

# JISC Infrastructure for Education and Research

## BRIEFING PAPER, October 2010

### Contents

<b>Introduction</b> .....	<b>2</b>
<b>Identity management pilots</b> .....	<b>3</b>
Toolkit Aims .....	3
Use Cases .....	3
Toolkit Sections .....	3
<b>Research information management</b> .....	<b>6</b>
Context.....	6
Related work .....	6
<b>Identifiers</b> .....	<b>8</b>
Context.....	8
Key resources.....	8
Related work .....	8
<b>Geospatial</b> .....	<b>11</b>
Context.....	11
<b>Infrastructure for resource discovery</b> .....	<b>15</b>
<b>Activity data</b> .....	<b>18</b>
Context.....	18
Glossary .....	18
Related work .....	19
<b>Repositories: take-up and embedding</b> .....	<b>21</b>
Context.....	21
Specific Examples:.....	23
Working with organisations.....	26
<b>Preservation tools</b> .....	<b>27</b>
Context.....	27
Preservation Tools .....	27
Preservation Activity .....	27
<b>Preservation of complex visual digital materials and environments</b> .....	<b>29</b>
Context.....	29
Related Work .....	29
Strand Scope and Related Work .....	29
<b>Sustaining ‘at risk’ online resources</b> .....	<b>32</b>
Context.....	32
Relevant links.....	32

## **Introduction**

This briefing paper is intended to support bidders drafting proposals against the October 2010 JISC Grant Funding Call 15/10 for Infrastructure for Education and Research. Bidders are strongly advised to read at least the relevant sections of this briefing paper alongside the main Call document. This document is intended to offer further information relevant to the programme areas in the Call; in all cases, the terms of reference given in the main Call document are definitive.

A 'Frequently Asked Questions' (FAQ) area will be set up on the 'Information Environment' blog, at: <http://infteam.jiscinvolve.org>, where prospective bidders can both review questions asked and answered previously, and raise questions of their own.

# Identity management pilots

## Toolkit Aims

The Identity Management Toolkit Project<sup>1</sup> ran from 1 January 2009 to 30 June 2010 and was based on direct experience gained by project partners from work undertaken in The Identity Project<sup>2</sup>. The aim of the Toolkit project was to produce downloadable and online versions of the Toolkit for use by executive and technical staff in Further and Higher Education institutions who wanted to review, assess and improve the performance of Identity Management in their organisations.

A secondary aim of the project was to support efforts by JISC to raise and maintain awareness of the importance and key issues of Identity Management for the UK academic community, particularly with respect to:

- business and process efficiencies
- good information governance and legal compliance
- participation in the UK Access Management Federation for Education & Research
- awareness of UK and international standards and trends in Identity and Access Management

The Toolkit was launched at the UCISA<sup>3</sup> and JISC<sup>4</sup> annual conferences in March/April 2010. The availability of the Toolkit has been highlighted in a number of publications and presentations at events relevant to the target audience. There has been a high level of interest in the toolkit both nationally and internationally.

## Use Cases

The Toolkit was field-tested during the project by use in institutional projects at a large university and a FE college, with detailed feedback from staff involved in those projects being used to refine the final versions of the Toolkit. Bidders may want to refer to the case study ("*Case Study: Identity Management at Cardiff University: Membership, Categories and Entitlement.*"<sup>5</sup>), on the Identity Management Toolkit website for further information. Although Cardiff was not one of the field-test sites for the Toolkit (and this use case was a supplementary piece of work produced after the Toolkit was published), it provides an example of a possible template for collecting use cases. Other formats including multimedia formats will be accepted.

## Toolkit Sections

The following sections are included within the Toolkit. JISC has not ranked these by importance and all bids should address one or more of these areas:

**Identity Management Governance and Policies:** Describes the roles, structures and policies required for Identity Management and how they relate to Identity Management systems and processes.

---

<sup>1</sup> <http://www.jisc.ac.uk/whatwedo/programmes/aim/idmtoolkit.aspx>

<sup>2</sup> <http://www.Identity-Project.org>

<sup>3</sup> <http://www.ucisa.ac.uk/events/2010/conference2010/programme.aspx>

<sup>4</sup> <http://www.jisc.ac.uk/events/2010/04/jisc10/programme/identity.aspx>

<sup>5</sup> <https://gabriel.lse.ac.uk/twiki/bin/view/Projects/IdMToolkit/UserCaseCardiffJul10>

**Identity Management Systems, Components and Functions:** The technical components and functions of Identity Management systems in an academic institution.

**Defining Institutional Requirements:** Functional requirements for each component of an Identity Management system, which may be useful in defining the objectives of an in-house implementation project or in detailed specifications to suppliers.

**Discovering and Auditing Current Institutional Identity Management:** A detailed guide to assessing the state of Identity Management in an institution with a comprehensive audit (based on work of the JISC Identity Project which developed and tested Identity Management audits in several universities).

**Gap Analysis:** Explains how to establish the current and desired states of Identity Management, gives a list of common gaps in Further and Higher Education Institutions, and suggests ways for developing a strategy.

**An Institutional Roadmap for Identity Management:** Producing an overall roadmap or programme plan. Prioritising major deliverables and milestones by achievability, cost and institutional impact.

**Designing and Managing an Identity Management Project:** Project management issues particular to implementing Identity Management, including key institutional benefits of improved Identity Management for use in an institutional business case.

**Selecting Supplier Solutions:** Where commercial procurement of systems or components is required, this section aims to help understand the Identity Management system solutions available, produce procurement criteria, and construct tender documents.

## Related Work

Although the Toolkit website is the main resource for information on this call, bidders should be aware that JISC Collections is funding a project that is looking at extending Access Management and Identity Management to external Business Community Engagement (BCE) partners. This project is based on the Extending Access Management report<sup>6</sup>. As this project is required to look at the Toolkit developments, successful bidders should be aware of the work being undertaken in this project and communicate with them as necessary. The project details are available from <http://www.jisc-collections.ac.uk/Our-projects/EAM2BCE/>.

## Further information

The Toolkit project has set up contact points for further engagement with the JISC community specific to Identity Management issues, including the website, an email enquiry point ([jisc-identity-management@jiscmail.ac.uk](mailto:jisc-identity-management@jiscmail.ac.uk)) for direct (private) communication with JISC support for Identity Management, and a public email discussion list ([identity-project-public@jiscmail.ac.uk](mailto:identity-project-public@jiscmail.ac.uk)). To contact the creators of the Toolkit, please use the [jisc-identity-management@jiscmail.ac.uk](mailto:jisc-identity-management@jiscmail.ac.uk) mailing list.

The Access and Identity Management (AIM) Programme uses the tag #jiscaim. Further information supporting this call will be available via the AIM blog

---

<sup>6</sup> Oakleigh Consulting (2009) Extending Access Management into Business and Community Engagement – Scoping Study: <http://www.jisc.ac.uk/media/documents/programmes/bce/extendingaccessmanagementreport.pdf>

<http://aimprog.jiscinvolve.org/wp/>. A collection of all tagged content is available at the Netvibes AIM page <http://www.netvibes.com/jiscaim#General>, where you can read blog postings, tweets and information on other AIM funded projects. If any bidder wants to comment, or raise an issue, during the proposal writing phase they are encouraged to do so and to use the above tag.

# Research information management

## Context

Research information is structured information about projects, researchers, research outputs and their value and impact, research impact more widely, funding streams, organisations such as funders and universities, and so on. A general background is available on the JISC website<sup>7</sup>. The intended outcome of work in this area is to bring some further coherence to data exchange activities between and within UK research stakeholders, with funding aimed at helping to reduce duplication, wasted effort and burden and to help develop effective partnerships across stakeholder groups.

## Related work

This programme area is based on a number of key resources, namely:

- JISC InfoNet has compiled support resources relevant to research information management: <http://www.jiscinfonet.ac.uk/research>.
- UKOLN offers technical support to the sector in this area: <http://www.ukoln.ac.uk/rim/>.
- The JISC EXRI (Exchanging Research Information) Report<sup>8</sup> examined various scenarios for the exchange of research information, and recommended that CERIF 2008<sup>9</sup> be the basis for the exchange of research information in the UK.
- This recommendation has been accepted by a cross-sector representative group, whose conclusions are summarised in a JISC Briefing Paper on Research Information Management<sup>10</sup>.

Evidence has been compiled on the business case for the adoption of CERIF in the UK higher education sector. The conclusion of this study is that such evidence is compelling and that the case is strong<sup>11</sup>.

A growing body of recent work has been undertaken based on CERIF. A document summarising this work is available<sup>12</sup>. This document is a key resource and it is strongly recommended that bidders read it closely in preparing proposals.

