JISC

Project Information			
Project Acronym	COMPASS		
Project Title	COastal Marine Perception Application Scientific Scholarship		
Start Date	3 December 2007	**End Date**	31 March 2009
Lead Institution	EDINA		
Project Director	Dr David Medyckyj-Scott, EDINA		
Project Manager & contact details	Dr Kristin Stock Centre for Geospatial Science Sir Clive Granger Building University Park University of Nottingham, NG7 2RD email: kristin.stock@nottingham.ac.uk telephone: +44 (0) 115 84 68413		
Partner Institutions	University of Edinburgh University of Muenster National University of Ireland, Galway Social Change Online		
Project Web URL	compass.edina.ac.uk		
Programme Name (and number)	*JISC Circular 01/07 e-Research: e-Infrastructure Call II* *Knowledge Organisation and Semantic Services*		
Programme Manager	Alison Turner		

Document Name			
Document Title	*JISC Project Plan*		
Reporting Period			
Author(s) & project role	Kristin Stock, Project Manager		
Date	21 February 2008	**Filename**	Compass Project Plan v0.1.doc
URL			
Access	✓ Project and JISC internal		☐ General dissemination

Document History			
Version	**Date**	**Author**	**Comments**
0.1	21 February 2008	Kristin Stock	First draft.
0.2	27 February 2008	Kristin Stock	Incorporation of comments from team members.
1.0	29 February 2008	Kristin Stock	Released version.

JISC

JISC Project Plan

Overview of Project

1. Background

The research and decision-making process in any discipline is supported by a vast quantity and diversity of scientific resources, including journal articles; scientific models; scientific theories; data sets and web services that implement scientific models or provide other functionality. Improved discovery and access to these scientific resources has the potential to make the process of using and developing scientific knowledge more effective and efficient.

Current scientific research or decision making that relies on scientific resources requires an extensive search for relevant resources. Published journal papers may be discovered using web searches on the basis of words that appear in the title or metadata, but this approach is limited by the need to select the appropriate words, and does not identify articles that may be of interest because they use a similar approach, methodology or technique but are in a different discipline, or that are likely to be helpful despite not sharing the same keywords. This approach also does not identify resources that do not appear in web pages or are not indexed by search engines, including resources that are stored in databases or digital libraries, or are registered as web services in data infrastructures such as, in the case of geospatial data, Spatial Data Infrastructures (SDIs).

The COMPASS project will create a prototype knowledge infrastructure that includes semantic information about scientific resources. The inclusion of such semantic information in the form of ontologies will enable more intelligent and sophisticated discovery of scientific resources and allow automatic access to and execution of resources where appropriate and possible. Specifically, such an approach will:

- reduce the need for scientists to use a particular term in searching for resources;
- allow scientists to discover related resources that may also be relevant;
- allow scientists to immediately execute web services and access data sets that relate to their research;
- aid scientists in identifying other scientists working in the same area, or scientists working in different areas but using similar approaches, models or theories and
- enable scientists to more easily visualise knowledge that relates to their work.

The storage of semantic information in such an infrastructure can be applied in a number of ways for different purposes, and this project will focus on providing a framework that may be later expanded to include more advanced functionality.

MarineXML[1] and Marine Overlays on Topography for Annex II Valuation and Exploitation (MOTIIVE)[2] provide much of the basis for this work. These projects developed and structured knowledge in the marine domain and designed basic infrastructures for the storage of simple semantic information in an interoperable context. Established connections with the coastal marine community (for example, the British Oceanographic Data Centre (BODC), the National Oceanography Centre, the International Hydrographic Organisation (IHO) and the Marine Data and Information Partnership (MDIP)[3]), will provide much of the context, domain expertise and raw material feeding into the project. A number of existing initiatives to model knowledge in the marine domain will also be used as a starting point,

[1] http://www.iode.org/marinexml/
[2] http://www.motiive.net/
[3] http://www.oceannet.org/mdip/

including the Natural Environment Research Council Data Grid vocabulary and ongoing work by the Australian Antarctic Data Centre.

