

Project Acronym: Resource Discovery Infokit
Version: 1.0
Contact: Gordon Dunsire
Date: 03/08/2007

Cover sheet for proposals JISC Invitation to Tender:
Project to create Resource Discovery Infokit

Name of lead institution/organisation	University of Strathclyde
List consultants/partners, if any	None
Full contact details for primary contact	
Name:	Gordon Dunsire
Position:	Depute Director (Infokit Project Director/Manager)
Email:	g.dunsire@strath.ac.uk
Address:	Centre for Digital Library Research Department of Computer and Information Sciences University of Strathclyde Livingstone Tower 26 Richmond Street Glasgow G1 1XH
Tel:	0141 548 2379
Fax:	0141 548 4523
Length of project:	5 months
Total cost to JISC over its life:	£29,979
Proposed project start and end dates:	2 July to 30 November 2007
Outline proposal description	
This project will:	
<ul style="list-style-type: none">• create a resource discovery information kit (or Infokit) on the JISC website;• offer views of the content of the Infokit according to key questions and themes;• explain clearly how to implement resource discovery solutions;• provide pointers to useful resources and projects;• develop a plan for the future maintenance of the Infokit.	



Resource Discovery Infokit: Project Plan

Overview of Project

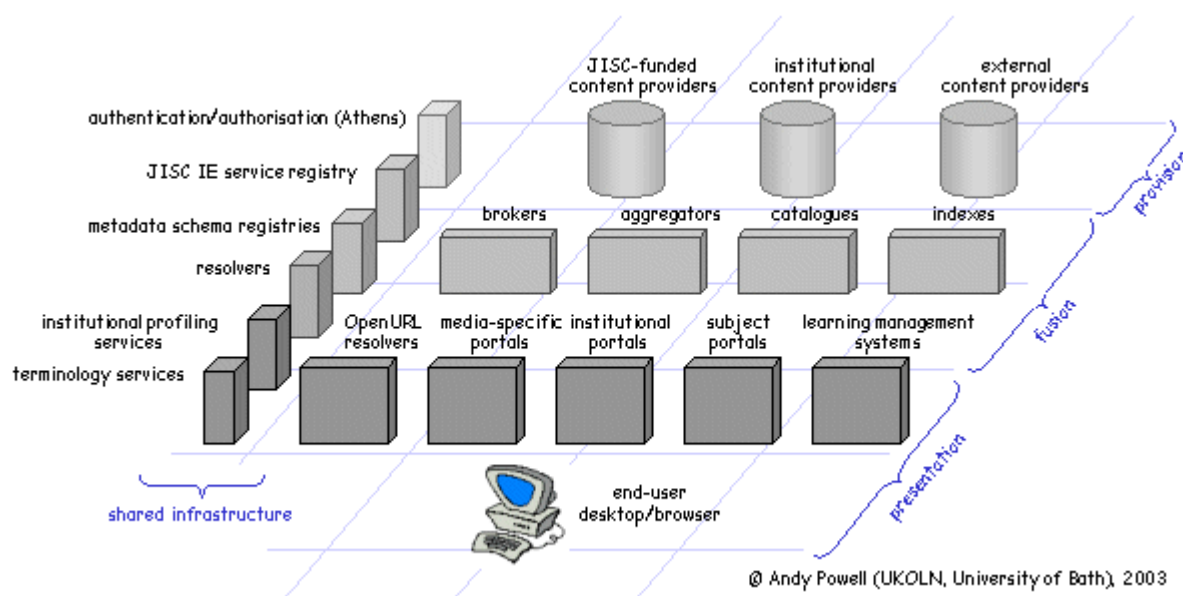
1. Background

1.1. Centre for Digital Library Research

The University of Strathclyde's Centre for Digital Library Research was established in 2000, with two key aims, to combine theory with practice in innovative ways and to be a centre of research excellence in the area of distributed digital information and learning environments. The CDLR has a clear guiding philosophy of combining research and development, such that its research is well grounded and readily applicable, and its development work is based on sound principles, demonstrates best practice and wherever possible consolidates and extends skills, experience and knowledge. This philosophy is reflected in the Centre's position within the University, with connections to both the Department of Computer and Information Sciences and the Information Resources Directorate.

