

JOINT INFORMATION SYSTEMS COMMITTEE

JISC Circular 04/06 : Appendix G

Repositories and Preservation Programme

Summary

G1. JISC invites proposals in the following areas:

- I Tools and Innovation (Strand B)
- II Discovery to Delivery (Strand C)
- III Repository Start-Up and Enhancement Projects (Strand D)
- IV Digital Preservation and Records Management (Strand H)
- V Shared Infrastructure Services (Strand I)

G2. For a complete list of all the programme strands, A to J, including those that are funded outside of this particular call, see the associated Repositories and Preservation Briefing Paper.¹

G3. Proposals are invited in the areas set out in the table below. The table also specifies the paragraphs within the call that specify the details of the projects and their scope.

Call Area	Theme/context	Description	Funds	Paragraph
Programme Strand B: Tools and Innovation				
I	Projects that look at developing tools (software or model practices) that can enhance the use of repositories or digital information systems. These tools and technologies can be innovative and new.	Projects may address more than one area. <ul style="list-style-type: none"> • Knowledge organisation systems and text mining • Semi-automatic indexing and classification • Metadata generation and validation • Tools to support added value services including learned societies • Usage metrics for impacts/trends • Assessment of user and management requirements • Linking, integration and re-use of diverse but related resources • Interchange between institutional repositories and information management systems • Personalisation tools and features for repositories • Use of RSS and ATOM • Common repository interfaces • Personalisation 	Total funds £500,000 Up to 8 projects Max £100,000 per project 6 months to 2 years duration	G39–G43

¹ Repositories and Preservation Briefing Paper http://www.jisc.ac.uk/circular04_06_briefing_papers

Call Area	Theme/context	Description	Funds	Paragraph
Programme Strand C: Discovery to Delivery				
II	Projects that develop solutions or demonstrate issues that will help offer improved discovery to delivery	<ul style="list-style-type: none"> a) Version identification framework b) Persistent identifier interoperability demonstrator c) Federated access management and repositories d) Semantic interoperability demonstrator 	<p>Total funds £440,000</p> <p>4 projects £110,000 per project</p> <p>10 months to 1 year duration</p>	G44–G56
Programme Strand D: Repository Start-Up and Enhancement Projects				
III	Start-up and enhancement funding for repositories	<ul style="list-style-type: none"> a) Repository start-up projects b) Repository enhancement projects 	<ul style="list-style-type: none"> a) Up to 20 projects of maximum funding of £30,000 each for Repository start-up projects and £60,000 for consortium projects b) Between 10 to 15 projects, between £80,000 and £300,000 for Repository enhancement projects <p>2 years duration</p>	G57–G68

Call Area	Theme/context	Description	Funds	Paragraph
Programme Strand H: Digital Preservation and Records Management				
IV	Projects that develop preservation assessment tools, models for preservation service provision and repository preservation capability.	<ul style="list-style-type: none"> a) Digital preservation assessment across the lifecycle b) Models and implementation of preservation services (2 projects) c) Develop preservation enhancement tools (2 projects) 	<p>Total funds £800,000</p> <p>5 projects</p> <ul style="list-style-type: none"> a) Digital preservation assessment – up to £150,000 b) Models and implementation of preservation services – up to £200,000 each c) Develop preservation enhancement tools – up to £125,000 each <p>18 months to 2 years duration</p>	G69–G88
Programme Strand I: Shared Infrastructure Services				
V	Shared infrastructure services for discovery, repositories and curation	<ul style="list-style-type: none"> a) Pilot implementation of a licence registry b) Pilot national name and factual authority service c) Scoping of an architecture to support digital policy management (rights) d) Scoping of a terminology registry 	<p>Total funds £480,000</p> <p>4 projects</p> <ul style="list-style-type: none"> a) and b) £200,000 max per pilot implementation project c) and d) £40,000 max per scoping study <p>The pilot implementations will run over a period of 2 years</p> <p>The scoping studies will run for 6 months</p>	G89–G110
Project proposals should be submitted under one call area only, but relationships across the areas should be highlighted within proposals.				
<p>This appendix must be read in conjunction with the main body of JISC Circular 4/06 which can be found at: http://www.jisc.ac.uk/funding_circular04_06</p>				

Structure of this document:

Summary.....	1
Background.....	5
Information Environment (IE)	7
e-Framework	7
Programme Context.....	9
Submission of Proposals	9
Eligibility	9
Introduction	9
Outcomes and Benefits of the Projects	11
Call Area I – Tools and Innovation Projects.....	12
Call Area II – Discovery to Delivery Projects	15
II a) Version Identification Framework	15
II b) Persistent Identifier Interoperability Demonstrator.....	16
II c) Federated Access Management and Repositories	16
II d) Semantic Interoperability Demonstrator	16
Call Area III – Repository Start-Up and Enhancement Projects	18
III a) Repository Start-Up Projects	19
III b) Repository Enhancement Projects	19
Call Area IV - Digital Preservation and Records Management Projects.....	21
IV a) Digital Preservation Assessment.....	22
IV b) Models and Implementation of Preservation Services	23
IV c) Preservation Enhancement Tools	24
Call Area V – Shared Infrastructure Service Projects.....	26
V a) Licence Registry Pilot Implementation Project	26
V b) Name and Factual Authority Pilot Implementation Project	27
V c) Digital Policy Management Scoping Study	28
V d) Terminology Registry Scoping Study	28
Repositories and Preservation Proposal Cover Sheet	30

Background

- G4. The development and implementation of digital repositories and their supporting infrastructure is important to the delivery of the JISC Strategy. JISC aims to develop:
- A common information and communications environment, including addressing intellectual property rights, interoperability standards, digital preservation and appropriate technology;
 - The effectiveness of scholarly communication and digital resources in support of research, learning and teaching, especially through sustainable content management.
- G5. Digital repositories are one aspect of the information and communications environment that can help achieve these aims. They can help institutions to develop coherent and coordinated approaches to the capture, identification, storage, preservation, retrieval and dissemination of their intellectual assets. These intellectual assets go beyond normal publishing regimes, and may include audiovisual objects, datasets, presentations, learning materials and research works. A managed approach to these assets enhances opportunities for efficient use of existing research, improved learning experiences and encourages collaboration within and between different disciplines and groups.
- G6. There are examples of repository development that are changing the global information landscape. ARROW in Australia² and DAREnet³ in the Netherlands show how joined-up national approaches can deliver innovation and value to the user. The JISC Repositories and Preservation Programme builds on previous work and aims to help to establish a network of digital repositories and supporting services that will help to radically improve content use and curation. The programme will continue to bring 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. The programme will also look to address issues of digital preservation and resource discovery alongside the development of repositories.
- G7. The programme will help to develop digital repositories so they not only provide the benefits mentioned above but so they form part of the JISC Information Environment.⁴ They can do this by supporting open standards and common implementations of those standards. The Information Environment is guided by a high-level architecture and set of standards and protocols that promotes interoperability for the discovery of digital resources and their management.
- G8. This programme⁵ is a £14 million investment in higher education (HE) repository and digital content infrastructure, which aims to establish a network of digital resources and services. It will help to equip the HE sector with the skills necessary to use and maintain digital assets created through their work. The funding will be allocated to different projects and calls for proposals over the next 12-month period and expenditure will be complete in 2009.
- G9. In April 2006 JISC issued a call for projects as part of this Repositories and Preservation Programme. Twelve projects were funded. These are referred to within

² ARROW: <http://arrow.edu.au/>

³ DAREnet: <http://www.darenet.nl/en/page/language.view/home>

⁴ JISC Information Environment: http://www.jisc.ac.uk/index.cfm?name=ie_architecture

⁵ Within the £14M funding allocation there will be the Repositories and Preservation Programme, which also includes activities that support the development of shared machine to machine infrastructure to support resource discovery and curation.

the Repositories and Preservation Briefing Paper.⁶ In addition there have been a number of programmes in the areas of digital repositories and preservation that JISC has previously supported, some of which are still on-going.

- G10. For the purpose of this call for proposals the following definitions should be used to guide the remit of projects.