Note that HEFCE have stated that the Research Excellence Framework (REF) will support the 'collect data once, use many times' principle. Research income and research student data will be aligned as far as possible with data collected by The Higher Educational Statistics Agency (HESA). Data on research staff and outputs will be broadly similar to that collected in the RAE. HEFCE has indicated that REF submissions may be made using the CERIF format.

Projects engaging with national and international infrastructure developments are encouraged. Such projects are likely to lead to a more flexible and standards-based

---

<sup>7</sup> <http://www.jisc.ac.uk/whatwedo/themes/informationenvironment/researchinfomgt.aspx>

<sup>8</sup> Rogers, N and Huxley, L and Ferguson, N (2010) Exchanging Research Information in the UK <http://ie-repository.jisc.ac.uk/448/>

<sup>9</sup> <http://www.eurocris.org/>

<sup>10</sup> <http://www.jisc.ac.uk/publications/briefingpapers/2010/bpexriv1.aspx>

<sup>11</sup> Bolton, S (2010) The Business Case for the Adoption of a UK Standard for Research Information Interchange <http://ie-repository.jisc.ac.uk/487/>

<sup>12</sup> Russell, R and Rogers, N. (2010) Research Information Management in the UK: CERIF and metadata alignment: <http://www.ukoln.ac.uk/rim/dissemination/2010/rim-cerif.pdf>

infrastructure that is able to support varied and evolving requirements. Such developments include (but are not limited to):

- The pilot UK name authority service Names <http://names.mimas.ac.uk/>. Contact Amanda Hill [amanda.hill@manchester.ac.uk](mailto:amanda.hill@manchester.ac.uk)
- Potential development of a shared research management and administration system <http://as.exeter.ac.uk/rmas/>. Contact Deborah Welland [D.Welland@exeter.ac.uk](mailto:D.Welland@exeter.ac.uk)
- Research Council award and output systems. Contact Dale Heenan [dale.heenan@esrc.ac.uk](mailto:dale.heenan@esrc.ac.uk), Robert Sanderson [RHS@nerc.ac.uk](mailto:RHS@nerc.ac.uk)
- UK PubMed Central <http://ukpmc.ac.uk>. Contact Kevin Dolby [k.dolby@wellcome.ac.uk](mailto:k.dolby@wellcome.ac.uk)
- OpenAIRE <http://www.openaire.eu>. Contact Mikael Elbaek [mke@dtic.dtu.dk](mailto:mke@dtic.dtu.dk)

Please note, that with regards to integrating research management, administrative and repository systems, bidders should read both the “Repositories: take-up and embedding” and “The Research Information Management” areas of the call document and briefing paper, and then decide which is more appropriate for their proposal to bid against.

# Identifiers

## Context

In March of 2001 JISC commissioned a TechWatch Report into the use of CMS systems for University Web Sites<sup>13</sup>. Since then we have seen an exponential increase in the number of institutional staff who have editorial control of the public-facing web pages. This has been one factor in a significant growth in the number of .ac.uk web pages created each day. While there is value in allowing public .ac.uk web pages to be created by institutional staff (especially highly-rated teachers and researchers) for the purposes of increasing the online visibility and reputation of the institution, there is also a need by the institution to have common structures and vocabularies by which to organise and maintain the overall expansion of the public-facing website as it grows in size and amasses more legacy content. One key aspect of that organisation is the use of identifiers for web pages and data.

The overall aim of the Identifiers programme area of work is to improve the extent to which identifiers for public .ac.uk websites are planned and managed within institutions, and contribute to the technologies and skills required to do that. These improvements may be simply better use of structured URIs for web pages, or may extend to the presentation of institutional information as linked data that will allow it to be easily used internally and externally to the institution<sup>14</sup>.

## Key resources

The following are key resources for bidders to consult in preparing any proposal:

1. There is a useful glossary of terms available from UKOLN:  
<http://www.ukoln.ac.uk/jisc-ie/blog/identifiers/identifiers-quick-reference/>
2. To help clarify the problem space JISC had a community consultation conducted into persistent identifiers, and the definitions and meeting notes from that meeting are of value: <http://identifiers2010.jiscpress.org/>
3. The Cabinet Office has released guidance on designing URI sets for public sector websites: [http://www.cabinetoffice.gov.uk/media/301253/public\\_sector\\_uri.pdf](http://www.cabinetoffice.gov.uk/media/301253/public_sector_uri.pdf)

## Related work

Please note that all of the example links below are listed under the following “jiscPID” tag: <http://www.delicious.com/tag/jiscpid>

The following relate to the “Objectives” section of the Call document (the Call document should be read prior to reading this Briefing Paper). The examples below provide context for these objectives. These examples are not prescriptive and are only an attempt to provide scope for the problem space. Proposals should not be restricted by these examples as they are only intended as a guide to previous work that has been done in the problem space.

---

<sup>13</sup> Browning, P. and Lowndes, M. (2001) Content management systems  
<http://www.jisc.ac.uk/whatwedo/services/techwatch/reports/horizonscanning/hs0102>

<sup>14</sup> <http://linkeddata.jiscpress.org>

Projects will need to engage a range of stakeholders from within the institution and make sure there are common communication methods for discussing the structuring of the university website or part's thereof.

- There have been experimental work done at UKOLN on the pragmatics of utilising linked data in a content management system: <http://www.ukoln.ac.uk/jisc-ie/blog/2010/09/15/consuming-and-producing-linked-data-in-a-content-management-system/>
- The Guardian has reported on how it has addressed structuring of data for journalism for its web platforms: <http://www.guardian.co.uk/news/datablog/2010/aug/10/government-data-information-architecture>
- JISC Strategic Content Alliance has published this report on the value for institutions increasing Search Engine Optimisation which persistent identifiers supports directly: <http://sca.jiscinvolve.org/wp/2010/01/16/download-new-seo-report-with-case-studies/>

Projects will need to be aware of how they can progress their proposed structure for the University's public web pages, some examples of how to track that progress are provided below:

- Tim Berners-Lee's 5 Stars of Linked Data: <http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/>
- The 4 Steps of structuring data as adopted by the Resource Discovery Taskforce: <http://www.ukoln.ac.uk/jisc-ie/blog/2010/08/19/aggregation-and-the-resource-discovery-taskforce-vision/>
- Paul Walk's reflection on his experience as a technical manager: <http://blog.paulwalk.net/2010/09/21/institutions-and-the-web-done-better/>
- Microsoft's John Udell talks about the value of re-using part of other organisations' identifier structures: <http://blog.jonudell.net/2009/08/31/the-joy-of-webscale-identifiers/>

Projects are encouraged to build on pre-existing structured vocabularies where possible. Please note that many of these vocabulary structures are still in draft and are subject to change. Projects are not expected to use OWL/RDF, but rather they should consider the list of vocabulary terms as part of their URI structure. Proposals should actively discuss the use of any vocabulary as part of the project (see above):

- Vocabulary structure for describing contact details for people: <http://www.foaf-project.org/>
- Vocabulary structure for the creation of researcher profiles: <http://vocab.ox.ac.uk/res/researchers.htm>
- Vocabulary structure for research grant funded projects: <http://www.dotac.info/mapping/>
- Vocabulary structure for describing online communities and their participants <http://sioc-project.org/ontology>
- Vocabulary structure for describing an organisation and the things it provides as part of its business: <http://www.heppnetz.de/projects/goodrelations/>
- Vocabulary structure for course related information: <http://www.xcri.org/>
- Vocabulary structure for events: <http://programme.ecs.soton.ac.uk/1.0/>
- Vocabulary for writing and annotating scientific publications: <http://salt.semanticauthoring.org/>
- Vocabularies for describing higher and further education institutions (UCAS, HESA, Athens): <http://www.jiscmu.ac.uk/news/view/189>
- Vocabulary for creating time based structures for web pages: <http://www.mementoweb.org/guide/>
- Vocabulary for the provenance of resources: <http://openprovenance.org/>
- Various vocabularies used with bibliographic metadata: [http://www.w3.org/2005/Incubator/ld/wiki/Library\\_Data\\_Resources](http://www.w3.org/2005/Incubator/ld/wiki/Library_Data_Resources)

- Vocabulary for describing basic resources: <http://dublincore.org/documents/dc-rdf/>

Projects will need to be mindful of how they are communicating their proposed structures to a range of people across the institution. It is especially important that senior managers are easily engaged in the proposed structure. Some examples of how that might be archived are below:

- The University of Southampton provides a mindmap for the structure of the data on their website: <http://mind42.com/pub/mindmap?mid=605c3bad-3980-4d4b-9155-75b33af8860d>
- A step-by-step walk-through for how data was worked with and identifiers created for various website data: <http://www.jenitennison.com/blog/node/145>
- A personal account of coming to terms with structured data at the start of the work is described here: <http://blogs.ukoln.ac.uk/locah/2010/09/22/creating-linked-data-more-reflections-from-the-coal-face/>

In building their prototype structured public-facing web pages projects will need to consider both the URI structures and how that data at those URIs are made available; examples of both are provided below:

