data it enters long-lived data repositories and therefore the public arena through presentation rather than publication.