*Digital Repository*⁷

‘A digital repository is differentiated from other digital collections by the following characteristics:

Content is deposited in a repository, whether by the content creator, owner or third party; the repository architecture manages content as well as metadata; the repository offers a minimum set of basic services, eg put, get, search, access control; the repository must [aim to] be sustainable and trusted, well supported and well managed.

Enhancing access to scholarly communications has been a major driver for establishing, both institutional and subject-based archives. Many, though by no means all repositories, support “open access”, at least in part. Open access repositories [have] the following characteristics: the repository must provide open access to its content (unless there are legal constraints), the repository must provide open access to its metadata for harvesting.”

Digital Preservation and Records Management

Digital preservation can be defined as ‘the series of managed activities necessary to ensure continued access to digital materials for as long as necessary’.⁸

Records management is a process for the systematic management of all records and the information or data that they contain. The core concept in records management is that of the ‘records lifecycle’, a model for ensuring the consistent and controlled management of records during all stages in their ‘life’: from creation through to final disposition, either in the form of destruction or permanent preservation.⁹

Digital information is subject to potential loss due to easy manipulation, hardware and software obsolescence. Digital preservation and records management are closely related activities, which will help ensure that content held within digital repositories remains accessible and retains its authenticity and integrity for as long as it is needed.

⁶ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

⁷ Definition adapted from Anderson, S. and Heery, R. (2005) *Digital Repository Review*. AHDS, UKOLN and JISC: http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf

⁸ Jones, M. and Beagrie, N. (2002) *Preservation Management of Digital Materials: A Handbook*. Digital Preservation Coalition, York: <http://www.dpconline.org/graphics/intro/definitions.html>

⁹ Cf. JISC infoNet: <http://www.jiscinfonet.ac.uk/InfoKits/records-management/what-is-records-mgt>

Information Environment (IE)

- G11. Projects must work in the context of the Information Environment and the e-Framework, which provide complementary service-oriented approaches for systems and services in education and research.
- G12. One of JISC's key objectives is to build an online IE that provides secure and convenient access to a comprehensive collection of scholarly and educational material. This is based on the Information Environment Technical Architecture,¹⁰ which specifies a set of standards and protocols that support resource discovery and curation of digital resources as part of learning, teaching and research activities. JISC is developing components of this architecture to test out and promote the use of these standards and to help provide appropriate services to the community. Projects will need to adhere to the IE technical standards and will also contribute to the development and implementation of the IE.

e-Framework

- G13. The e-Framework for Education and Research is an international initiative, by JISC and Australia's Department of Education, Science and Training (DEST), to explore the potential benefits of applying a service-oriented approach to the provision of ICT infrastructure for education and research, and where successful to support its broader adoption by institutions and their suppliers (see: e-Framework Overview Briefing¹¹).
- G14. Its main provision is an evolving knowledge base, presented as the e-Framework web site, containing information and links to further information on services and their effective use. There is thus technical information about open service standards covering pre-specification prototypes, specifications and standards under development, those being implemented and those in general use. There is also information about the usage of services, the domains and context of use, the human level, tasks and processes being supported, scenarios and case studies of how humans make use of service-based applications to accomplish these and technical information about the ways in which the services were brought together (see: Domain, Process and Service Models Briefing¹²).
- G15. These two aspects interact and are expected to evolve as it becomes clear which areas benefit most from a service-oriented approach, and which less. Also commonalities across tasks will refine the services that support them and the provision of services will enable more flexible implementations that allow new practices and processes to evolve.
- G16. Projects are expected to work within the e-Framework by making use of its available information and by contributing to its further development, with the emphasis on the latter in the early stages. How this happens will depend on the nature of the project.
- G17. Where projects carry out technical development, this should be done within the service-oriented approach of the e-Framework and, where possible, should expose and consume functionality via Web Services (SOAP or REST). However, other technical approaches are permissible, where appropriate, eg where existing standards are already in use (such as Z39.50), or where Web Services do not yet meet

¹⁰ JISC Information Environment Technical Architecture

<http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>

¹¹ e-Framework and SOA Briefing http://www.jisc.ac.uk/circular04_06_briefing_papers

¹² e-Framework DP and SM Briefing http://www.jisc.ac.uk/circular04_06_briefing_papers

performance or functional needs (such as for secure transactions) (see: Web Services & SOA Briefing¹³).

- G18. Other projects that will not be providing services themselves are encouraged to use Web Service-enabled tools and applications within their own environment. All projects should be able to contribute to the knowledge base that the e-Framework is developing. This can include domain, practice and process models, scenarios and use cases, and good practice guidelines on the internal and cross-institutional implementation of the technology, as well as information about the service definitions they have used or developed.
- G19. At the interface between users and services there is emerging a thinner but more flexible and capable technical user environment layer. This is taking two forms: the continued development of portal technology and the so-called rich client platform. Both of these provide a capable plug-in software framework that can take much of the work out of developing the user interface, allowing concentration on the coordinating functionality of the tool or application (see: Service Oriented Application Integration Layer Briefing¹⁴).
- G20. The goal is to record both relevant project outputs and outcomes, in order to support those seeking to implement a service-oriented approach. By sharing developments and experiences internationally, we hope to be able to do this more effectively and rapidly than if done alone, and by developing and adopting open standards, establish a wider and more open market enabling costs to be reduced.
- G21. Where appropriate, all projects will contribute the following to JISC and the e-Framework:
- Domain maps that reflect consensus-based practices, processes and supporting systems. These will take the form of a functional specification and technical architecture model of the services and components.
 - Service usage models (SUMs). These provide a description of the needs, requirements, workflows, management policies and processes within a domain and the mapping of these to a design of a structured collection of service genres, service expressions and other resources, that can be used to implement software applications within the domain.
 - Services – where not already recorded in the e-Framework, additional service genres or expressions identified in the course of the projects, or any emergent interoperability specifications that could become part of a service expression, should be contributed.

¹³ e-Framework Web Services and SOA Briefing http://www.jisc.ac.uk/circular04_06_briefing_papers

¹⁴ e-Framework Service Oriented Application Integration Layer Briefing
http://www.jisc.ac.uk/circular04_06_briefing_papers

Programme Context

- G22. Please refer to the Repositories and Preservation Briefing Paper¹⁵ for a brief summary of all of the Repositories and Preservation Programme areas (A–J) that are planned under the total programme expenditure. A table within the briefing paper identifies, at a high level, the relationships between programme strands. This particular call asks for proposals in areas B, C, D, H and I. Bidders should be aware that elements of the programme are still being scoped. It is, however, hoped that this table is a useful indication of the potential dependencies and relationships across the programme.
- G23. The briefing paper includes information on relevant JISC work, technical and organisational considerations, including the Information Environment and the e-Framework for Education and Research and some other work outside of JISC that bidders should be aware of. Relevant work is referenced to provide context for bidders.

Submission of Proposals

- G24. Information on the bidding process and submission of proposals is set out in the main text of the circular. Bids in response to this call for projects should be sent to repositories-bids@jisc.ac.uk, with the name of the lead institution in the subject line. If more than one bid is submitted by an institution, these must be submitted in separate messages.

Eligibility

- G25. The eligibility criteria for the projects in this appendix are as those set out in the overarching circular.

Introduction

- G26. A number of projects are invited in five of the ten Repositories and Preservation Programme strands. These are:
- I. Tools and Innovation (Strand B)**
 - II. Discovery to Delivery (Strand C)**
 - III. Repository Start-Up and Enhancement Projects (Strand D)**
 - IV. Digital Preservation and Records Management (Strand H)**
 - V. Shared Infrastructure Services (Strand I)**
- G27. The Repositories and Preservation Briefing Paper¹⁶ gives details of all the strands within the programme.
- G28. The Tools and Innovation area of projects is concerned with the development of software and management tools to assist in the use of digital repositories to support end-users, workflow and specific business requirements. Projects in this area might also test out or explore less mature technology and techniques to investigate how these can be applied to digital repositories to improve their effectiveness. Up to £500,000 is available under the call for this area. A maximum of eight projects will be funded. The maximum funding for any one project is £100,000. It is anticipated that projects will vary in length from six months to two years, commencing in February 2007.
- G29. The Discovery to Delivery area of projects looks specifically at creating the technical conditions required for interoperability for resource discovery and access to resources held within digital repositories. The projects asked for in this call are concerned with persistent identifiers, version identification, federated access management and

¹⁵ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

¹⁶ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

semantic interoperability. Up to £440,000 is available under this call. Four projects will be funded up to a maximum of £110,000 each. Projects should last between ten months and one year in duration commencing in February/March 2007. To further the implementation of interoperability, JISC is commissioning separate work to develop application profiles for content.