- Public URI sets: [www.cabinetoffice.gov.uk/media/301253/public\\_sector\\_uri.pdf](http://www.cabinetoffice.gov.uk/media/301253/public_sector_uri.pdf)
- Cool URIs <http://www.w3.org/TR/cooluris/>
- How the BBC makes websites: [http://www.bbc.co.uk/blogs/radiolabs/2009/01/how\\_we\\_make\\_websites.shtml](http://www.bbc.co.uk/blogs/radiolabs/2009/01/how_we_make_websites.shtml)
- An example for URI structures is available from The Web Science Trust, for example how they have provided the information about each of their people: <http://webscience.org/people.html>
- University of Southampton publishes its researcher's profiles as machine readable data: <http://www.ecs.soton.ac.uk/people/lac> => <http://graphite.ecs.soton.ac.uk/browser/?uri=http://id.ecs.soton.ac.uk/person/60>
- Projects should consider the value of open data structure standards, e.g. CSV, RSS/ATOM, JSON, RDFa, RDF, etc. Advice is available via UKOLN's developer contact community forum: <http://devcsi.ukoln.ac.uk/dev-contact/>

While the re-structuring of the identifiers and data at the [.ac.uk](http://www.ac.uk) public pages should directly benefit the institution it should also benefit the wider sector, and the following references suggest how that

- Webscale identifiers and how they can add value to an organisations website: [http://www.bbc.co.uk/blogs/radiolabs/2008/06/the\\_simple\\_joy\\_of\\_webscale\\_id.shtml](http://www.bbc.co.uk/blogs/radiolabs/2008/06/the_simple_joy_of_webscale_id.shtml)
- Use of common human readable key-value pairs as identifiers: <http://patterns.dataincubator.org/book/shared-keys.html>
- Curation of websites and the long term value in structuring them that way: <http://derivadow.com/2010/03/11/some-thoughts-on-moving-beyond-the-resource/>

# Geospatial

## Context

JISC has invested in geospatial data and supporting infrastructure services over a number of years. For example back in 1996 JISC invested in trialing infrastructure to serve Ordnance Survey data to the academic community; this resulted in the Digimap service<sup>15</sup>. Investment in resources and infrastructure has continued, resulting in the creation of multiple geospatial resources, tools and infrastructure<sup>16</sup>. The use of geospatial information has become prevalent in the day-to-day lives of students, researchers and teachers; however there is more to be done to exploit the potential benefits for research and learning. The proposals requested in this programme area should aim to make geospatial tools and data more ubiquitous in all subject areas so that end users can directly benefit from their use and re-use. There are a wide variety of applications of geospatial data, tools and analysis suitable for use in a wide range disciplines such as: ecology, criminology, health sciences, archaeology, marketing, business, economics, architecture and design.

It is the intention that the projects commissioned under this funding opportunity will also help to inform requirements for an Academic Spatial Data Infrastructure. This will not be a primary focus of the projects but as set out in the terms of reference it is a relevant concern and collectively the lessons from the projects should help to inform requirements.

The JISC's on-going interest in this area is informed by the work of its Geospatial Working Group (GWG)<sup>17</sup>. The GWG has recently been reconstituted to advise on all aspects of JISC geospatial activity, including the licensing of Ordnance Survey data, the interoperability of services and the take-up of resources to benefit learning, teaching and research. The GWG is currently revising its vision statement in view of providing a fit for purpose UK Academic Spatial Data Infrastructure. The GWG performs a horizon scanning role to indicate where JISC can and should provide leadership and innovation in the use of geospatial technologies.

This briefing paper provides relevant examples of implementations of geospatial solutions and activity that have previously explored the problem space, and all the links are available under the "jiscgeo" tag on delicious<sup>18</sup>. It sets out some of the related tools, services, policies and issues that bidders should take into account when preparing proposals. The list of examples and information given is not exhaustive and its intention is not to limit projects but instead it provides bidders with an overview of the area. Therefore bidders are encouraged to go beyond the suggestions presented here, so long as the proposal clearly meets the 'aim of the programme-area' as set out in the call for proposals. This briefing paper should be read in conjunction with the grant funding call; it provides supplementary contextual information to the terms of reference.

## Supporting background information for bidders

A number of reports and resources that set out directions in the geospatial arena are listed below:

---

<sup>15</sup> <http://en.wikipedia.org/wiki/Digimap>

<sup>16</sup> <http://www.dlib.org/dlib/may04/reid/05reid.html>

<sup>17</sup> <http://www.jisc.ac.uk/aboutus/committees/workinggroups/geospatial>

<sup>18</sup> <http://www.delicious.com/tag/jiscgeo>

- This public broadcast video by Penn State University provides a video overview and history of geospatial technology: <http://geospatialrevolution.psu.edu/>
- European legislation is pushing forward the requirements from the INSPIRE Directive to openly publish geospatial data from public sector organisations: <http://inspire.jrc.ec.europa.eu>
- DEFRA in the UK is taking forward the INSPIRE Directive as applied to Public Sector organisations including H/FEIs: <http://location.defra.gov.uk>
- The Foresight study looks at the future of geospatial data until 2015: <http://www.agi.org.uk/foresight/>
- ESRC has published a report including recommendation for the use of geospatial data: [http://www.esrc.ac.uk/ESRCInfoCentre/Images/Geospatial%20report%20with%20cover%20Dec09\\_tcm6-35008.pdf](http://www.esrc.ac.uk/ESRCInfoCentre/Images/Geospatial%20report%20with%20cover%20Dec09_tcm6-35008.pdf)
- A JISC Tech Watch report has been published that presents a horizon scan on data mash-ups and the future of mapping: <http://www.jisc.ac.uk/whatwedo/services/techwatch/reports/horizonscanning/hs1001>
- Some of the issues to be considered in relation to an academic spatial data infrastructure are set out in the report of a related workshop that was held in 2008: <http://shapingthefuture.pbworks.com/SDI-Tech-Architecture-workshop>

The call seeks projects that define end user requirements, some examples of where this has been done in previous projects follow. Bidders should note that these are only examples of a similar project method and they do not necessarily demonstrate a specific way forward, rather they are highlighted here just as useful examples of starting with an end-use problem to define development.

- Please note these projects were small lightweight projects producing lightweight prototypes and are only an example of stating end user problems: <http://code.google.com/p/jiscrid/downloads/list?q=label:jiscGEO>
- The JISC website has had several projects investigating usability-centric design: <http://www.jisc.ac.uk/whatwedo/topics/personalisation>

Projects should actively look to re-use existing tools, services and datasets. The following examples are potential re-usable components:

- The ShareGeo enables sharing of geo data, a briefing overview is available here: <http://www.sharegeo.ac.uk/docs/about.pdf>
- Deposit of data and resources is encouraged in Open Geo , for those derived from the Digimap Share Geo may be the appropriate place for deposit due to licensing restrictions : <http://sharegeo.ac.uk/>
- Geoparser and open data gazetteer: <http://unlock.edina.ac.uk/>
- Landmap <http://landmap.mimas.ac.uk/>
- Go Geo <http://www.gogeo.ac.uk/cgi-bin/index.cgi>
- GeoDoc <http://geodoc.stottlerhenke.com/>
- Twitter's location APIs provides several features for integrating data with geospatial references: <http://support.twitter.com/entries/194473-twitter-places-and-how-to-use-them>

As referred to above deposit of data is encouraged in Open Geo or Share Geo; however projects may have other suitable repositories that adequately support curation and re-use.

All projects should consider the reuse of technology components to assure machine interoperability; this can often come in the form of working with the communities who support

re-usable components. Some example of geospatial technologies that are widely used in the sector include:

- The Open Street Map community: <http://oliverobrien.co.uk/2010/08/openstreetmap-101/>
- The Open Geospatial Consortium supports the creation of new geospatial standards: <http://www.opengeospatial.org/>
- The Location strategy for the UK has published a recommendation for designing URIs for geospatial data: [http://location.defra.gov.uk/wp-content/uploads/2010/04/Designing\\_URI\\_Sets\\_for\\_Location-Ver0.5.pdf](http://location.defra.gov.uk/wp-content/uploads/2010/04/Designing_URI_Sets_for_Location-Ver0.5.pdf)
- Various Free and Open Source Geospatial tools are listed here: [http://wiki.openstreetmap.org/wiki/Category:Free\\_software](http://wiki.openstreetmap.org/wiki/Category:Free_software)
- The Ushahidi open source platform: <http://www.ushahidi.com/>
- The Open Source Geospatial Foundation supports software creation and community development: <http://www.osgeo.org/>
- The Open Layers project supports the creation of layered maps: <http://openlayers.org/>
- This Digital Preservation Briefing paper provides an overview of common file formats used in the geospatial sector: [http://www.dpconline.org/component/docman/doc\\_download/363-preserving-geospatial-data-by-guy-mcgarva-steve-morris-and-gred-greg-janee](http://www.dpconline.org/component/docman/doc_download/363-preserving-geospatial-data-by-guy-mcgarva-steve-morris-and-gred-greg-janee)
- Projects should consider the merits and interoperability of widely used geospatial file formats such as: GPX, GeoRSS, GML, DXF, GeoJSON, GML, KML, etc.