1.2. The JISC and the JISC Information Environment

The Joint Information Systems Committee (JISC) is a committee of the UK further and higher education funding bodies that, through the funding of services and the management of development programmes, aims to help and facilitate institutions in their use of the Internet and other Information and Communications Technology applications. One of the JISC's key objectives is to build an online information environment that provides secure and convenient access to a comprehensive collection of scholarly and educational material. This is based on the Information Environment Technical Architecture (shown below) which specifies a set of standards and protocols that support resource discovery as part of learning, teaching, and research activities. The JISC is developing components within this architecture to test out and promote the use of these standards and to help provide appropriate services to the community.



1.3. Resource Discovery Infokit

The project to develop a Resource Discovery Infokit will consolidate and synthesise resource discovery information from JISC Development Programmes and JISC Services within the Information Environment. Relevant work in the area of resource discovery includes but is not limited to:

- the Portals Programme;
- the Presentation Programme;
- the Virtual Research Environments (VRE) Programme;
- the e-Learning Frameworks and Tools Programme;
- the Distributed National Electronic Resource (DNER 5/99) Learning and Teaching Programme;
- the Shared Infrastructure Services Programme;
- the User and Innovation Programme;
- the Semantic Grid and Autonomic Computing Programme.

The project will synthesise the resource discovery outputs from relevant JISC programmes and services and so provide a single point of quick and easy access to information, advice, services, tools, case studies, software and related standards and technology to support resource discovery development. Rather than being a comprehensive programme synthesis, the Infokit is intended to provide answers to the key questions which users might wish to ask and should be selective in terms of the information to which it points.

When creating any kind of digital resource, it is important to consider its sustainability, longevity, accessibility, interoperability and reusability, thereby ensuring not only that it is fit for its immediate purpose, but also that it can readily be maintained and that it can support the widest possible range of uses beyond the immediate context of its creation. For the creation of the Infokit, the CDLR will adopt an approach based on faceted information retrieval via an extensible taxonomic framework, which will not only enable users to discover and access resources quickly and easily, but also enable the JISC to maintain and sustain the Infokit in the longer term, and indeed to extend and add value to it in the future.

1.4. Terms of reference

- To scope the requirements of a resource discovery Infokit. This will require comprehensive desk research including regular liaison with the JISC Executive and UKOLN.
- An early output will be defining the key questions and themes that the Infokit will need to address and a site map based on this analysis. As part of this addressing institutional and enterprise requirements, national requirements such as the JISC Information Environment should be accounted for.
- Consolidation and synthesising of the relevant JISC projects and services that deal with resource discovery with the aim of identifying the most relevant resources and information to address the key questions. As well as focusing on JISC programme and service outputs the information should, within reason, take account of other relevant developments and dependencies e.g. emerging standards and technologies such as semantic web.
- To draft supporting information and to bring together programme outcomes and outputs according to the Infokit requirements.
- To create the web page(s) in conjunction with the JISC Communications and Outreach Team.
- As part of this work, supporting FAQ's and briefing papers are also anticipated.
- To design optimum display for the summarised information (pointing to various relevant developments).
- To consider maintenance – it might be that the Infokit should just reflect a moment in time, or perhaps there are ways in which it could be easily maintained.

2. Aims and Objectives

The main aim of the Infokit is to make it easier for institutions and resource discovery providers to find information on resource discovery solutions.

The main audiences for the Infokit are computing and library staff, learning technologists and others implementing resource discovery services in UK HE; and JISC projects.

The primary objectives are to:

- create a resource discovery information kit (or Infokit) on the JISC website;
- offer views of the content of the Infokit according to key questions and themes;
- explain clearly how to implement resource discovery solutions;
- provide pointers to useful resources and projects;
- develop a plan for the future maintenance of the Infokit.

The primary focus will be on resource discovery solutions that support the JISC Information Environment and that have been developed via JISC Development Programmes and Services.

3. Overall Approach

The methodology chosen for the Infokit is based on the development of a faceted taxonomic framework. This approach to organising and displaying information within web pages allows the user to discover and access resources in a number of ways – navigational browsing via vocabularies and taxonomies, contextual browsing via FAQs, narrative browsing via briefing papers, and of course free text searching.