G30. The Repository Start-Up and Enhancement Projects are expected to offer sustainable and embedded growth of interoperable repositories for learning, teaching and research. Tie-in with the developing national (international) context is important as is progress in the deposit of content and supporting institutional policies. Enhancement Projects will also help to develop and evaluate enhanced repository solutions. For both types of projects JISC will only fund up to 50% of the total project cost.

- For Repository start-up projects the maximum funding available is £30,000 and it is anticipated that approximately 20 of these projects will be funded. For consortium start-up projects the maximum funding available is £60,000. Value for money is an important criterion in the evaluation of project proposals.
- Funding for Repository enhancement projects is between £80,000 and £300,000. Project proposals are welcome for enhancement project funding. Please note that the upper end of this funding scale is to enable consortia project proposals. It is anticipated that between 10 and 15 projects will be funded. Value for money is an important criterion in the evaluation of project proposals.

G31. The Digital Preservation and Records Management area of projects is concerned with the development of preservation assessment tools and methodology, models for preservation service provision and enhancement of repository preservation capability throughout the lifecycle of digital information. Up to £800,000 is available under this call for this area and it is anticipated that five projects will be funded. One project is sought for digital preservation assessment and will be funded up to £150,000. Two projects are sought for models and implementation of preservation services and funding of up to £200,000 is available for each project. Two projects are sought within preservation enhancement tools. Funding of up to £125,000 is available for each project. The projects are expected to run for 18 months to 2 years, commencing in February/March 2007.

G32. The Shared Infrastructure Services area of projects will develop machine to machine services as part of the Information Environment. Shared Infrastructure Services support the integration of networked services and digital resources and provide shared functions to deliver interoperability and economies of scale. This area builds on a programme of previous funding called 'Shared Services'¹⁷ and is informed by a recent Shared Infrastructure Services Synthesis Report.¹⁸ Repositories and digital preservation can both benefit from this approach. An example of an existing Shared Infrastructure Service is the Information Environment Service Registry (IESR).¹⁹

Four projects are invited under this call. Two pilot implementation projects are called for and are expected to run over a period of two years. These are a pilot licence registry and a pilot national name and factual authority service. The pilot implementations will be funded up to a maximum of £200,000 each. The two other invited projects are scoping studies that are expected to run over a period of six to eight months. These are a digital policy management architecture study and a terminology registry study. The two scoping studies will be funded up to a maximum of £40,000 each.

¹⁷ *Shared Services Programme* http://www.jisc.ac.uk/index.cfm?name=programme_shared_services

¹⁸ *Shared Infrastructure Services Review* http://www.jisc.ac.uk/Shared_Infrastructure_Services_Review_Sep_06

¹⁹ *Information Environment Service Registry (IESR)* http://www.jisc.ac.uk/project_iesr.html

- G33. Bidders should refer to the Repositories and Preservation Briefing Paper²⁰ for important contextual information.

Outcomes and Benefits of the Projects

- G34. Through the *Tools and Innovation* projects, informed change will be stimulated by piloting new technologies and approaches. The benefits will be: improved understanding of new technologies and their application to repositories; increased flexibility and usability of repository systems; and improved digital content tools.
- G35. Through the *Discovery to Delivery* projects, models for using technology and applying standards will be provided that can be used at departmental, institutional, regional, national and trans-national levels. The benefits are better understanding and practice relating to rich search and discovery of content within repositories. Access will be enhanced through the project outcomes directly informing the anticipated Intute Repositories Search Service Project that is being taken forward outside of this call for proposals.
- G36. Through the *Repository Start-Up and Enhancement Projects*, more repositories will be established that apply good practice in their start-up and management of repositories, particularly in building content deposit into workflows in learning, teaching and research. They will be supported in adopting relevant standards in order to improve interoperability between repositories and other services. Other benefits will be: helping institutions manage digital assets effectively; an increase in the volume of content held in interoperable repositories; and through the enhancement projects, further development of solutions that help to improve digital repositories in both technical and organisational areas.
- G37. Through the *Digital Preservation and Records Management* projects, more clarity will be achieved in terms of the roles and responsibilities for digital preservation. Those that manage digital repositories will feel more confident and supported when assessing and undertaking digital preservation actions.
- G38. Through the *Shared Infrastructure Services* projects, there will be increased interoperability and integration of networked services and content. The user will consequently be provided with a more coherent experience within the diverse digital environment, enabling use of richer and more appropriate content. Duplication of effort in service provision will be reduced, and economies of scale will be achieved. Further implementations and understanding of machine to machine services within the Information Environment and how they can support repositories and preservation to underpin education and research will be delivered.

²⁰ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

Programme Strand B: Tools and Innovation

Call Area I – Tools and Innovation Projects

- G39. JISC acknowledges that a wide range of content formats, implemented systems, technologies and established practice already exists in the area of digital repositories. However, there is a perceived requirement for tools to be further developed to improve the efficiency and quality of 'repository functions' and use of repositories. Within this area, projects concerned with software and tools that might be related to a variety of digital repository functions will be funded. These 'repository functions' include adding content to repositories and managing and using repository content. Under this area projects should either deal with tools that support workflow and use, or new technologies that are less mature in terms of implementation. 'Tools' refers to software and/or to the articulation of approaches, processes and criteria that can be applied to help manage and use digital repositories. Projects therefore may produce software, models for practice or in some cases evaluative reports that help to develop practice and understanding.
- G40. Projects might support exploration and development of tools within the following broad areas; projects may also address more than one area:
- Interactions between material held within institutional repositories and institutional bibliographic databases, research and information management systems and records management systems etc, particularly with a view to establishing workflow and common practices
 - Linking up or re-using data and/or resources held within different types of repositories (eg ePrints with research data, learning objects with publications) or repositories holding multiple media (text, images, video, data), and to investigate the challenges involved in doing this
 - Provision of value added services including the role of learned societies and/or learned society publishers. Scoping and development of demonstrator tools, for example in partnership with learned societies, to illustrate the ways in which value might be added to repository content. An example might be a tool to notify users that repository content had a particular measure of quality, as defined and applied by a trusted organisation
 - Semi-automatic solutions to indexing and classification in pilot projects. Investigate currently available tools for automatic indexing and classification. Bidders are asked to note the current AHDS Project²¹
 - Demonstrate relationships between knowledge organisation systems (KOS) and text mining and how these approaches might benefit digital repositories. For example, KOS can support text mining and text mining can be used to update and enhance KOS. Bidders should refer to the UKOLN Terminology Services and Technology report²²
 - Investigate metadata generation and validation tools, particularly in the ingest function within the content lifecycle. Consideration should be given to the roles of authors, cataloguers, librarians and other other institutional roles involved in the deposit and ingest in repositories etc. It would be valuable for a project to investigate the availability of tools that enable the checking of object metadata against a local schema. This might include checking against controlled

²¹ <http://ahds.ac.uk/about/projects/metadata-generation/index.htm>

²² *Terminology Services and Technology Report*

http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06

vocabularies, automatic checking of mandatory fields and or support for checking local authority files. These tools need to be embedded in the repository systems and workflows and should deal with both individual items and batch processing. Bidders are asked to note the call for automatic metadata extraction or generation tools for preservation metadata in area IV of this call. Proposals that consider the automation of both types of metadata are allowed, but should be submitted under one strand and cross reference the fact that they address both issues