Several examples of how research and/or teaching and learning data can be applied to geospatial data to encourage multidisciplinary work can be found at the following:

- Scientific work as applied to geospatial: <http://www.bgs.ac.uk/opengeoscience/>
- Government electoral data as applied to geospatial <http://blog.casa.ucl.ac.uk/?p=88>
- Weather data as applied to geospatial: <http://blog.casa.ucl.ac.uk/?p=80>
- Population data as applied to geospatial: <http://blog.casa.ucl.ac.uk/?p=69>
- City sustainability analysis as inferred from geospatial mashups: <http://blog.casa.ucl.ac.uk/?p=55>
- Timelapse data as applied to geospatial (London cycle hire): <http://oliverobrien.co.uk/2010/09/animation-of-cycle-hire-patterns/>
- Twitter and geospatial mashup prototype: <http://www.digitalurban.org/2010/09/san-francisco-tweetography-twitter.html>
- Social behaviour and geospatial data mash-up: <http://infteam.iiscinvolve.org/wp/2010/09/22/the-city-that-never-sleeps-neiss-projects-analyses-geolocated-twitter-data/>
- Population data in space/time map: <http://oliverobrien.co.uk/2010/09/animation-of-cycle-hire-patterns/>
- Social data in geospatial maps (ambulance data): <http://oliverobrien.co.uk/2010/09/london-data-in-maptube/>
- Historical images as applied to geospatial data: <http://www.historypin.com/>
- Sound data as applied to geospatial data: <http://sounds.bl.uk/Maps.aspx>
- Use of geospatial in teaching: [http://www.le.ac.uk/gg/staff/academic\\_jarvis.html](http://www.le.ac.uk/gg/staff/academic_jarvis.html)
- Geospatial data as applied to facts over time: <http://eps.berkeley.edu/~saekow/chronozoom/>
- Visualisation technique for various data over space/time: <http://gisagents.blogspot.com/2010/09/space-time-dynamics-in-scaling-systems.html>
- Geospatial data as it applies beyond our planet: <http://www.worldwidetelescope.org/>

Examples of technical provision for geospatial use are briefly demonstrated at the following examples:

- The Google Maps API has been one of the most popular APIs for developers to use across “Web 2.0” applications: <http://twitter.com/googlemapsapi>
- The Google Maps Team continues to push forward the functionality for working the geo data: <http://googlegeodevelopers.blogspot.com/2010/07/checking-in-with-places-api.html>
- Government organisations are considering the use of linked data to share data across disparate organisations and with other countries: <http://geoservice.psi.enaktinq.org/>
- UCL’s CASA department has been experimenting with the use of augmented reality tools: <http://blogs.splintdev.geog.ucl.ac.uk/tag/augmented-reality/>

All projects will need to consider the relevant policy and trends that are part of any change solution. The following are examples of policies and trends to be taken into account:

- Ordnance Survey Open Creative Commons policy: <http://www.ordnancesurvey.co.uk/oswebsite/opendata/licence/docs/licence.pdf>
- There is growing interest in linked data technologies in relation to geospatial data: <http://location.defra.gov.uk/wp-content/uploads/2009/12/UKLP-Linked-Data-Guide-Final.pdf>
- There are community concerns for the “openness” of ubiquitous geo services such as Google, Yahoo and Bing maps: <http://brainoff.com/weblog/2010/03/16/1541>
- Concerns for the use of “tagging technologies”: <http://fields.eca.ac.uk/totem/wp-content/uploads/2010/08/TaggingTechnologies.pdf>

The call for proposals states that projects will work with associated JISC programme mechanisms, such as the community synthesis project that will be established, along with other JISC programme activities. Projects are also encouraged to participate actively in relevant communities, as this is a way to develop and disseminate ideas and project experiences. Some examples of relevant communities are listed below:

- UKOLN hosts the “Developer Community Supporting Innovation (devCSI) project which supports and trains developers across the UK HE & FE sector: <http://devcsi.ukoln.ac.uk/>
- CETIS holds several community events across various learning and teaching technologies: <http://jisc.cetis.ac.uk/events/>
- The annual Open Street Map Conference is a common gather point for the geospatial community <http://stateofthemap.org/>
- There are local chapter groups who actively participate in geospatial activity: [http://wiki.openstreetmap.org/wiki/Foundation/Local\\_Chapters](http://wiki.openstreetmap.org/wiki/Foundation/Local_Chapters)
- Another large and distributed group is the geo-caching community: <http://www.geocaching.com/>

## Infrastructure for resource discovery

In 2009 JISC and RLUK convened a group of Higher Education library, museum and archive experts to think about what national services were required for supporting online discovery and reuse of collection metadata. This group was called the resource discovery taskforce (RDTF) and it met four times throughout 2009. The taskforce produced a vision and an implementation plan focused on making metadata about collections openly available therefore supporting the development of flexible and innovative services for end users. You can read summaries of the RDTF meetings and the vision and implementation plan on the RDTF blog <http://rdtf.jiscinvolve.org>. This blog will also be used to post news about progress.

The RDTF vision sets targets to meet by the end of 2012. JISC will be funding a range of projects, communication and support activities designed to meet those targets. This work will fall into four rough categories: institutional level, aggregation level, service level and support. JISC has begun to fund a range of work in these categories:

### Institutional:

- A guide to Open bibliographic data for librarians: <http://rdtf.jiscinvolve.org/wp/2010/08/05/update-on-current-work/> This is due to be delivered in November and will be useful to successful bidders.
- The projects funded under this call will begin to address the issues of what needs to be done at an institutional level to realise the RDTF vision.
- There are 3 projects funded under the JISCexpo call <http://code.google.com/p/jiscexpo/> that will produce metadata and knowledge that will be relevant to the RDTF vision:
  - Linked Open Copac Archives Hub: <http://blogs.ukoln.ac.uk/locah/>
  - Linking University Content for Education and Research Online: <http://lucero-project.info/lb/2010/07/hello-world-2/>
  - Openbib: <http://openbiblio.net/p/jiscopenbib/>

### Aggregation

- A study investigating the issues with aggregation of metadata about images and time based media: <http://rdtf.jiscinvolve.org/wp/2010/08/05/update-on-current-work/>
- Projects will be funded at Edina and Mimas to enhance existing aggregations like Suncat <http://www.suncat.ac.uk/> and Copac <http://copac.ac.uk/> in line with the RDTF vision. Projects are expected to start in December.

### Services

- Work on this level will be funded later in the process, but JISC will be working with stakeholders in this area to define requirements, communicate possibilities and identify opportunities.

### Support

- JISC has issued an invitation to tender (ITT) for a management framework for the RDTF. This project will involve setting up a central website for engaging with RDTF work and analysing and providing advice and guidance on key issues such as sustainability, licensing, standards and technology. All projects will be expected to dedicate time to engaging with the management framework. The framework project is expected to start in November.
- JISC will be issuing an open ITT for a support role dedicated to managing the relationships with the many stakeholders listed in the implementation plan and communicating effectively with those stakeholders. This project is expected to start in December.

Institutional projects funded under this call are part of this larger programme and have been purposefully designed to be short to permit future projects to build on the successful approaches developed by the projects and to allow the metadata produced to be reused by subsequent projects.

Libraries, museums and archives throughout Europe are engaging in work on open bibliographic or open metadata projects that are relevant to JISC work on the RDTF vision. Projects and developments are too numerous to mention in this briefing paper but here are some particularly useful links:

- Announcement of OKFN working group on open bibliographic data: <http://blog.okfn.org/2010/03/03/new-working-group-on-open-bibliographic-data/>
- The OKFN mailing list is a useful tool for keeping up to date on open bibliographic data matters: <http://lists.okfn.org/mailman/listinfo/open-bibliography>
- Karen Coyle's blog post is a useful summary of developments with open bibliographic data in Europe: <http://www.opencontentalliance.org/2010/03/17/going-open/>
- During the RDTF, Rightscom produced a useful report which investigated relevant activity happening elsewhere in the world. The report is available from the RDTF blog <http://rdtf.jiscinvolve.org/wp/2009/09/11/information-gathering-report/>
- The British Library has made some of their bibliographic datasets available for research purposes: <http://www.bl.uk/bibliographic/datafree.html>
- The Technology Strategy Board has released a competition focused on metadata: <http://www.innovateuk.org/content/competition/metadata-increasing-the-value-of-digital-content-m.ashx> and <http://www.innovateuk.org/content/competition/metadata-increasing-the-value-of-digital-content-f.ashx>
- The Collections trust have produced the Culture Grid, a service that aggregates content from public museums, libraries and archives: <http://www.culturegrid.org.uk/>

The methodology described in this call was shaped by a meeting that UKOLN organised to discuss technical approaches for the RDTF vision. Bidders are encouraged to read the summaries of the meeting:

- <http://www.ukoln.ac.uk/jisc-ie/blog/2010/08/19/aggregation-and-the-resource-discovery-taskforce-vision/>
- <http://efoundations.typepad.com/efoundations/2010/08/resource-discovery-revisited.html>

The management framework will build upon the advice from this meeting and will establish a technical group to provide advice and guidance throughout the RDTF vision implementation.