Some scoping work is now underway to confirm the suitability of the approach outlined above and to begin to identify the vocabularies or taxonomies required and the key questions and themes to be addressed. An initial listing of resources is being compiled in Access, following which we will have a clearer idea of scale. It will be possible to apply either a hard or a soft definition of resource discovery – the latter will produce a very large number of resources, some of which are invaluable but the totality of which might serve to dilute and devalue the Infokit as a whole – so it will be necessary to set some definitional boundaries and determine priorities for inclusion.

As part of the scoping work, the project team are also investigating whether the Infokit should be integrated fully or partially within the JISC's content management system, or whether it should be completely separate; a range of technical, practical and economic factors are being considered and a decision on the preferred technical solution will be taken in consultation with the JISC.

Following on from the scoping work, the project will develop a series of vocabularies for the various facets by which users might wish to navigate the content, as follows:

- institutions and organisations – identify and select from existing (ideally maintained) listings – hierarchical eg CDLR, University of Strathclyde;
- JISC programmes and projects – obtain from the JISC/Eduserv and retain ID numbers to ensure integration – again hierarchical and may also have other internal structure;
- taxonomy to reflect the architectural and functional attributes of the JISC Information Environment – will require some research effort and external evaluation (see Evaluation Plan) – will be inclusive of all components of a resource discovery solution – will have internal relationships and dependencies;
- some relationships and dependencies may best be handled by additional vocabularies eg institution type, user type, technologies and standards used, etc.

Within each vocabulary or taxonomy, the degree of detail will in part be determined by the number of resources to be described, although the approach is extensible in that further degrees of detail or granularity can be applied as and when required. Notes on the completeness and volatility of each vocabulary will be provided.

The findings of the initial scoping work, together with the completed vocabularies and taxonomies, will be presented in an interim report in mid-September. It may be useful to update the project plan at this time.

The second phase of the project will develop the content of the Infokit and implement the preferred technical solution. Content development will have two main strands – the preparation of FAQs and briefing papers, and the identification and description of the resource discovery tools and other resources which are to be included in the Infokit. The vocabulary of JISC programmes and projects and the taxonomy of attributes of the JISC IE are broadly equivalent to, but more formally expressed than, the 'key questions and themes' mentioned in the terms of reference. Thus, the topics for FAQs will be driven by the taxonomic analysis whilst the topics for the briefing papers will be abstracted from JISC programme specifications. When preparing FAQs and briefing papers, the policy will be to reuse JISC text wherever possible and then to supplement with original text as necessary.

The initial listing of resources for inclusion will be revised and terms from the vocabularies and taxonomies will be applied, enabling the resource discovery and access mechanisms described above to be implemented. At this stage, it is not possible to plan in detail the technical implementation; however, it is assumed that a database will be required to store vocabularies, resource descriptions and metadata, and to either publish the web pages locally or export the data for publication elsewhere.

In the final stages of the project, a plan for the ongoing maintenance of the Infokit will be drawn up in consultation with the JISC, and with Eduserv as appropriate (see Exit and Sustainability Plan).

To carry out the project successfully, a series of meetings with the JISC Programme Manager, members of the JISC Communications and Outreach Team, and possibly relevant staff at Eduserv, will be required (either face-to-face or by phone). At this stage, it is clear that one or two meetings will be required to reach a decision about the preferred technical solution, and to finalise the vocabularies. The requirement and schedule for meetings beyond that will be determined at the meeting to discuss the interim report.

One further task will be carried out alongside the programme of work outlined above. It is not critical to the successful completion of the project, but has the potential to add considerable value to the work from both short and long term perspectives. In addition to applying terms from the vocabularies and taxonomies within the database, we will also consider the feasibility of using Simple Knowledge Organisation Systems (SKOS: <http://www.w3.org/2004/02/skos/>). This has the capacity to include cross-references, scope notes and so on, which would make the Infokit more readily accessible to a cross-community user base. Furthermore, the SKOS editor enables visualisation of taxonomies which could prove useful for evaluation and dissemination purposes. Our experiences will be presented in the final report and may warrant further dissemination in the form of a journal article or conference paper (see Dissemination Plan)

4. Project Outputs

The deliverables formally required from the project are:

- a project plan with full details of the methodology, workplan, etc to be delivered within a month of the project start date;
- an interim report detailing the key questions and themes to be addressed, identifying the most relevant resources for inclusion, proposing an approach for displaying this information within a series of web pages and recommending a technical solution that can achieve this, to be delivered by 14 September 2007;
- the design and content of the web pages, including relevant FAQs and briefing papers, to be delivered such that web pages will be live by the project end date;
- a final report, including a plan for the future maintenance of the Infokit, to be delivered by the project end date.