In the case of geospatial interoperability and semantics, the outputs from the Semantic Web services Interoperability for Geospatial decision making (SWING)[4] project, in combination with MOTIIVE, will provide added means to ensure that COMPASS builds on best practice from the open geospatial interoperability standards community (that is, the Open Geospatial Consortium (OGC)[5] and ISO/TC211[6]) and the latest research in the semantics field. Semantic interoperability in the digital libraries area is less mature than the geospatial data realm, but existing work by Digital Enterprise Research Institute (DERI) (National University of Ireland, Galway) will be expanded in this project.

2. Aims and Objectives

The aim of the project is to build an ontologically supported knowledge infrastructure within a specific domain (coastal marine environments) for a number of Communities of Practices (CoP)[7] in order to investigate the development and use of semantic tools to assist in discovery, access and use of scientific resources such as data; journal articles; scientific models and web services.

The objectives of the project are:

- to create a domain ontology that expresses expert knowledge taking existing ontologies and feature type catalogues into account for a selected subset of the marine domain;
- to explore and compare the use of ontologies with the use of informal tagging approaches to assist users in discovery of scientific resources;
- to develop an interoperable semantic registry to store knowledge infrastructure content and metadata (including ontologies) using international standards for storage and implementing existing international standard interfaces;
- to develop or extend international standards (for example, ebRIM[8]) where needed to accommodate the semantic content of the knowledge infrastructure;
- to develop methods and tools for discovery of scientific resources from the knowledge infrastructure;
- to implement methods and tools to allow the knowledge infrastructure to interoperate with digital libraries;
- to implement and populate the knowledge infrastructure and demonstrate its operation using a selected subset of the marine domain;
- to make recommendations to JISC concerning the sustainability of the knowledge infrastructures approach and how the methodology may be extended to other areas within the UK research community;
- to evaluate the knowledge infrastructure within the CoPs and
- to work within the context of relevant standards organisations to gain maximum synergy and sustainable outputs.

3. Overall Approach

3.1. Strategy and Methodology

The project brings together a group of people with established expertise in the following areas:

---

[4] http://www.swing-project.org/
[5] http://www.opengeospatial.org/
[6] http://www.isotc211.org/
[7] http://home.att.net/~discon/KM/CoPOverview.pdf
[8] OASIS ebRIM, *ebXML Registry Information Model Version 3.0*, OASIS Standard (May 2005), available [online]: <http://www.oasis-open.org/committees/download.php/13591/docs.oasis-open.orgregrepv3.0specsregrep-rim-3.0-os.pdf>

- *semantics and ontologies* (Dr Femke Reitsma, (School of Geosciences, University of Edinburgh), the University of Muenster Semantic Interoperability Lab (MUSIL), Dr Kristin Stock, Social Change Online UK (SCO));
- *ontological/semantic registries* (SCO, EDINA)*, and the use of semantics for automated service invocation* (MUSIL, SCO, EDINA);
- *Semantic Service Description frameworks* (MUSIL, Digital Enterprise Research Institute (DERI) (National University of Ireland, Galway));
- *information resources, data, metadata and web and GRID services* (EDINA, DERI);
- *involvement in standards development* (EDINA, SCO, DERI).

COMPASS will combine technologies and standards in the areas of metadata, registries, ontologies and digital libraries with a focus on (but not exclusive to) the geospatial domain in order to build the knowledge infrastructure.

The knowledge infrastructure content will be stored in a semantic registry using open source software (if possible). The registry will be implemented in ebRIM with an OGC WRS[9] interface. This registry will store the information used by the infrastructure to find, manage and execute scientific resources, including the ontologies, and will be fully interoperable with other registries so that its content can be accessed by other knowledge infrastructures and SDIs. It will be implemented using existing OGC standards that will be extended to include the components necessary to store ontologies. This extension will be published as another standard to achieve the long term objective of a standardised approach to the inclusion of semantic information in registries.

Knowledge management and reasoning software will be used to interpret the registry content (particularly the semantic content) and support intelligent queries. Knowledge Smarts[10] is being considered as a package to be used for this, with the added advantage of allowing multiple data sources to be integrated and access through APIs. If Knowledge Smarts proves unsuitable, alternative open source tools will be investigated (for example, Pellet[11]).

A discovery interface will be implemented to allow users to perform queries and view and access scientific resources that relate to their query. The query may be performed by selection of any of the following search criteria:

- marine domain ontology concept or relationship (by browsing or keyword search of the ontology);
- geographical location (on a map);
- temporal location (on a timeline);
- scientific theory and/or scientific model;
- informal user tag (by browsing or keyword search in user tag clouds) or
- criteria representing various relaxations of semantic content (for example, searching not only by a specific geographic feature type, but also by related geographic feature types).