- Personalisation tools and features and how these might be applied to repositories. Bidders should refer to the JISC Personalisation report²³
 - The use of RSS and ATOM as additional standards to aggregate metadata and content. Bidders should refer to the Linking UK Repositories report²⁴
 - Usage metrics that help to indicate impact and trends in repositories. Projects in this area should help to elucidate the complex issues surrounding this area and might include pilot demonstrators. Bidders should refer to previous work²⁵
 - Development and testing of common repository interfaces that can enhance the use of repositories. Work in this area should be cognisant of the deposit API work being lead by UKOLN and a recent workshop on Augmenting Interoperability Across Scholarly Repositories, which has now resulted in some related work funded by the Mellon Foundation²⁶
 - Methods and tools to assess the use and requirements of repository services, for example how might deep log analysis be used as a method for more efficient repository management or management of digital services that are considered part of the information environment within institutions or at a shared national level
- G41. Projects under the Tools and Innovation strand should seek to examine the use and development of tools within the context of use cases and business requirements, helping to articulate and establish the benefits of the tools, and to develop solutions to the issues explored. Project proposals in this area will be considered in terms of their contribution to the programme as a whole; proposals therefore should highlight opportunities for cross-project working. It will be essential for projects under Tools and Innovation to also report on cultural and technical issues including usability and interoperability. Sustainability should also be addressed, for example how will the tools be adopted, how obstacles to long-term take-up might be tackled. This might involve dissemination or testing with, and the influencing of, software systems working within the digital repository arena. It is anticipated that most of the projects in this area will work with the e-Framework, by following service-oriented approaches and sharing information and contributing documentation to the JISC DEST e-Framework initiative²⁷

Further Information

- G42. All enquiries regarding this section of the call should be sent to Phil Vaughan (tel: 07810 814476; email: rpteam@jisc.ac.uk).

²³ http://www.jisc.ac.uk/index.cfm?name=project_personalise

²⁴ www.jisc.ac.uk/uploaded_documents/Linking_UK_repositories_report.pdf

²⁵ Interoperable Repository Statistics project <http://irs.eprints.org/> and see the work of Johann Bollen <http://www.dini.de/veranstaltung/workshop/oaimpact/programm.php>

²⁶ http://www.ukoln.ac.uk/repositories/digirep/index/Deposit_API and see <http://msc.mellon.org/Meetings/Interop/>

²⁷ e-Framework for Education and Research www.e-framework.org

G43. Any enquiries regarding the proposal submission process should be sent to Nike Holmes (tel: 0117 931 7427; email: n.holmes@jisc.ac.uk).

Programme Strand C: Discovery to Delivery

Call Area II – Discovery to Delivery Projects

G44. In this area JISC is seeking to fund four projects that develop solutions or demonstrate issues that will help offer improved discovery to delivery. Projects are sought in the following areas:

- a) Version identification framework**
- b) Persistent identifier interoperability demonstrator**
- c) Federated access management and repositories**
- d) Semantic interoperability demonstrator**

G45. Funding of up to £110,000 is available for each project.

G46. The results of these projects will support the Information Environment and enhance the implementation of standards and protocols to deliver a richer discovery to delivery experience with particular relevance to the evolving network of repositories within the UK. They should also help to inform advice in the Repositories Support Project and the proposed JISC search project across repositories`. The projects will also be able to contribute to service definitions and usage models within the e-Framework where appropriate.

G47. It should be noted that a few projects are being directly funded outside this call for proposals to create application profiles in the area of images, moving pictures and sound and geospatial data. It is also intended that studies in the areas of scientific data and learning materials will be commissioned outside of this call in order to ascertain what is required for interoperability for distributed repositories (see the table within the Repositories and Preservation Programme briefing paper²⁸).

G48. The scope of the four Discovery to Delivery projects is outlined below.

II a) Version Identification Framework

G49. A project is sought to build and produce a version identification framework.

G50. This area needs to take into account the River Project recommendations,²⁹ in which it states 'that JISC should research definitive sets of version identification requirements from researchers, teachers, learners, information professionals and other stakeholders before finalising policies in this area. Specifically this should be to gain a better understanding of the way people are working with documents, images, learning objects (LOs) and other resources, including the tools that they are using and the communities of practice that they work in. Before moving into standards development it will be essential to understand how people are using their desktops, LANs Web Pages, Content Versioning Systems (CVS) and Wikis that allow sophisticated handling of workflows, revision and version identification. An understanding of social interoperability is critical to the successful implementation of technical interoperability.' The project will need to be aware of the work of the NISO/ALPSP versions working group³⁰ and the Versions Project³¹ at the London School of Economics. Once a more rigorous requirements exercise has been completed then a framework needs to be developed that will enable users to consistently identify versions of digital objects within digital repositories, bearing in mind the diversity of content types that repositories might hold. The framework should build on the straw man developed in

²⁸ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

²⁹ *River Project Report* http://www.jisc.ac.uk/uploaded_documents/RIVER%20Final%20Report.pdf

³⁰ *NISO/ALPSP versions working group* http://www.niso.org/committees/Journal_versioning/JournalVer_comm.html

³¹ *Versions Project* www.lse.ac.uk/library/versions

the River project. The project will need to work in collaboration with proposed Intute Repositories Search Project³², as version identification may be an aspect of the search criteria.

II b) Persistent Identifier Interoperability Demonstrator

- G51. A project is sought to build a persistent identifier interoperability demonstrator across the variety of identifier schemes.
- G52. Considerable amounts of work have been done on standardising identification schemes for a particular community or sector. Equally, much work has been done on creating standard or reference metadata sets that can be used to associate key metadata descriptors with content. Much less work has been done on the impact of cross-sector (by sector we mean a particular media sector's) working. Relatively little is understood about the effect of using one industry's identifiers in another, or on attempting to import metadata from one identification scheme into a system based on another. In the long term it is clear that interoperability of all these media identifiers and metadata schemes will be required. A variety of identifier schemes are adopted across digital repositories in order for a network of repositories to flourish, interoperability across schemes will be required. One project, perhaps working in partnership across a consortium to build a demonstrator, will be funded. The project will need to create use cases and to build an interoperability demonstrator in order to test interoperability across different identifier systems, eg Handle, DOI etc. Bidders should take note of the Digital Object Identifier for Publishing and the e-Learning Community Report.³³ This area of work is likely to feed into the application profiles that will be developed outside of this call for funding and work in collaboration with the proposed Intute Repositories Search Project³⁴ as persistent identifiers may be an aspect of the search criteria.

II c) Federated Access Management and Repositories

- G53. A project is sought in the area of Federated access management. The project should include managing user identities, enabling repositories to act as service providers to manage authorised user actions, attributes and roles for typical workflows in the use of institutional repositories, and linking this to externalised authorisation. Potential projects may build on tools developed and implemented within the JISC Core Middleware Technology Development programme,³⁵ including the JISC Shibboleth Project,³⁶ which looked at making the ePrints.org software Shibboleth-enabled. Projects in this area will be invited to also consider how access management can be handled at a collection level and also at an item level within a repository. For example, a repository may have items within a given collection, which are restricted as well as open. How would this issue be dealt with? Briefing papers on JISC's developments in the area of Federated Access Management are available.³⁷

II d) Semantic Interoperability Demonstrator

- G54. A project is sought to build a demonstrator to test the combination and comparison of controlled and folksonomy approaches to semantic interoperability. The project should examine the usefulness of the approaches and evaluate cost and benefit issues. The

³² Intute Repository Search Project (Interim URL) <http://irs.ukoln.ac.uk>

³³ www.jisc.ac.uk/index.cfm?name=project_tso

³⁴ Intute Repository Search Project (Interim URL) <http://irs.ukoln.ac.uk>

³⁵ Core Middleware Technology Development Programme www.jisc.ac.uk/index.cfm?name=programme_middleware

³⁶ Shibboleth Project <https://gate-test.library.lse.ac.uk/dspace/bitstream/1988/2809/1/eprints-fin>

³⁷ Federated Access Management Infrastructure www.jisc.ac.uk/federation.html

demonstrator should investigate how these approaches might enhance resource discovery across repositories. The project should identify and test use cases. The recent Terminologies Services and Technology Report³⁸ reviews different kinds of vocabularies, according to their structure and their intended purpose. This report should form the starting point for the project. The following extracts from the report are given as an indication of potential project scope, however bidders should not be constrained by this and should develop a methodology for the demonstrator that will offer a robust assessment that can then be taken forward in the development of digital repository networks.