This call encourages projects that engage with Linked Data. The following links may be helpful to bidders wishing to explore Linked Data for their project:

- Linkeddata.org: <http://linkeddata.org/>
- When thinking about how to publish the uri sets developed during projects, bidders should take account of the Cabinet Office guidelines: [http://www.cabinetoffice.gov.uk/media/301253/public\\_sector\\_uri.pdf](http://www.cabinetoffice.gov.uk/media/301253/public_sector_uri.pdf)
- The W3C has set up a group on Linked Data and libraries and there is a useful wiki that accompanies this group: [http://www.w3.org/2005/Incubator/ld/wiki/Main\\_Page](http://www.w3.org/2005/Incubator/ld/wiki/Main_Page)
- Bidders are encouraged to look at the relevant JISC Linked Data projects listed above when preparing their submission and if possible, talk to those involved.

These projects are being funded to allow exploration of making library archive and museum metadata openly available on the web. This exploration will support the discovery of approaches that will work for other Higher Education institutions. It will also permit the exploration of the major issues involved in making metadata available under an open licence, issues such as licensing, schema, provenance, authority and technical issues. Projects will be required to discuss these issues on their project blogs and JISC and relevant

RDTF projects will collect these lessons and process them into appropriate formats so that the sector as a whole benefits from the projects. Successful bidders will need to pay close attention to the JISCexpo strand of projects mentioned above (<http://code.google.com/p/jiscexpo/>) as these projects will be investigating a similar set of issues.

For these projects JISC's stance on licensing is that the ODC-PDDL licence is the preferred option for releasing the data. You can read more about the ODC-PDDL licence <http://www.opendatacommons.org/licenses/pddl/1-0/>. This is a recommendation, not a requirement. However, projects choosing a different approach must discuss the reasons for this on their project blog so that JISC can collect and distribute the lessons. Projects will be supported in addressing licensing issues by the RDTF management framework project. This will be the topic of a meeting in late January or early February 2011 for successful projects.

As stated above the purpose of these projects is to identify approaches that work. Therefore JISC does not require bidders to have answers for every technical issue at the bid stage. There will be room for projects to develop their approach as they work and JISC will be supporting projects on technical issues through programme meetings and contact with the RDTF management framework project. This will also be a topic at the meeting in late January or early February 2011 for successful projects.

## Activity data

### Context

Companies such as Amazon have proved that careful use of data about users' actions and attention can improve services to end users and provide advantages in managing and targeting the service. Many systems in institutions store data about the actions of students, teachers and researchers. The purpose of this funding is to experiment with this data with the aim of improving the user experience or the administration of services.

Projects funded under this strand will identify tools and techniques that prove beneficial to university services. Once identified, JISC will fund further development to enable as many Higher Education institutions as possible to benefit from them.

### Glossary

The call for projects uses a number of specific terms that are worth describing in more detail.

We are using three related but distinct definitions for different types of activity data:

- Business intelligence - the use of data to provide historical, current and predictive views of business operation and the use of those views to support better decision making.
- User activity data - a record of a user's actions on a website or software system or other relevant institutional service.
- Attention data - the record of what a user has viewed on a website or software system or other relevant institutional service.

This call is focused on experimentation with user activity data and attention data. We are using the umbrella term activity data to refer to both of these data types. Projects funded under this call should seek to use activity data to improve institutional services or

There is a separate JISC call related to business intelligence data:

[http://www.jisc.ac.uk/fundingopportunities/funding\\_calls/2010/09/grant1210.aspx](http://www.jisc.ac.uk/fundingopportunities/funding_calls/2010/09/grant1210.aspx). It is seeking projects that use internal and external sources of data to address strategic institutional concerns and help improve decision making. Implementation of specific business intelligence software systems is also a major part of the call. We advise all bidders to read both calls carefully to assess which fits their project most closely. JISC will reject bids if the same or very similar bid is submitted to both calls.

The following criteria should help bidders decide which call to submit to:

The activity data call is for projects:

- Improving the management of an institutional service
- Improving the end user experience
- Targeting an institutional service more effectively
- Working with the data stored in existing institutional systems

The business data call is for projects:

- Addressing key institutional strategic concerns
- Seeking to improve decision making
- Implementing a system to manage business intelligence data
- Seeking to use internal and external sources of data

We are asking all bidders to state an institutional problem or hypothesis in their bids and make this central to the project. All problems or hypotheses should be practical and testable.

In July 2011 we will be looking for successful projects to build on, and the problem or hypothesis and the data to support or deny it will be a key part of the evaluation of successful projects. Therefore, bidders should carefully plan the way in which the hypothesis or problem is evaluated and include these plans in their bids.

As stated in the call, bids should have a strong focus on the improvement of an institutional service. Service is a word that can have subtle variations of meaning. In this case the word is used to cover both of the following variations:

- A piece of software or website used to provide a service to end users
- A service department such as the library or research management

We are looking for projects that can provide improvements in both of these types of services. Possible improvements can be from a user perspective or from an administrative perspective. This service improvement can form part of the hypothesis and, like the hypothesis, this improvement must be practical and measurable and projects will be judged on it in July 2011. Evaluators will be looking for a well scoped service improvement with realistic plans for measurement of the improvement.

Possible improvements could include:

- Interface development
- New functionality for end users
- Collection or stock management
- Study of behaviour to reveal new information
- Prediction of behaviour or demand and altering of service to meet it

Funding is available for projects focusing on institutional improvements and also for projects seeking to provide a national, regional or consortial shared service. Shared services offer opportunities to provide efficiency savings or greater effectiveness for the institutions involved or should provide a better service to end users than an institutional service could offer. Projects will need to address these issues during the lifecycle of the project and should collect and disseminate knowledge gained via the project blog. Projects will also be encouraged to focus on the business cases for shared services that would persuade institutions to engage with the proposed service.

This section of the call is broad in scope and is seeking bids that involve any institutional service or software system. Services or systems that could be interesting to investigate include:

- Library management systems
- Electronic resources
- Websites
- Research management systems
- Reference management Systems
- Virtual learning environments
- Student record systems
- Repositories and other types of digital resource management systems
- Identity management systems

This list is not intended to be exhaustive and bidders are free to include systems and services that are not included on the list.

## **Related work**

There is a range of related work that bidders may be interested in:

- The JISC Mosaic project investigated the possibilities of collecting and reusing library circulation data. The project final report is relevant to all bidders seeking to work with

library data <http://ie-repository.jisc.ac.uk/466/>. The project also organised a developer competition that may be a source of potential ideas <http://devcsi.ukoln.ac.uk/demonstrator/2009/10/22/jisc-mosaic-project-competition-winners/>.

- Infonet have produced an infokit on business intelligence that may have relevant information for bidders to this call as well as the business intelligence call <http://www.jiscinfonet.ac.uk/bi>.
- Mark Stubbs and team at Manchester have been experimenting with the data stored by Virtual Learning Environments to explore the issue of student progression<sup>19</sup>.
- The Knowledge Exchange has a group that is thinking about usage statistics in the retrieval of scholarly information. This group has produced a number of useful outputs including a briefing paper. All outputs can be found on their website: <http://www.knowledge-exchange.info/Default.aspx?ID=365>
- JISC is funding a project that is developing a usage statistics portal for libraries to manage statistics about electronic journal usage. More information can be found on the project website: <http://www.iusp.mimas.ac.uk/>
- JISC is funding a project called PIRUS which is investigation the extension of Counter statistics to cover article level usage of electronic journals: <http://www.jisc.ac.uk/whatwedo/programmes/pals3/pirus.aspx>
- The JISC RAPTOR project is investigating ways to explore usage of e-resources: <http://www.jisc.ac.uk/whatwedo/programmes/aim/raptor.aspx>. This project is part of the AIM programme and bidders interested in activity data from identity management systems are advised to browse this programme: <http://www.jisc.ac.uk/whatwedo/programmes/aim.aspx>
- The Strategic Content Alliance has produced an audience analysis toolkit that bidders may find useful: <http://sca.jiscinvolve.org/wp/audience-publications/>
- The Technology Strategy Board are releasing a competition exploring the possibilities of extracting value from large and diverse sources of data <http://www.innovateuk.org/content/competition/harnessing-large-and-diverse-sources-of-data.ashx>. There may be overlaps from the TSB competition with the work funded by this JISC call.
- Edina released a study that, among other things, considered the possibility of using data from openurl link servers: [http://edina.ac.uk/projects/Shared\\_OpenURL\\_Data\\_Infrastructure\\_Investigation\\_summary.html](http://edina.ac.uk/projects/Shared_OpenURL_Data_Infrastructure_Investigation_summary.html). JISC and Edina are investigating the possibility of more work in this area and it is possible that this work will run concurrently with projects funded under this call.
- JISC held an event that explored the issue of activity data in July of this year: <http://www.jisc.ac.uk/events/2010/07/businessintelligence.aspx>. The report from this event is available and bidders are strongly recommended to read it: <http://ie-repository.jisc.ac.uk/486/>. Bidders may also be interested in the reflections on the event from Balviar Notay: <http://infteam.jiscinvolve.org/wp/2010/09/29/event-report-gaining-business-intelligence-from-user-activity-july-2010/>

This list highlights some relevant work. It is not intended to be an exhaustive account of all work in this area.