The interface will use the reasoning software to gather results, and will allow users to directly execute web services returned in response to a query, or to navigate to other related resources using hyperlinks.

The knowledge infrastructure will include four different ontologies or groups of ontologies:

- for the domain knowledge (a selected part of the marine domain focussing on instruments for measurement of parameters of the marine environment and related concepts);
- scientific knowledge, including theories, models and methods;
- scientific sources, including journal articles, data sets and web services (possibly incorporating or adapting existing web service ontology languages) and
- application ontologies to link a particular scientific source to the related scientific knowledge and domain knowledge.

---

[9] Open Geospatial Consortium. (2005). OGC Catalogue Services - ebRIM (ISO/TS 15000-3) profile of CSW.
[10] http://www.imagemattersllc.com/products/knowledgesmarts.php
[11] http://pellet.owldl.com/

(The term scientific resources used throughout this document refers to the combination of the scope of the scientific knowledge and scientific source ontologies and includes the scientific resources that may be of interest to users in discovery, browsing or interrogation).

The domain ontologies will be developed using a workshop with marine experts, in which their knowledge will be collected to create the ontology from generic concepts using ontology engineering methods developed by MUSIL for the SWING project.

The scientific knowledge ontology will be developed on the basis of previous research by project partners on the SKI ontology (to be published in the future) in developing ontologies of scientific knowledge, and will focus on models, methods and theories.

The scientific sources ontology will be developed by project partners.

An alternative method for annotating scientific resources will also be used in the project. This method will allow users to tag scientific resources in the knowledge infrastructure with terms that are meaningful to them using open source gnizr[12] software. The tags created in gnizr will be provided in the discovery interface to allow users to search based either on tag clouds (combining the tags of multiple users) or on the more formal ontology. This part of the project will evaluate the two alternatives.

An digital library interface will be created using existing standards. Software components will be developed to allow the digital library to interoperate with the knowledge infrastructure, and to allow digital library information to be returned through the discovery interface of the knowledge infrastructure. Our initial view is that digital library content will not be stored in the COMPASS semantic registry, but will be queried on-demand. However, the project will re-assess this position as it progresses.

The project will incorporate content based on the selected sub-domain of scientific instrumentation for measurement and sampling of the marine environment. These instruments measure and sample a range of parameters in the marine environment and will be identified at the workshop. The created domain ontology will model the knowledge about the instruments and the parameters they measure. A range of scientific resources will be selected that relate to research using and about those instruments. Such scientific resources will include journal articles, scientific models and theories, data sets and web services. Data sets and web services already held by project partners or easily accessible will be used where possible, but a limited number of new web services may be implemented if necessary to make relevant data available for the knowledge infrastructure.

This method was selected for a number of reasons:

- it meets the project objectives;
- interoperability is one of the important goals of the project and one of the areas in which many team members are active, and the use of the selected standards (for example, OGC and Digital Libraries standards) will enable an interoperable solution;
- ontologies were selected for the representation of semantics because they are capable of representing semantics richly, support some reasoning capability and are well supported by existing research and software tools;
- the use of domain, scientific knowledge, scientific source and application ontologies allows users to have access to a range of related resources without knowing the domain in detail or using the correct terminology. It also allows users to browse scientific knowledge using an intuitive structure and permits semantically-related resources (including web services) to be identified through inference, reducing the need to store all links to related resources;

[12] http://www.imagemattersllc.com/products/gnizr.php Gnizr is a semantic web tool that allows users to create tags, attach them to resources, and relate them to one another.
[12] http://marinemetadata.org/

- open source software is preferred for COMPASS to make the solution adoptable by other parties without expense and to ensure that we have full access to software code for customisation;
- the approach allows different methods for knowledge representation to be evaluated and compared;
- the inclusion of digital libraries data on-demand using recognised standards provides a solution that makes a large amount of information available in an extensible manner, in contrast to solutions that are based upon the idea that digital library information is stored directly within the registry.

Interoperability is one of the important goals of the project, and will be achieved by the adoption of international standards, particularly for the interfaces to content storage. Interfaces using the OGC standards and the digital library standards will be provided and where necessary, the existing standards will be extended to include ontology content. The extensions will be submitted to become part of the relevant standards ratification processes.