One type of terminology service attempts to increase consistency and improve access to digital collections and Web navigation systems via vocabulary control. Vocabulary control aims to reduce the ambiguity of natural language when describing and retrieving items for purposes of information searching. Another type of terminology service is not concerned with consistency but with making it easier for end-users to describe information items and to have access to other users' descriptions. This results in vocabularies (folksonomies) that may not be controlled, at least initially.

Recommendations:

- *Experiment with combination of KOS-based controlled indexing with an established vocabulary and free (social) tagging for research purposes in a specific discipline, optimised for discovery and retrieval*
- *Experiment with potential for automatic linking of tags to facets, controlled vocabularies and authorities*
- *Integrate tagging with existing services such as repositories, OPACs, (RDN/Intute) subject gateways, Digital Libraries, KOS creation and management systems, museum exhibitions and catalogues, metadata enhancement services etc*
- *Comparison study between different types of user participation: annotation, recommendation, personalisation, restructuring of information, categorisation, concept space, concept maps, topic map tools. This could inform a prototype integrating different types of user participation with social tagging*

Further Information

- G55. All enquiries regarding this section of the call should be sent to Balviar Notay (tel: 020 7848 2670; email: rpteam@jisc.ac.uk).
- G56. Any enquiries regarding the proposal submission process should be sent to Nike Holmes (tel: 0117 931 7427; email: n.holmes@jisc.ac.uk).

³⁸Terminology Services and Technology Report

http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06

Programme Strand D: Repository Start-Up and Enhancement Projects

Call Area III – Repository Start-Up and Enhancement Projects

- G57. JISC wishes to increase the number of interoperable repositories and the content held within them and to give institutions the opportunity to contribute to, and make use of, the growing UK repositories network. Some institutions that have implemented repositories have encountered cultural barriers to embedding repositories and challenges surrounding deposit. Whilst good progress has been made and approaches have emerged through experience, JISC intends these projects to contribute to a step change in overcoming these challenges and helping to populate repositories.
- G58. The aim of this area is to stimulate the population of existing repositories, enhance interoperability and embed repositories within institutional working practices such as research and learning, as well as supporting the establishment of repositories where this will be sustained in the future.
- G59. There are two types of projects sought in this area in recognition of the different stages that institutions are at in terms of repository implementation. These are as follows:
- a) Repository Start-Up Projects**
 - b) Repository Enhancement Projects**
- G60. For both types of projects a range of repositories are eligible for funding. For example: institutional open access repositories of research outputs; repositories of teaching and learning materials, integrated with content management and/or content delivery systems such as Virtual Learning Environments (VLEs); subject-based or media-based repositories that address an identified need within or across institutions. Repositories may also cater for a broad range of content, or in some cases it is recognised that an institution may have requirements for more than one repository. Repository projects might work in collaboration with user communities, national data centres, research councils or academic centres. Relationships between national data centres/archives and institutional repositories are of interest to the programme. JISC will seek to fund a coherent set of projects that helps to address a range of issues as well as fulfilling the aims of the programme.
- G61. It is essential for both types of projects to demonstrate institutional commitment to the long-term use and maintenance of digital repositories in supporting the institutional mission. JISC will therefore only fund up to 50% of the total cost of the project. Evidence of institutional commitment should be clearly addressed in proposals, including a commitment to producing an open access policy during the life of the project. Project proposals for both types of projects must outline how the project fulfils an institutional need in terms of learning, research or administration and show their strategies for content population or in the case of start-up projects how they intend to tackle this issue. All projects will need to work within the framework of technical standards that are supported by the Information Environment and those that emerge through the programme in the interest of interoperability. A list of some of the issues that projects might need to consider in establishing or enhancing repositories is contained within the Repositories and Preservation briefing paper.³⁹
- G62. All projects will be expected to share lessons and to participate in the evolving JISC-funded network of digital repositories. For example this will mean exposing content via the proposed Intute Repositories Search Project where appropriate and working to improve or use guidance developed via the Repositories Support Project (Programme Strand A) that is led by SHERPA⁴⁰ at the University of Nottingham and was funded as

³⁹ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

⁴⁰ SHERPA Project www.sherpa.ac.uk/about.html

a result of the April 2006 programme call. Projects will also need to work with reference to the developing Shared Infrastructure Services and Digital Preservation and Records Management programme strands where appropriate; this might be in terms of inputting requirements or using these services as part of their repository solution.

III a) Repository Start-Up Projects

- G63. These projects should be concerned with the establishment and population of repositories.
- G64. Funding of up to a maximum of £30,000 is available for (single institutional) start-up projects. It is expected that a approximately 20 projects will be funded in this area, however, this will depend on the quality of proposal and the amount of funding requested. All projects will be aided by the new Repositories Support Project and should not duplicate this support infrastructure. However, institutions may wish to consider sharing technical expertise in order to achieve economies of scale. Where such arrangements can exist, each institutional bid should indicate this and illustrate the resulting savings. Alternatively, if consortia may wish to submit bids, the funding limit for such consortium start-up projects is £60,000. Value for money will be a significant evaluation criterion. Projects will need to engage with the JISC programme from their start date until the end of the programme in March 2009.

III b) Repository Enhancement Projects

- G65. Projects from institutions with established repositories are invited. As well as more general institutional repositories there may be specialist community requirements that need to further develop a supported distributed network of repositories for subject specific or media requirements.⁴¹ Enhancement projects should be concerned with embedding of repositories within institutional policies and learning, research and administration workflows in order to improve the value of repositories in terms of the resources they offer and functions they help staff undertake. These projects are encouraged to undertake further technical and organisational repository solutions and practices. Projects should wherever possible take a service-oriented approach and contribute to enhanced interoperability and common repository service interfaces, for example the deposit API⁴² that is in development (this may feed into work of the Repositories Research Team at UKOLN⁴³ and/or the e-Framework⁴⁴). Issues outlined in the other areas within this call may also be relevant to these projects, in particular those listed under areas B and H. Proposals for Enhancement Projects must include a URL of the existing repository/repositories on which they are building. If this is not possible, proposals should include descriptions and screenshots of the existing repository or where possible a guest log-in for evaluators.
- G66. Funding for repository enhancement projects will be between £80,000 and £300,000. It is anticipated that the maximum number of projects funded will be between 10 and 15. Consortia projects are welcome for enhancement project funding. Value for money will be a significant evaluation criterion. Please note that the upper end of this funding scale is to enable consortia projects.

⁴¹ There may be circumstances that vary from those scenarios mentioned above that the JISC Evaluation panel will consider within Repository Enhancement projects.

⁴² Deposit API http://www.ukoln.ac.uk/repositories/digirep/index/JISC_Digital_Repository_Wiki

⁴³ Repositories Research Team at UKOLN http://www.ukoln.ac.uk/repositories/digirep/index/Deposit_API

⁴⁴ E-Framework for Education and Research www.e-framework.org

Further Information

- G67. All enquiries regarding this section of the call should be sent to Amber Thomas (tel: 07920 534 933; email: rpteam@jisc.ac.uk).
- G68. Any enquiries regarding the proposal submission process should be sent to Nike Holmes (tel: 0117 931 7427; email: n.holmes@jisc.ac.uk).