In addition, at least some of the vocabularies or taxonomies developed for the Infokit could have a wider application and will therefore be useful outputs in their own right.

Less tangible outputs include the knowledge and experience gained through investigating the use of SKOS to evaluate, apply and share vocabularies and taxonomies. Similarly, the use of the taxonomic framework to organise and display the content of the Infokit, and thus to allow the user to discover and access resources in a number of ways, will be of value beyond the project.

5. Project Outcomes

Outcomes of the project are envisaged in three key areas:

- improved discovery of and access to resources relating to resource discovery, resulting in more efficient and/or effective implementation of resource discovery solutions within the JISC community;
- a improved understanding, in the form of a taxonomy or ontology, of the JISC Information Environment's technical and functional attributes, together with an improved understanding of how we might use SKOS to share this understanding with other communities, potentially resulting in greater cohesion within and across similar communities worldwide;
- a high profile demonstration of the use of taxonomic frameworks to structure collections of information resources on the web, potentially resulting in improved practice elsewhere.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
JISC	Synthesis of outputs	High
JISC projects and services	Ready access to resource discovery info Vocabularies	Medium Medium
HE/FE institutions	Ready access to resource discovery info Vocabularies Demo of taxonomic framework approach	Medium Low Medium
Resource discovery providers	Ready access to resource discovery info Vocabularies	High Medium

7. Risk Analysis

The primary risk to this project is in the area of staffing. If any of the members of staff named in this tender are unable to complete their contribution within the given timescale, the CDLR has a larger pool of suitably qualified staff on which to draw, thereby minimising any risk of delay due to staffing issues. One such change of staff has already occurred.

The lack of a clearly defined technical solution at the outset of the project poses some organisational and technical risks but a number of technical solutions are possible, including an in-house solution which is known to be achievable within budget and timescale. The methodology permits work on the content to progress without waiting for a decision on the preferred technical solution, thereby avoiding unnecessary delay.

A further risk arises in the area of contributions from other organisations, primarily the JISC's Communications and Outreach Team, Eduserv and UKOLN. These organisations also have a pool of suitably qualified and experienced staff on which to draw, so any risk is minimal.

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing	2	2	4	Adopt team approach
Organisational - technical solution may not be known until well into the project	4	2	8	At least one good solution known to exist Adopt flexible methodology
Technical - availability of skills to implement preferred technical solution	1	3	3	Additional technical skills available within CDLR
External suppliers - Eduserv unable or unwilling to deliver	1	4	4	Draw up contract Alternative solutions possible

8. Standards

It is not possible to provide a comprehensive list of standards to be used within the project until the technical solution has been determined, but it is anticipated that JISC guidelines will be followed and appropriate non-proprietary standards used throughout.

9. Technical Development

As stated in Section 3, we will deploy database software to underpin the technical development of the project, as this offers maximum flexibility ahead of any decision about the preferred technical solution. Resource descriptions, metadata and vocabulary terms will be recorded in an MS Access database, from which we can generate static web pages for local publication or export data and/or metadata in a variety of formats for publication or reuse elsewhere.

The choice of MS Access has been determined by the availability of suitable skills and experience at all levels within the project team, by the existence in house of VB scripts for generating static web pages or exporting in various file formats, and because it does not preclude the possibility of migrating the database to another platform mid-project.

10. Intellectual Property Rights

All outputs specified as deliverables of the project will be made freely available for widespread dissemination by the JISC in perpetuity. The rights to any additional outputs, such as research papers, will be retained by the CDLR and/or assigned to publishers as appropriate.

Project Resources

11. Project Partners

The project has no formal partners or subcontractors at present. Depending on the technical solution selected for the Infokit, a formal contract with Eduserv may be required for the implementation stage of the project; this would have to be in place no later than the end of September.

12. Project Management

The project will be directed and managed by Gordon Dunsire, the CDLR's Depute Director, who will contribute approximately one day per week, at least 15 days in total. Emma McCulloch will contribute a similar amount of time; her role will be focused in the area of taxonomies and ontologies, where she has considerable expertise and experience. The day-to-day work of the project will be carried out by one or more junior researchers from the CDLR, with support, in the form of technical assistance or training, from more experienced CDLR staff as and when required.