Evaluation is also an important issue for the project. The project intends to evaluate the use of semantic information in knowledge infrastructures; the use of ontologies to represent semantics in knowledge infrastructures and methods for the representation of knowledge. The project will achieve this by implementing a system that includes all of these components and comparing it to other work by the project team and by other researchers. The deliverables for the project include a number of evaluation and comparison reports to this end.

The continuation and extension of work in the area of SDIs is an important aspect of this project. The JISC has already committed significant resources to the establishment of an academic SDI but its uptake, as in many SDI's, will be limited by the lack of knowledge of specific spatial data domains; COMPASS represents another stage in the progression towards semantically-enabled SDIs. It is therefore important that the outcomes of the project be publicised and disseminated within not only the academic community but also the wider geographic information using community through publications and presentations.

3.2 Scope and Boundaries

The COMPASS knowledge infrastructure will be implemented using content in a selected subset of the marine domain. This selected subset will limit the scope of the ontologies developed and will be centred around a limited set of marine instruments.

The domain ontologies will include information about these instruments, the parameters they measure and any other concepts necessary for modelling knowledge about those instruments (including possibly generalisations of those instruments), but will not include other instruments or unrelated parameters and aspects of the marine domain.

The scientific knowledge ontology will focus on theories, methods and models.

The scientific source ontology will focus on journal articles, data sets and web services.

Within the scope of the selected instruments, a set of scientific resources will be chosen and will include journal articles, scientific models, theories, data sets and web services that include research and other work relating to those instruments and parameters, either directly or indirectly. COMPASS does not intend to exhaustively include all scientific resources that relate to these instruments, but to use a sample to demonstrate and evaluate its operation. This will include enough resources to allow users to:

- follow several links between resources using various paths (through scientific theories, scientific models, ontology terms);
- navigate through the ontology to look at related research (thus the ontology will be sufficiently complex to allow this);

- select from a range of web services and data sets for different scientific resources, including different types of OGC web services (for example, web feature services and web map services) and non-spatial web services, and
- view related resources in the digital library.

The project will not attempt to dynamically generate ontology content from natural language texts of other documents. Ontologies will be built manually.

The discovery tools will allow searching based on a range of ontology terms, returning related resources and will allow navigation among and between resources from the returned results (without executing a new search), but will not necessarily include advanced reasoning or inference mechanisms for the knowledge. Such advanced discovery and visualisation will be included if time permits, but is not a mandatory part of the project.

The project will not directly address social networking and trust management but will be referenced in reports where appropriate.

The project will develop simple tools for either harvesting or entry of the content as required for the project, depending on the most efficient method for populating the system with content. This component of the project is not intended to be part of the final, offered knowledge infrastructure, but will be included if time permits and if it is suitable for more general use.

3.3 Critical Success Factors

The critical success factors for the COMPASS project are interoperability, intelligence and scalability.

Interoperability refers to the sharing of information between users and systems, and this will be achieved in COMPASS if:
- content can be shared between the knowledge infrastructure and digital libraries (demonstrating interoperability with a broader range of information infrastructures);
- international standards are used for storage of and access to content;
- the architecture allows the content from other systems that use appropriate international standards to be included and
- the full (including semantic) content from the knowledge infrastructure can be included in other systems through selected international standards that define the interface and content of the knowledge infrastructure.

Intelligence refers to the type of knowledge infrastructure we are aiming to create. Existing infrastructures that support scientists (for example, Google Scholar) have limited intelligence in that they are based only on the use of search terms. This success factor will be achieved if:
- scientific resources that relate to the area of interest are found, even if they do not use the same exact terms;
- scientific resources that use similar theories or models are found, even if they are not in the same area of research;
- web services conforming to a limited set of standards can be directly executed from the discovery interface;
- users can create tags as part of their interaction with the scientific resources themselves; and
- future options for extension to more advanced reasoning and visualisation are possible.

Scalability refers to the ability of the knowledge infrastructure to become a wider infrastructure for scientific knowledge across a range of domains. This will be achieved if:
- the solution is interoperable, so other, existing systems can be communicated with;
- large amounts of knowledge can be potentially accommodated;
- the solution provides interfaces to other information repositories, rather than copying their content into the COMPASS repository.