Programme Strand H: Digital Preservation and Records Management

Call Area IV - Digital Preservation and Records Management Projects

- G69. Digital preservation is a relatively new activity for which there exists neither much experience nor commonly agreed best practice. Many unsolved organisational and technical issues make digital preservation a challenging task for those managing digital repositories. As an organisation working on behalf of the funding councils and the academic community, JISC has to date undertaken and funded much work to address the challenges of digital preservation and to advance the UK digital preservation agenda. This includes national initiatives such as the Digital Curation Centre (DCC),⁴⁵ the 'Supporting Digital Preservation and Asset Management in Institutions' programme⁴⁶ and membership of the Digital Preservation Coalition (DPC).⁴⁷ There is however a perceived need to embed digital preservation as an integral part of the workflow and to further develop the repositories' digital preservation capacity. Records management, owing to its lifecycle concept, its expertise and existing good practice in appraising records' archival value, contributes greatly to the ultimate goal of long-term preservation and should therefore be brought into the equation when considering digital preservation.
- G70. JISC intends to fund projects that build the preservation capacity for digital repositories. This will draw on experiences from various JISC-funded and non-JISC-funded development and encompass a wide range of activities. A number of digital preservation-related projects have already been funded under the April 2006 Repositories and Preservation Programme call, these are referred to in the supporting briefing paper.⁴⁸ Projects should avoid duplication and build on each others' work where appropriate.
- G71. Up to £800,000 is available for digital preservation projects under this call and it is anticipated that five projects will be funded. The projects are expected to run for 18 months to two years, commencing in February/March 2007. Projects are invited as follows:
- a) 1 project for digital preservation assessment – up to £150,000**
 - b) 2 projects on models and implementation of preservation services – up to £200,000 each**
 - c) 2 projects to develop preservation tools – up to £125,000 each**
- G72. Projects should where appropriate employ open and emerging preservation standards such as the Open Archival Information Systems (OAIS) Reference Model⁴⁹ and the PREMIS Data Dictionary for Preservation Metadata⁵⁰.
- G73. Using repositories to manage institutional and organisational intellectual assets requires that all are able to trust the ability of the repositories to secure the information over the long term. 'The fact that digital information is found within a trusted repository

⁴⁵ Digital Curation Centre (DCC): <http://www.dcc.ac.uk>

⁴⁶ Supporting Digital Preservation and Asset Management in Institutions:
http://www.jisc.ac.uk/index.cfm?name=programme_404

⁴⁷ Digital Preservation Coalition: <http://www.dpconline.org/graphics/index.html>

⁴⁸ *Repositories and Preservation Briefing Paper* http://www.jisc.ac.uk/circular04_06_briefing_papers

⁴⁹ *OAIS Reference Model* <http://www.ukoln.ac.uk/repositories/digirep/index/OAIS>

⁵⁰ *PREMIS Metadata* <http://www.loc.gov/standards/premis/>

may become the base upon which all further assessment of action builds'.⁵¹ Trust is an issue that could become a significant barrier for repositories and increase the complexity of digital preservation. The US Research Library Group (RLG) and OCLC have made significant progress in defining the attributes and responsibilities of trusted repositories,⁵² which include not only aspects of longevity of the digital objects stored within the repository but also the viability and financial sustainability of the organisation that operates the repository, the need for accountability, and the human aspects of competency and trustworthiness. The RLG and the US National Archives and Records Administration (NARA) released a draft Audit Checklist for the Certification of Trusted Digital Repositories in August 2005 for determining whether a digital repository can be certified as a trusted location for digital collections.⁵³ Although this work and the related audit and certification mechanisms are still in development, projects need to be aware of the requirements for becoming trusted digital repositories and where appropriate consider and incorporate the trust element into work.

- G74. The potential areas of funding under each heading are elaborated below. However, bidders are encouraged to put forward their own ideas provided that they have demonstrable relevance to the headings above.

IV a) Digital Preservation Assessment

- G75. Digital preservation assessment involves selection and appraisal, which are criteria and processes for identifying information for long-term retention and preservation. Selection and appraisal are iterative decision-making points within the lifecycle of a digital object connected to its active use and can be applied to help ensure resources are devoted to the long-term availability of the most valuable digital assets. These decisions concern strategic, organisational, economic, legal, regulatory and technical aspects of the digital objects, and are often determined and influenced by various internal and external factors. For analogue materials, there are existing preservation methodologies that help organisations to assess and prioritise their preservation needs. One example is the Preservation Assessment Survey (PAS), developed by the National Preservation Office (NPO) at the British Library.⁵⁴ The JISC-funded Digital Preservation Assessment Tool (DAAT) project has been adapting PAS for digital assets and investigating the necessary attributes for a digital preservation assessment tool.⁵⁵ Other examples of relevant work in this area include the Archive Ingest and Handling Test (AIHT) project at Stanford University, funded by the US Library of Congress as part of the National Digital Information Infrastructure and Preservation Program (NDIIPP),⁵⁶ and Decision Tree for Selection of Digital Materials for Long-term Retention, developed by the Digital Preservation Coalition (DPC) in the UK.⁵⁷
- G76. Making decisions on what digital objects to keep and how long to keep them inevitably requires that factors such as risks, value, benefits and costs are carefully considered.

⁵¹ Hirtle, Peter (2000), Archival authenticity in a digital age, in: *Authenticity in a Digital Environment*. Council on Library and Information Resources, Washington DC, pp. 8–23. Available at: <http://www.clir.org/pubs/reports/pub92/pub92.pdf>.

⁵² RLG-OCLC (2002), *Trusted Digital Repositories: Attributes and Responsibilities* – An RLG-OCLC Report, RLG, Mountain View, California. Available at: <http://www.rlg.org/legacy/longterm/repositories.pdf>.

⁵³ RLG-NARA (2005), *An Audit Checklist for the Certification of Trusted Digital Repositories*, RLG, Mountain View, California. Available at: <http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf>

⁵⁴ <http://liber.library.uu.nl/publish/articles/000035/index.html>

⁵⁵ DAAT is expected to complete by the end of September 2006: <http://www.ulcc.ac.uk/daat>

⁵⁶ AIHT: http://www.digitalpreservation.gov/technical/ndiipp_aiht_final_report.pdf

⁵⁷ Beagrie, Jones, and Digital Preservation Coalition, *Decision Tree for Selection of Digital Materials for Long-term Retention*: <http://www.dpconline.org/graphics/handbook/dec-tree.html>

Some of these are defined by fixed external drivers, such as legal and regulatory requirements, whilst others are more intangible and difficult to assess. Three JISC-funded projects, LIFE (Lifecycle Information for E-Literature), eSPIDA (effective Strategic model for the Preservation and disposal of Institutional Digital Assets) and the JISC Study of the Records Lifecycle, should provide useful insight into these issues and help repositories assess and justify the costs for digital preservation. LIFE is a collaborative project between the British Library and University College London, which investigates methods to attribute costs of digital preservation to identified stages within the lifecycle of digital collections. The eSPIDA project⁵⁸ is based at the University of Glasgow. It is developing a business-focused model to assess the value and the associated risks and costs of digital assets and explore ways to bring digital preservation to the core of institutional strategic planning. The JISC Study of the Records Lifecycle provides a generic records retention schedule for institutions and includes versions for both higher and further education institutions. It defines the functions, activities and transactions undertaken by institutions and the records they generate. It then maps these records against relevant legal and regulatory requirements or best practice to define the most appropriate retention period for such information.⁵⁹ A project is currently underway to revise this resource, with the updated version available from January 2007. Any projects funded in this area will be expected to use the most current version on its release.

- G77. It is anticipated that only one project will be funded under this area. The envisaged project should synthesise and build on the relevant work to date and develop a generalised digital preservation assessment methodology and/or retention and assessment decision-making tool, which incorporates consideration of types of content, file format and related technical issues, internal and external retention drivers and cost/benefit factors. It is essential that the methodology and tool are tested and supported by case studies and are applicable to a broad range of digital assets, from e-literature and learning objects to administrative and business records. The finished outputs should provide institutions with a decision path enabling them to reach an informed decision regarding both the required retention periods for each specific type of digital asset and the required preservation actions at each stage of their lifecycle.
- G78. The scenario of external preservation services provision should also be taken into consideration. The proposed project should therefore investigate and where possible quantify the criteria that digital repositories can employ to decide whether digital preservation should be performed in-house or outsourced to external preservation service providers.⁶⁰
- G79. Funding of up to £150,000 is available for the proposed project for a period of eighteen months to two years.