---

<sup>19</sup> Hardman, Julie; Paucar-Caceres, Alberto; Urquhart, Cathy; and Fielding, Alan, "Predicting Students Progression Using Existing University Datasets: A Random Forest Application" (2010). AMCIS 2010 Proceedings. Paper 272. <http://aisel.aisnet.org/amcis2010/272>

# Repositories: take-up and embedding

## Context

As indicated in the Call, JISC has been working towards a vision of a rich scholarly communications and learning environment that is supported by an extensive network of repositories, where content is widely available and can be re-used. JISC believes that there is now a solid foundation for a UK repository infrastructure. This infrastructure has been developed from JISC programmes/projects and independent work undertaken by institutions and other organisations in the UK and internationally.

Note that the following examples in this briefing paper are just a selection of a wide range of tools applications and services developed by JISC funding and others.

This infrastructure consists of:

1. The building and enhancement of institutional repositories, subject repositories – there are now 178 UK repositories (according to OpenDoar <http://www.opendoar.org/>) and UK PubMed Central <http://ukpmc.ac.uk/> is an example of an open access subject based repository.
2. The development of shared services and projects such as
  - SHERPA RoMEO <http://www.sherpa.ac.uk/romeo/>
  - OpenDOAR <http://www.opendoar.org/>
  - ROAR <http://roar.eprints.org/>
  - Repository Junction <http://edina.ac.uk/projects/oa-rj/index.html>
  - The Depot [http://edina.ac.uk/projects/depot\\_summary.html](http://edina.ac.uk/projects/depot_summary.html)
  - Jorum <http://www.jorum.ac.uk/>
  - The Repositories UK aggregation <http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/repubk.aspx>
  - Institutional Repository Search <http://irs.mimas.ac.uk/demonstrator/>.
3. The development of software platforms – examples of these are ePrints <http://www.eprints.org/>, DSpace <http://www.dspace.org/>, Fedora <http://fedora-commons.org/>. Please note that both DSpace and Fedora are now part of the same organisation DuraSpace <http://duraspace.org/index.php>.
4. Software tools and applications such as MePrints <http://allaboutme.eprints.org/> and SNEEP <http://sneep.ulcc.ac.uk/>.
5. The development of preservation tools such as the Drambora Toolkit <http://www.repositoryaudit.eu/>.
6. Providing training and skills through JISC projects and in particular the Repositories Support Project, RSP <http://www.rsp.ac.uk/>.
7. The development of national support and guidance - examples of these are RSP, the United Kingdom Council of Research Repositories, UKCoRR <http://www.ukcorr.org/>, the Digital Curation Centre, DCC <http://www.dcc.ac.uk/>, UKOLN <http://www.ukoln.ac.uk/> and CETIS <http://jisc.cetis.ac.uk/>.
8. International collaboration and activities - examples of this include
  - the Confederation of Open Access Repositories <http://coar-repositories.org/>
  - DRIVER <http://www.driver-repository.eu/>
  - OpenAIRE <http://www.openaire.eu/>

- PEER <http://www.peerproject.eu/>
- SONEX <http://sonexworkgroup.blogspot.com/>
- collaborative work via relevant “Knowledge Exchange” Groups <http://www.knowledge-exchange.info/>.

The following is the list of JISC programmes that have contributed to building of this infrastructure and proposals should build on these programmes and their outputs.

1. Digital Repositories programme 2005-7  
<http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2005.aspx>. This programme brought together people and practices from across various domains (research, learning, information services, institutional policy, management and administration, records management, and so on) to ensure the maximum degree of coordination in the development of digital repositories, in terms of their technical and social (including business) aspects.
2. Repositories and Preservation Programme  
<http://www.jisc.ac.uk/whatwedo/programmes/reppres.aspx>. This was an investment in Higher Education repository and digital content infrastructure. It funded initiatives to develop the Information Environment supporting digital repositories and preservation, including cross-searching facilities across repositories; funding for institutions to develop a critical mass of content, preservation solutions and advice for the development of repositories.
3. Users & Innovation: Personalising Technologies  
<http://www.jisc.ac.uk/whatwedo/programmes/usersandinnovation.aspx>. Based on the needs of real end-users, the programme scoped and developed areas where innovative user technologies and practices could increase efficiency.
4. Digital preservation & records management programme  
<http://www.jisc.ac.uk/whatwedo/programmes/preservation/assetmanagement>. This programme was designed to support institutions in long-term digital asset management and preservation.
5. Information Environment Programme 2009-11  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11.aspx> Specifically the following strands of work:
  - a. Repositories enhancement  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/sue2.aspx>
  - b. Preservation <http://www.jisc.ac.uk/whatwedo/programmes/inf11/digpres.aspx>
  - c. Deposit <http://www.jisc.ac.uk/whatwedo/programmes/inf11/jiscdepo.aspx>
  - d. Rapid Innovation  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/jiscris.aspx> and the ‘Toolshed’ map summary at [http://ie-repository.jisc.ac.uk/449/2/jiscris\\_map1.html](http://ie-repository.jisc.ac.uk/449/2/jiscris_map1.html)
  - e. Resource Discovery, especially projects working in the area of automatic metadata generation and text mining  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/resdis.aspx>

## Specific Examples:

### Administrative, ingest and service workflow processes

1. The University of Glasgow is a good example of an embedded repository within an institution as they have linked the repository with the research system, and it feeds information for the staff pages. They have also done some back-end work to take full advantage of tools and good practice e.g. author disambiguation in order to fully realise the benefits of integration. The disambiguation work at Glasgow has enabled them to build new services such as using the publications in Professorial Professional Development. An embedded repository is not just one that is technically linked to other institutional systems but that is demonstrably/clearly embedded within the policy and processes of the wider institution. This example could prove to be an interesting model to follow.  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/sue2/enrich>
2. The Universities of St Andrews and Aberdeen have jointly procured an Atira system (the PURE research information system <http://atira.dk/en/pure/>) and have done a lot of work in embedding it with their repositories.  
<http://www.abdn.ac.uk/documents/Pure-launch-email.pdf>
3. The CRISPool <http://www.jisc.ac.uk/whatwedo/projects/crispool.aspx> project used CERIF-XML to integrate heterogeneous research information from several institutions into a single portal. This project forms part of the Research Information Management strand of projects.  
<http://www.jisc.ac.uk/whatwedo/themes/informationenvironment/researchinfomgt.aspx>

Please note, that with regards to integrating research management, administrative and repository systems, bidders should read both the “Repositories: take-up and embedding” and “The Research Information Management” areas of the call document and briefing paper, and then decide which is more appropriate for their proposal to bid against.

### Technical processes

1. SWORD is a lightweight protocol for depositing content from one location to another. It stands for Simple Web-service Offering Repository Deposit and is a profile of the Atom Publishing Protocol (known as APP or ATOMPUB). As mentioned above, a number of repositories and services are using the SWORD API as part of their workflow, nationally and internationally.  
<http://swordapp.org/>
2. The MERLIN project has used the TerMine tool to extract keywords from the full text digital objects harvested by the University of London's repository aggregator, LASSO. These keywords are presented as part of an interactive 'cloud' display to create an innovative, subject-specific search interface.  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/resdis/merlin.aspx>

<http://www.nactem.ac.uk/software/termine/>

3. Repository Junction. The aim of this project is to assist open access deposit into, and interoperability between, existing repository services, by developing a deposit broker system. <http://edina.ac.uk/projects/oa-rj/index.html>

### Links with other systems and applications

1. The Readiness for REF project (R4R) at King's College London is developing an interoperable system to support REF, based on repository-CRIS integration using CERIF, a model that will be shared with REF stakeholders - a group of 100 HEIs. <http://www.jisc.ac.uk/whatwedo/programmes/inf11/sue2/r4r>  
See also note above with respect "Research Information Management".
2. Trinity College, Dublin. This is an example of an institution that has developed excellent interoperability between its institutional repository (TARA) and its current research information system. <http://www.tcd.ie/Library/riss/tara/>
3. Open University repository<sup>20</sup> can be seen as another exemplar of good practice for embedding a repository within an institution. This example could also prove to be an interesting model to follow.