The primary point of contact will be: Gordon Dunsire
(g.dunsire@strath.ac.uk; 0141 548 2379).

Address for correspondence: Centre for Digital Library Research
Department of Computer and Information Sciences
University of Strathclyde
Livingstone Tower
26 Richmond Street
Glasgow G1 1XH

13. Programme Support

There are two main areas where support from the programme manager would be invaluable: liaison with the JISC Communications and Outreach Team, and with Eduserv and UKOLN, as necessary;

and contributions, in the form of feedback and possibly liaison with other stakeholders, during the evaluation phase of the project.

14. Budget

A revised budget for the project is attached at Appendix A.

Detailed Project Planning

15. Workpackages

A detailed programme of work is given in the form of a series of workpackages at Appendix B. The workpackage relating to the technical implementation in the second half of the project is necessarily sparse at this stage; they will be updated following the delivery of the interim report in mid-September.

16. Evaluation Plan

A formative evaluation of the taxonomies will be carried out in conjunction with the main stakeholders in the project; the findings and outcomes of this evaluation will form a significant part of the interim report. The primary stakeholders are of course the JISC and UKOLN; a full list of the stakeholders who will be asked to evaluate the taxonomies will be determined during the project scoping process.

The project team will also carry out an informal peer review of the FAQs and briefing papers, drawing on the expertise of colleagues in CDLR, at the JISC and elsewhere (the CDLR has an extensive network of contacts who might be called on to review documents on an informal basis).

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Early Sept	Taxonomies	Fitness for purpose	Peer review Implementation	Acceptance by stakeholders Successful implementation
Early Nov	FAQs/ briefing papers	Fitness for purpose	Peer review	Acceptance by stakeholders

17. Quality Plan

In addition to the evaluation activities described above, there are two areas which will be subject to quality assurance procedures during the project, namely database content and web page creation.

The creation of the database of resources will be conducted in line with normal quality assured content and metadata creation processes, via the implementation of a simple workflow, to include the application of controlled vocabularies, checks for internal consistency, etc. It is anticipated that completeness may be an issue within the timescale and resourcing available; a policy decision on how to resolve this and any similar issues will be taken as part of the scoping of the project.

The development of web pages necessarily requires compliance with various standards, and with accessibility legislation; this compliance will be verified in conjunction with the JISC Communications and Outreach Team towards the end of the project and ahead of the publication of web pages.

Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities
End Nov	Various database quality criteria	Implementation of simple workflow	Effective resource discovery and access	Policy decisions may be required
End Nov	Web standards	Adoption of best practice	Compliant web pages	External partners may be involved

18. Dissemination Plan

The web pages to be created as a deliverable of the project will be published via the JISC website on or before the project end date and will be freely available to all potential users both within and beyond the JISC community.

The project team will write one or more short articles or announcements to raise awareness of the project in its early stages and promote the resulting Infokit towards the end of the project, to be disseminated via publications such as Ariadne and on appropriate mailing lists.

The project team would also expect to produce one or more research papers on the methodology and outcomes of the project.

Timing	Dissemination Activity	Audience	Purpose
Jul/Aug	Promotional article	JISC community	Raise awareness of project
Nov/Dec	Promotional article	JISC community	Promote outputs of project
Nov/Dec	Research paper	Practitioner/ research communities	Report and evaluate use of taxonomic framework
Nov/Dec	Research paper	Practitioner/ research communities	Report and evaluate use of SKOS

19. Exit and Sustainability Plans

There are three main outputs which should be considered as part of the project's exit strategy: the Infokit itself, comprising both the published web pages and the underlying database, the associated vocabularies, and the learning from the project.

Project Outputs	Action for Take-up & Embedding	Action for Exit
Database and web pages	To be determined as part of the Sustainability Plan, which is a deliverable of the project	
Vocabularies		
Learning and experience	Documentation and dissemination	See Dissemination Plan

It is not possible to provide detailed information on sustainability at this stage, as this forms part of the work of the study. However, the approach described in this project plan has been selected precisely because the Infokit produced by this method will be extensible and, assuming some kind of commitment from the JISC, could readily be updated on an ongoing or periodic basis. A more detailed plan with resourcing options will be included in the final report.