IV b) Models and Implementation of Preservation Services

- G80. JISC has a broad and inclusive vision for digital repositories and believes that they have an important role for digital preservation. Researchers, students, staff and institutions will require their ongoing availability and confidence in the future accessibility of the content within the repositories. Those managing repositories therefore naturally have the responsibility to ensure this for the content they are entrusted with managing by their institutions and researchers. JISC however acknowledges that preservation will not be the primary goal for all repositories and not

⁵⁸ eSPIDA: <http://www.gla.ac.uk/espida/>

⁵⁹ JISC Study of the Records Lifecycle for HE institutions: www.jisc.ac.uk/index.cfm?name=srl_structure
JISC FE Retention Schedule: http://www.jiscinfonet.ac.uk/projects/records-retention-fe/index_html/view

⁶⁰ The Arts and Humanities Data Service (AHDS) and The UK Data Archive (UKDA) are examples of external preservation service providers: <http://ahds.ac.uk/>; <http://www.data-archive.ac.uk/>

all repositories will choose, be able, or should be expected to carry out the full digital preservation activities themselves. Some of these will be fulfilled by external service providers and some may be shared between repositories. The JISC-funded SHERPA DP⁶¹ and PRESERV⁶² projects are in the forefront of exploring different models for the provision of digital preservation services for institutional repositories.

- G81. There is however no commonly agreed practice as to how this should be done and the relationship and interactions between the repositories and external preservation providers need to be investigated in more detail.
- G82. Projects are invited to develop, implement and test various models for the provision of preservation services for digital repositories. The models must clearly define roles and responsibilities at technical and organisational levels and investigate how trust can be established and built into the provision and assessment of services. Where appropriate, the project should as much as possible make use of existing infrastructural services that support digital preservation. Examples of such infrastructural services include PRONOM,⁶³ developed by The National Archives, and Representation Information Registry, developed by the DCC.⁶⁴ The projects are also strongly encouraged to make use of existing preservation services and collaborate with national bodies and organisations such as The National Archives, the national libraries and research council-funded data services. In addition, the roles and responsibilities of the institutions, data creators, depositors and users should also form part of the overall consideration.
- G83. It is anticipated that two projects will be funded under this area, with funding of up to £200,000 each.

IV c) Preservation Enhancement Tools

- G84. Projects under this area will be concerned with the development and testing of tools that help increase productivity and ease the task of digital preservation throughout the lifecycle of digital objects. Such tools should support those involved with managing, using and creating content for repositories. Some examples of the types of projects that will be considered under this area are listed below. However, bidders are encouraged to put forward their own ideas.

- Automatic metadata extraction or generation tools for preservation metadata

The National Library of New Zealand Te (NLNZ) has for example developed a tool to programmatically extract preservation metadata from the headers of a range of file formats based on the NLNZ Preservation Metadata Schema. Similar tools could be usefully developed using as their basis the PREMIS Data Dictionary for Preservation Metadata. Such tools may also combine elements of a human-mediated interrogative tool to collect contextual metadata such as provenance, perceived value and intended use, which may be challenging to generate automatically.

Please note that metadata generation tools for resource discovery have also been called for under strand B of this call, Tools and Innovation. Proposals which consider the automation of both types of metadata are allowed, but should be

⁶¹ SHERPA DP: Creating A Persistent Preservation Environment for Institutional Repositories

(<http://ahds.ac.uk/about/projects/sherpa-dp/>)

⁶² PRESERV – Preservation Eprint SERVices (<http://preserv.eprints.org/>)

⁶³ PRONOM: <http://www.nationalarchives.gov.uk/pronom/>

⁶⁴ <http://registry.dcc.ac.uk/>

submitted under one strand and cross reference the fact that they address both issues.

- Tools that embed digital preservation into personal working environments or content creation process (eg desktop tools/pre-ingest tools) to improve the preservability of digital content that is to be ingested into repositories.⁶⁵
- Enhancement of or addition to functionalities of existing repository software to support preservation and to embed digital preservation into repository workflow. The PRESERV project⁶⁶ has for example adapted EPrints software to allow the collection and dissemination of preservation-oriented metadata. Similar approaches could be usefully developed for other repository software.
- Tools that enable consistent outputs of content from repositories that use different repository software systems; or tools that enable the extraction and management of repository content by preservation service providers.

G85. It is anticipated that two projects will be funded under this area, with funding of up to £125,000 each for a period of eighteen months to two years.

G86. Projects under the Digital Preservation and Records Management strand will need to consider issues of sustainability, dissemination and take-up. Cultural and technical issues are important as is working within the overarching programme. Some projects will have relevance to the e-Framework, especially where common interfaces and technical services are being developed and defined.

Further Information

G87. All enquiries regarding this section of the call should be sent to Helen Hockx-Yu (tel: 020 7484 1803; email: rpteam@jisc.ac.uk).

G88. Any enquiries regarding the proposal submission process should be sent to Nike Holmes (tel: 0117 931 7427; email: n.holmes@jisc.ac.uk).

⁶⁵ The National Library of Australia has for example developed XML Electronic Normalising of Archives (XENA), an Open Source Digital Preservation Software which uses the OpenOffice.org office suite to convert certain files into XML for preservation: <http://xena.sourceforge.net/index.html>.

⁶⁶ PRESERV Project www.jisc.ac.uk/index.cfm?name=project_preserv

Programme Strand I: Shared Infrastructure Services

Call Area V – Shared Infrastructure Service Projects

- G89. One of the strategic aims of the JISC Strategy 2004–2006 is: 'To develop solutions that enable the United Kingdom education and research communities to keep their activities world class through the innovative use of ICT – by providing a first-class sustainable infrastructure.' A key priority within this strategic aim is 'to develop a common, integrated information and communications environment'.
- G90. Shared Infrastructure Services are essential building blocks for an efficient and effective information and communications environment, as they help the user navigate through, and make best use of, a complex digital environment. JISC commissioned two reports from UKOLN relevant to this strand: the Shared Infrastructure Services review⁶⁷ and the Terminology Services and Technology review.⁶⁸ These reports synthesise efforts to date in this area and make recommendations for future JISC strategy. These have informed this element of the September 2006 call.
- G91. JISC intends to develop an Information Environment (IE) Testbed⁶⁹ that will help to develop understanding and use of the IE and, in particular, Shared Infrastructure Services will be part of this development. Projects might be required to work with the Testbed. Projects in this strand are also likely to contribute use cases and service definitions to the e-Framework and the IE.
- G92. Four projects are sought under this strand:
- a) **1 pilot implementation of a licence registry – up to £ 200,000**
 - b) **1 pilot implementation of a name and factual authority file service – up to £200,000**
 - c) **1 scoping project to develop a technical architecture for Digital Policy Management – up to £40,000**
 - d) **1 scoping project to assess requirements for and the potential of a terminology registry – up to £40,000**
- G93. The potential areas of funding under each heading are elaborated below. However, bidders are encouraged to put forward their own ideas working within the scope provided here.

V a) Licence Registry Pilot Implementation Project

- G94. JISC invites proposals to develop a machine-readable registry of licences in common use within UK Higher Education. The pilot should include licences that are related to digital repositories, commercial content or digital resources more generally. Currently resource metadata can include identifiers relating to licences under which resources are made available. However, this is not a solution to making licence conditions available in a reliable and consistent way, as identifiers vary and often point to different forms of licence information, or do not result in any reliable information being available. It is proposed that the development of a registry would help address this and contribute to the development of identification standards for licence terms. A registry will help to record licences that are in use and give a persistent record of what is

⁶⁷ *Shared Infrastructure Services Review* http://www.jisc.ac.uk/Shared_Infrastructure_Services_Review_Sep_06

⁶⁸ *Terminology Services and Technology Report*
http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06

⁶⁹ http://www.jisc.ac.uk/funding_ietestbed.html

licensed under what conditions, allowing users to automatically associate resources with licence terms. A range of licences will need to populate the registry pilot. Examples of licences that might be included are the JISC Model Licence and Creative Commons licences.

- G95. This project will need to draw on available literature and undertake a requirements gathering stage prior to implementing a pilot registry. The use of the pilot registry will need to be tested and future use and issues of embedding and sustainability addressed. The project should be delivered in two main phases, the first being a scoping phase to specify the pilot registry, the second being the implementation of the registry pilot. The project will need to seek input from JISC and its representatives in terms of how the pilot project progresses and the shape it should take as the pilot is implemented.
- G96. One project will be funded up to £200,000 for a period of two years.