### Improvement to interfaces

1. KULTUR was a JISC-funded project which created a model of an institutional repository for use in the creative and applied arts. The project investigated a policy and technical framework for creating a multimedia, multifunctional repository, applicable both to specialist institutions and departments across the sector. <http://kultur.eprints.org/index.htm>
2. The FAROES project developed a lightweight repository that modern language practitioners can incorporate into their everyday work, with a user interface that follows the best-practice principles of Web 2.0 sites (social, interlinked, evolving and flexible). <http://www.faroes.ecs.soton.ac.uk/>
  - a. The winning entry for the Developer Challenge at the Open Repositories Conference 2010 Madrid was from the University of London's Computer Centre. This work is developing a proof-of-concept demonstrator of a simple framework for activating links from embedded semantic metadata, using an example based on the Linnean Online Collections repository. <http://vimeo.com/13172548>

---

<sup>20</sup> Smith, Colin; Yates, Christopher and Chudasama, Sheila (2010). Open Research Online - a self-archiving success story. In: The 5th International Conference on Open Repositories, 6-9 July 2010, Madrid, Spain. [http://oro.open.ac.uk/22321/3/OR2010\\_handout.pdf](http://oro.open.ac.uk/22321/3/OR2010_handout.pdf)

Please note, that although the examples above refer to ePrints developments we will be seeking projects that work with any of the major repository platforms to improve user interfaces based on good practice from other institutions.

As well as interface improvements, usability improvements are in scope. However bidders must seek to implement proven good practice from other institutions and not undertake novel or innovative work.

## Digital preservation

1. KeepIt project: Preservation into ePrints platform adapted and applied proven documented approaches and training to develop preservation plans, and developed the KeepIt training resources.  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/digpres/keepit.aspx>  
[http://wiki.eprints.org/w/KeepIt\\_Training\\_Resources](http://wiki.eprints.org/w/KeepIt_Training_Resources).
- b. The Embedding Institutional Data Curation Services in Research (EIDCSR) project is addressing the research data management and curation challenges of three research groups in the University of Oxford.  
<http://eidcsr.oucs.ox.ac.uk/>
- c. PLANETS project <http://www.planets-project.eu/> produced a suite of preservation tools, including a test bed, an interoperability framework and the PLATO planning tool. The tools are being maintained by the Open Planets Foundation <http://www.openplanetsfoundation.org/>.
- d. Biophysical Repositories in the Lab project (BRIL) developed at the Centre for eResearch (CeRch) at Kings College, London is working with researchers to enhance repository workflows.  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/digpres/bril.aspx>
- e. The Preservation Exemplar at Kings (PEKin) project will investigate scope and implement a solution for the management of electronic business records and research materials in an HE institution.  
<http://www.jisc.ac.uk/whatwedo/programmes/inf11/digpres/pekin>
- f. Example Preservation Tools include:
  - JHOVE provides functions to perform format-specific identification, validation, and characterization of digital objects.  
<http://hul.harvard.edu/jhove/index.html>
  - PRONOM is a resource for anyone requiring impartial and definitive information about the file formats, software products and other technical components required to support long-term access to electronic records and other digital objects of cultural, historical or business value.  
<http://www.nationalarchives.gov.uk/pronom/>
  - DROID (Digital Record Object Identification) is an automatic file format identification tool

<http://sourceforge.net/projects/droid/>

- DRAMBORA toolkit is intended to facilitate internal audit by providing repository administrators with a means to assess their capabilities, identify their weaknesses, and recognise their strengths.

<http://www.repositoryaudit.eu/>

As mentioned above this is just a selection of good practice and if you want to look for more information you can go to JISC Programme webpages, UKCoRR, RSP or the JISC Programme Manager (Balviar Notay [b.notay@jisc.ac.uk](mailto:b.notay@jisc.ac.uk))

## **Working with organisations**

1. Successful projects will be expected to work closely with the Repositories Support Project, as they will be involved with much of the dissemination of project outputs. This will be vital as the Call requires best practice to be shared to the broader community.
2. The United Kingdom Council of Research Repositories (UKCoRR) is a professional organisation for those working in the field of repositories in the UK. It provides a forum for discussion amongst practitioners. Successful projects will also be expected to participate in dialogue with this organisation. Please note that this list is for repository managers and you will need to register to participate in this forum.
2. UKOLN facilitate a Metadata Forum (meetings, blog etc) for interested practitioners. Successful bids working in the area of metadata would be expected to contribute to this forum.  
<http://blogs.ukoln.ac.uk/themetadataforum/>
4. If you would like to seek advice in advance of submitting a proposal for this Call, you are encouraged to contact the RSP and/or the Metadata Forum. [support@rsp.ac.uk](mailto:support@rsp.ac.uk)  
Tel: 0845 257 6860

## Preservation tools

### Context

The projects in this strand are geared towards institutions, groups or individuals who perceive themselves to have specific problems with the short, medium or long-term preservation of digital materials. The funding is designed to enable bidders to engage with the range of tools and methods that have been developed over the years to assist people with the sort of information management tasks that are relevant to sustaining the integrity of, and access to, digital materials over time.

### Preservation Tools

The following resources contain references to many of the sorts of tools that are in scope for this strand. In aggregate, they should be fairly comprehensive of the variety of current tools available. They should not, however, be regarded as strictly prescriptive of the tools that should be used.

The Digital Curation Centre (DCC) maintains a list containing short descriptions of tools and resources. This is lengthy and it will be apparent that it contains a variety of (sometimes duplicate) resources over a number of different categories. Many of them cannot strictly be classified as preservation *tools* but may well be useful in the broader context of information management tasks that will need to take place in conjunction with specified preservation actions. <http://www.dcc.ac.uk/resources/external/software-and-hardware/tools>

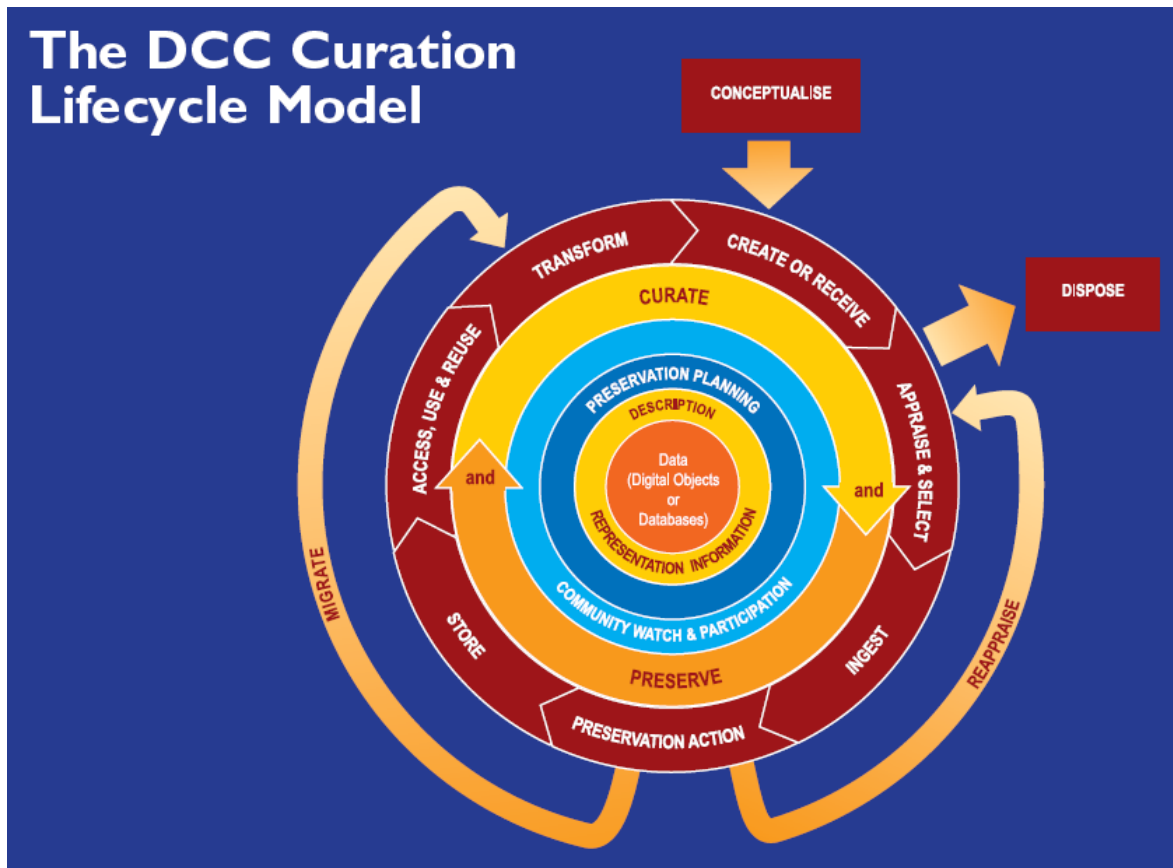
The National Digital Information and Infrastructure Preservation Programme (NDIIPP) based at the US Library of Congress also provides a very useful list of tools. <http://www.digitalpreservation.gov/partners/resources/tools/index.html>

Another useful resource that contains references to tools but which may also help to elucidate the principles and scope of preservation activity is provided by the KeepIt Project, a JISC Preservation Exemplar project funded as part of the INF11 Programme. [http://wiki.eprints.org/w/KeepIt\\_Training\\_Resources](http://wiki.eprints.org/w/KeepIt_Training_Resources)

### Preservation Activity

The Digital Curation Centre has created a summary diagram that helps to define how digital preservation and curation align with a lifecycle management approach to digital information.