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Database and web pages	Ongoing value to target audience	With JISC - to be determined	Ongoing maintenance required
Vocabularies	Value extends beyond project	With JISC - to be determined SKOS	

Project Acronym: Resource Discovery Infokit
Version: 1.0
Contact: Gordon Dunsire
Date: 03/08/2007

Appendixes

Appendix A. Project Budget

(Removed for data protection and privacy)

Appendix B. Workpackages

WORKPACKAGES	July 07	Aug 07	Sept 07	Oct 07	Nov 07
1: Project administration					
2: Project scoping					
3: Content development					
4: Technical development					
5: Dissemination					

Key milestones and deliverables					
Project planning					
Promotional article					
Initial database of resources					
Assessment of technical solutions					
Draft vocabularies and taxonomy					
Evaluation of taxonomy					
Selection of topics for FAQs and briefing papers					
Interim report					
Draft FAQs and briefing papers					
Evaluation of FAQs and briefing papers					
Full database of resources					
Quality assurance of database					
Preparation of web pages					
Quality assurance and publication of web pages					
Investigation into using SKOS					
Preparation of sustainability plan					
Final report					
Journal article(s)					

Project Acronym: Resource Discovery Infokit
 Version: 1.0
 Contact: Gordon Dunsire
 Date: 03/08/2007

Project start date: **2 July 2007**

Project completion date: **30 November 2007**

Duration: **5 months**

GD = Gordon Dunsire

GM/EM = George Macgregor / Emma McCulloch

RA = Junior research assistant

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables in bold)	Milestone	Responsibility
WORKPACKAGE 1: Project administration					
<u>Objective:</u> Manage project					
1. Project plan	2 July	3 Aug	Project Plan	Project Plan accepted	GM/GD
2. Interim Report	27 Aug	14 Sept	Interim Report	Interim Report accepted	GD
3. Final Report	5 Nov	30 Nov	Final Report	Final Report accepted	GD
4. Internal management meetings	2 July	30 Nov	Effective project management		GD
5. External project meetings	2 July	30 Nov	Effective decision making		GD
WORKPACKAGE 2: Project scoping					
<u>Objective:</u> Scope project					
6. Set definitional boundaries and priorities for inclusion of resources	5 Aug	14 Sept	Section in Interim Report	Interim Report accepted	GD
7. Identify vocabularies required	16 July	14 Sept	Section in Interim Report	Interim Report accepted	GD/EM
8. Identify key questions and themes	16 July	14 Sept	Section in Interim Report	Interim Report accepted	GD/EM
9. Assess technical solutions	16 July	14 Sept	Section in Interim Report	Interim Report accepted	GD
10. Assess feasibility of using SKOS	16 July	14 Sept	Section in Interim Report	Interim Report accepted	GD/EM

WORKPACKAGE 3: Content development					
<i>Objective: Develop content of Infokit</i>					
11. Create vocabularies	16 July	14 Sept	Vocabularies	Vocabularies implemented	GD/EM/RA
12. Create and evaluate taxonomy of IE	16 July	14 Sept	Taxonomy	Taxonomy implemented	GD/EM
13. Populate initial database	16 July	14 Sept	Initial database	Policy decisions taken	RA
14. Populate full database	17 Sept	30 Nov	Database	Web pages published	RA
15. Create FAQs and briefing papers	17 Sept	30 Nov	FAQs and briefing papers	Web pages published	RA
WORKPACKAGE 4: Technical development					
<i>Objective: Produce Infokit website</i>					
16. Set up initial database	16 July	14 Sept	Initial database	Supports content creation	GD/RA
17. Set up full database	17 Sept	30 Nov	Database	Supports web page creation	GD/RA
18.					
19.					
20.					
WORKPACKAGE 5: Dissemination					
<i>Objective: Disseminate project outputs</i>					
21. Awareness raising article	5 Aug	14 Sept	Article	Article accepted	GD/EM
22. Article reporting project outcomes	5 Nov	30 Nov	Article	Article accepted	GD/EM/RA
23. Research paper on methodology	3 Dec	25 Jan	Research paper	Paper accepted	GD/EM
24. Research paper on use of SKOS	3 Dec	25 Jan	Research paper	Paper accepted	GD/EM