V b) Name and Factual Authority Pilot Implementation Project

- G97. The JISC Shared Infrastructure Services report and the Terminology Review both highlight the importance of name authority in disambiguation and resource discovery. The following extract from the Shared Infrastructure Services Review sets out the problem to be addressed in relation to repositories:

Name authority is a particularly significant problem for repositories. It appears that many people depositing materials in an institutional repository will not be represented in library name authority files, because they have not produced books or other materials which have been catalogued. In addition, repositories rely on a large proportion of self-archiving, where authors are tasked with inputting their own information and barriers to data input must be kept low. It is unlikely that many repositories offer a facility for 'picking' names from an authoritative source. There are also issues of multiple authorship, often with authors at different institutions...

- G98. JISC therefore wishes to investigate the potential for the development of a Name Authority Service and factual authority for digital repositories, to support cataloguing, metadata creation and resource discovery in the repository environment. This should include author name, department, institution, and deal with the issues of provenance. This project will scope and pilot a solution exploring benefits and issues. There is potential here for cooperation with other organisations within the UK such as the British Library⁷⁰ and The National Archives.⁷¹ This project will need to engage with current international developments such as OCLC Research LC Name Authority Service,⁷² and relevant work within DareNet.⁷³ It is proposed that any project aims to offer a service for institutional and departmental names where this is found to be beneficial; these are requirements that are not part of the other initiatives outlined above.
- G99. This project will need to draw on available literature and undertake a requirements gathering stage prior to implementing a pilot solution. The use of the pilot will need to be tested and future use and issues of embedding and sustainability addressed. The project should be delivered in two main phases, the first being a scoping phase to specify the pilot solution, the second being the implementation of the pilot. The project

⁷⁰ <http://www.bl.uk/>

⁷¹ <http://www.nationalarchives.gov.uk/>

⁷² <http://www.oclc.org/research/researchworks/authority/default.htm>

⁷³ <http://www.darenet.nl/en/page/language.view/search.page>

will need to seek input from JISC and its representatives in terms of how the pilot project progresses and the shape it should take as it is implemented.

G100. One project will be funded up to £200,000 for a period of two years.

V c) Digital Policy Management Scoping Study

G101. Digital policy management has been identified as a key factor in enabling users to access a wide range of resources. There are barriers to this including the uncertainty of who owns the rights and what they are and the complexity of different rights models. JISC is partly addressing this by the commissioning of the Licence Registry set out above. However, building on the recommendations in the Shared Infrastructure Services Review, further work is required to fully address this area. A scoping study that will result in a technical architecture for digital policy management and that builds on the considerable work already undertaken is therefore invited. The study will have two main phases. These are a gap analysis of the area of the digital rights and permissions environment and a scoping of a technical architecture to support Digital Policy Management.

G102. The object of the gap analysis is to investigate and identify areas in which JISC itself needs to act as a priority and those where there are appropriate developments elsewhere or where other stakeholders might take the lead. The purpose of the resulting architecture is to develop a technology framework that will be able to support the range of digital policies that need to be in place to support the diverse exchange, creation and curation of resources.

G103. The work will have to liaise with the proposed pilot licence registry as far as possible. The project should build on the findings of the Shared Infrastructure Services Review, however bidders should not be constrained by this and should develop a methodology that will offer a robust assessment that can then be taken forward in the development of digital repository networks and the broader Information Environment.

G104. One project will be funded up to £40,000. The project should take place over a six-month period.

V d) Terminology Registry Scoping Study

G105. There is little doubt that terminologies are important in assisting discovery to delivery. The terminology review also highlighted the problems in this area, for example that terminologies span many different subject areas, vocabularies and communities. Terms can have several different meanings in different subject areas, leading to ambiguity and confusion.

G106. JISC therefore wishes to commission a scoping study on the potential delivery of a Terminology Registry. A registry would aim to provide a central resource of all existing vocabularies in common use within higher education. Such a registry should look to subject headings, thesauri, classification schemes, offering 'look-up' functionality and providing descriptions of each terminology and information on context and usage in domain or discipline. Providing machine-to-machine access to information about terminologies and terminology services would encourage exploitation of existing vocabularies and enable innovative interfacing with applications from 'other domains'. Policies would need to be established covering the use of the registry and the associated terminologies. The scoping study would need to establish the requirements for such a registry, examining use cases and developing use scenarios, taking into account the benefits and potential return on investment, intellectual property rights and business models and issues of persistence and quality.

G107. Various other initiatives might be relevant, for example the eScience GRIMOIRES⁷⁴ and the Becta Vocabulary Tool⁷⁵ as well as relevant work in the National Science Digital Library (NSDL).⁷⁶ As proposed in the supporting terminology review, the JISC Information Environment Service Registry (IESR)⁷⁷ may be a relevant registry to host such a service. The study should also consider relationships to the JISC Information Environment Metadata Registry (IEMSR).⁷⁸

G108. One project will be funded up to £40,000. The project should take place over a six-month period.

Further Information

G109. All enquiries regarding this section of the call should be sent to Phil Vaughan/Helen Hockx-Yu (tel: 07810 814476 / 020 7484 1803; email: rpteam@jisc.ac.uk).

G110. Any enquiries regarding the proposal submission process should be sent to Nike Holmes (tel: 0117 931 7427; email: n.holmes@jisc.ac.uk).

**JISC Executive
September 2006**

⁷⁴ www.ecs.soton.ac.uk/research/projects/grimoires

⁷⁵ www.becta.org.uk/vocab/faq.cfm

⁷⁶ www.becta.org.uk/vocab/faq.cfm

⁷⁷ <http://iesr.ac.uk/>

⁷⁸ www.ukoln.ac.uk/projects/iemsr

Repositories and Preservation Proposal Cover Sheet

Cover Sheet for Proposals (All sections must be completed)		JISC Capital Programme
Name of Capital Programme: Repositories and Preservation Programme		
Bid for Call Area : (Please tick ONE BOX ONLY, as appropriate)		
Tools and Innovation (Strand B)		
<input type="checkbox"/>	Call Area I – Tools and Innovation Projects	Please specify area of proposed project eg <i>'metadata generation and validation'</i>
Discovery to Delivery (Strand C)		
	Call Area II – Discovery to Delivery Projects	<input type="checkbox"/> a) Version identification framework <input type="checkbox"/> b) Persistent identifier interoperability demonstrator <input type="checkbox"/> c) Federated access management and repositories <input type="checkbox"/> d) Semantic interoperability demonstrator
Repository Start-Up and Enhancement (Strand D)		
	Call Area III – Repository Start-Up and Enhancement Projects	<input type="checkbox"/> a) Repository start-up projects <input type="checkbox"/> b) Repository enhancement projects
Digital Preservation and Records Management (Strand H)		
	Call Area IV – Digital Preservation and Records Management Projects	<input type="checkbox"/> a) Digital preservation across the lifecycle <input type="checkbox"/> b) Models and implementation of preservation services <input type="checkbox"/> c) Preservation tools development
Shared Infrastructure Services (Strand I)		
	Call Area V – Shared Infrastructure Services Projects	<input type="checkbox"/> a) Pilot implementation of licence registry <input type="checkbox"/> b) Pilot national name and factual authority service <input type="checkbox"/> c) Scoping an architecture to support digital policy management <input type="checkbox"/> d) Scoping a terminology registry
Name of Lead Institution:		
Name of Proposed Project:		
Name(s) of Project Partner(s):		

Full Contact Details for Primary Contact:		
Name:		
Position:		
Email:		
Address:		
Tel:		
Fax:		
Length of Project:		
Project Start Date:	Project End Date:	
Total Funding Requested from JISC:		
Funding Broken Down over Financial Years (April – March):		
Apr06 – Mar07	Apr07 – Mar08	Apr08 – Mar09
Total Institutional Contributions:		
Percentage Contributions over the Life of the Project:	JISC	PARTNERS
Outline Project Description		
I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)		
	YES	NO
I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)		
	YES	NO