# The DCC Curation Lifecycle Model



Please see the DCC website for a more detailed explanation of the model.  
<http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>

## Specific requirements

The purpose of this strand is therefore threefold:

- a) to enable bidders to take practical steps in preserving digital information
- b) to test the fitness for purpose of the available tools and to critique, validate and/or demonstrate their use
- c) to communicate the lessons learnt and evidential value of the project more widely

# Preservation of complex visual digital materials and environments

## Context

The purpose of this work is to identify, clarify and disseminate emerging good practice relating to the preservation of digital materials that have predominantly visual properties, and which may feature: properties of interactivity; complex interdependencies between objects (often of different formats); and time-based components.

## Related Work

The Call text cites some areas of related work that may have some bearing on the nature and scope of the work in this strand.

The sustainability of software has most recently been addressed by the JISC-funded project, Software Preservation Project, managed by Curtis & Cartwright in partnership with the Software Sustainability Institute: <http://softwarepreservation.jiscinvolve.org/wp/> and <http://www.software.ac.uk/what%20do%20we%20do/sustaining-software?>

This builds on previous work undertaken by staff at the Science and Technology Facilities Council (STFC) working on the SoftPres Project <http://www.e-science.stfc.ac.uk/projects/software-preservation/softpres8985.html>

And prior work by the same team on the 'Significant Properties of Software' <http://sigsoft.dcc.rl.ac.uk/twiki/bin/view/Main/SigSoftFramework>

Links to additional work in the area of the significant properties of digital objects is available at: <http://www.jisc.ac.uk/whatwedo/programmes/preservation/2008sigprops>

The emulation of computing environments is a well established technique for preserving complex digital materials and has been the focus of various initiatives over the years. A current major EC-funded project looking at this area of work is the KEEP Project <http://www.keep-project.eu/ezpub2/index.php>

## Strand Scope and Related Work

The three specific types of outputs which projects in this strand should focus on are:

1. Simulations and visualizations
2. Gaming environments and virtual worlds
3. Software based art

The following are illustrative links to definition, description and examples of work that pertain to the above categories:

### Simulations and Visualisations

The National e-Infrastructure for Social Simulation Project (NEISS)

<http://www.geog.leeds.ac.uk/projects/neiss/about.php>

- Visualisation of geo-related data <http://www.maptube.org/>
- VizNet is working towards creating a UK National Visualisation Network <https://wiki.viznet.ac.uk/bin/view/VizNET/WebHome>

- The King's Visualisation Lab specialise in visual representations of archaeology and historic buildings <http://www.kvl.cch.kcl.ac.uk/>

### **Gaming Environments and Virtual Worlds**

- The best known virtual world environment is Second Life <http://secondlife.com/>
- A JISC-funded study on the use of 'Serious Virtual Worlds' and their role in education <http://www.jisc.ac.uk/publications/reports/2008/seriousvirtualworldsreport.aspx>
- A JISC-funded study on 'Learning in Immersive Worlds: a review of game-based learning' [http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport\\_v3.pdf](http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport_v3.pdf)
- The NDIIPP-funded Preserving Virtual Worlds Project led by the University of Illinois <http://pvw.illinois.edu/pvw/>

### **Software-Based Art**

The term 'software-based art' is used in this context to refer to works of art where software is integral to the artwork. The term excludes works of art where the software is simply part of a control device, for example in the display of multiple channels of video or audio, or in the production of a work of art in a non digital form such as a digital print.

Software-based art often involves code which draws upon a range of assets such as images, sound, text or other data. It includes works which are stand alone, networked, interactive or generative. Essentially networked works are linked to external data sources and include works connected to the internet, a local area network or a closed network. 'Interactive' refers to works which require a person to engage with it, and 'generative' is where the data is changed during the function of the work of art.

Examples of software based art include:

- Rafael Lozano Hemmer 'Subtitled Public' <http://www.lozano-hemmer.com/projects.php>
- Yael Kanarek 'World of Awe' <http://www.worldofawe.net/thejournal/landing/>
- Michael Craig Martin 'Things Change'
- (No obvious link readily available but an interesting critique of the function and display of this work is available in a Guardian review <http://www.guardian.co.uk/artanddesign/2007/dec/30/art>)
- And the works in the Decode exhibition at the V&A <http://www.vam.ac.uk/microsites/decode/exhibition/code>

Organisations engaged with this field of preservation/conservation include:

- The Tate Gallery – Matters in Media Art <http://www.tate.org.uk/research/tateresearch/majorprojects/mediamatters/>
- Documentation and Conservation of Media Arts Heritage (DOCAM) <http://www.docam.ca/en/case-studies/rokey-machine-for-taking-time.html>

### **EC-funded Activity relating to the preservation of complex objects**

- The KEEP Project is addressing emulation topics <http://www.keep-project.eu/ezpub2/index.php>
- The PLANETS project has produced various reports, various of which address emulation issues [http://www.planets-project.eu/docs/presentations/Planets\\_iPRES\\_jrvanderhoeven\\_emulation.pdf](http://www.planets-project.eu/docs/presentations/Planets_iPRES_jrvanderhoeven_emulation.pdf)
- PrestoSpace – Preservation towards storage and access: Standardised Practices for audiovisual contents in Europe

- <http://www.prestospace.org/>
- CASPAR – Cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval
- <http://www.casparpreserves.eu/>
- ICT Work Programme 2009-2010: Objective 4.1: Digital Libraries and Digital Preservation (section b1):
- [ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/digicult/outcomes-a-and-b\\_en.pdf](ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/digicult/outcomes-a-and-b_en.pdf)

## Sustaining “at risk” online resources

### Context

This call is an opportunity to sustain resources otherwise at risk, and to tackle their sustainability in such a way that lessons can be shared across the sector. It is anticipated that initial planning and preparation will have been undertaken and bids should show an understanding of the work involved. Projects should make use of available advice on good practice in the management of digital resources.

Bidders will need to identify the new home for the content. Examples of potential homes for sustaining at risk resources might include public and private organisations such as:

- Commercial organisations
- Publicly-funded organisations
- JISC-funded services such as: Mimas <http://mimas.ac.uk/portfolio/current/>, Edina <http://edina.ac.uk/>, Jorum <http://www.jorum.ac.uk> and advice and guidance services at JISC Advance <http://www.jiscadvance.ac.uk/>
- Consortia

Whatever the new home, the content must be made available to FE/HE on an ongoing basis, to ensure continued access, use and maintenance.

As stated in the Call, potential bidders should contact potential new homes to assess feasibility of transfer.

### Relevant links

The following list is not exhaustive, but it indicates sources of advice particularly relevant to this work

- Understanding sustainability and doing it better <http://www.jisc.ac.uk/events/2010/07/jif10/virtualgoodybag/understandingsustainability.aspx>
- What to do when a service provider closes <http://www.ukoln.ac.uk/cultural-heritage/documents/briefing-82/>
- <http://www.dpconline.org/advocacy/knowledge-base/594-digitisation-programme-digital-preservation-study> Follow link to a downloadable PDF p18-19 is a self assessment questionnaire
- Preservation of Web Resources Handbook: <http://jiscpower.jiscinvolve.org/wp/guide/>
- Clearing the rights in order to re-release under a Creative Commons license [http://sca.jiscinvolve.org/wp/files/2009/10/sca\\_ipr\\_toolkit-v2-01\\_intro.pdf](http://sca.jiscinvolve.org/wp/files/2009/10/sca_ipr_toolkit-v2-01_intro.pdf)
- Sustainability and business planning <http://sca.jiscinvolve.org/wp/business-modelling-publications/>