4. Project Outputs

4.1 Deliverables

The project will deliver the following:

- A workshop to capture domain knowledge from experts in the marine domain.
- A project web site describing the project, including project and technical documents, links to related projects and other information of interest to third-parties.
- A scenario and use case description.
- A report describing the range of options for the architecture and the selected option and reasons for the selection.
- A research report describing the detailed information architecture.
- A standards specification for the semantic registry to be submitted to the OGC.
- A report describing the tools that are available for ontology engineering and management.
- Installed, configured and operating software components for:
    - ontology creation and management;
    - tagging;
    - a registry;
    - a WRS interface;
    - knowledge management, integration and reasoning;
    - scientific resource discovery allowing tagging and returning content from the knowledge infrastructure and digital library;
    - an interface to the digital library.
- A report synthesising the lessons learnt regarding the discovery user interface into best practice and advice for the wider community.
- A report analysing standards for communication between knowledge infrastructures and digital libraries and, if possible, a description of an extension to the existing digital library standards to include semantics.
- A report describing the work to integrate knowledge infrastructures and digital libraries and future research directions.
- A report synthesising the lessons learnt about the overall architecture into best practice and advice for the wider community.
- A report reviewing existing domain ontologies.
- Domain ontologies for selected marine instruments and parameters (to also be donated to MMI).
- A set of semantically annotated resources.
- A report comparing the use of ontologies and the use of informal user tags and describing lessons learnt and advice for the wider community.
- A scientific resource ontology for the resources included in the knowledge infrastructure, to be published.
- Application ontologies for the project, linking the scientific resources with relevant marine domain ontology concepts and other scientific resources.
- An operating knowledge infrastructure, populated with content from the marine domain, supported by domain, scientific knowledge and source ontologies and accessible through an intelligent discovery interface.
- Quarterly and final project reports as required by JISC.

4.2 Knowledge and Other Outputs

The project team expects to gain significant knowledge during the course of the project in ontology engineering, the structure of scientific knowledge, the design of semantic registries and knowledge management. The project deliverables have been designed to ensure that this knowledge is shared both in academic and commercial environments. Several academic papers will be written, but reports will also be created that describe best practice and lessons learnt.

The knowledge we gain in the coastal marine domain will be communicated through the domain ontology, and we also expect to impart some knowledge in ontology engineering and knowledge modelling to domain experts through the workshop process.

We also intend to donate the marine domain ontologies created to the Marine Metadata Interoperability Project[13] (MMI) to enable their dissemination among marine scientists.

All deliverables will be made available on the web site, and will be publicly accessible.

5. Project Outcomes

The COMPASS project will have four main enduring outcomes. Firstly, it will develop more advanced and richer knowledge infrastructures, and it is expected that our work will set the scene for future development in this area. The software and ontologies that we produce will in themselves be useful outcomes, but the work is also likely to stimulate further development and enhancement in the area.

Secondly, the project will extend the work being done in the inclusion of semantic information in SDIs, and make the work known to others working in the area. This is likely to further the development of semantic spatial data infrastructures and the semantic geospatial web, as well as having impacts on the non-spatial area.

Thirdly, COMPASS will develop work in the area of formal versus informal structures for the representation of knowledge (ontologies vs. informal tagging), and will explore options for this. The results will be published in an academic context, and will thus influence future work in the area.

Finally, the work will create a significant bridge between digital libraries and knowledge infrastructures, demonstrating how the two can support each other and promoting overall interoperability between different types of systems.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
JISC	The project represents a new development in e-infrastructure, specifically with regard to digital libraries and geospatial interoperability, and will affect future work in these areas.	Medium
Universities and researchers	The project will provide a foundation for a more advanced tool to help researchers do their work.	High
Marine scientists	The project will develop marine ontologies and capture marine knowledge, as well as demonstrating a system to help with marine research.	High
Other scientists	The project will provide a framework for a more advanced tool to help scientists do their work.	High
Librarians	The project is a development in the management of information, and if successful and extended by others, will provide a new approach to be used in libraries and will impact on existing digital library projects.	Medium

[13] http://marinemetadata.org/

Semantics and ontologies researchers	The project will develop new ways of capturing and using semantics information.	Medium
Other knowledge infrastructure projects	The project will advance developments in the field.	High
Other digital library projects	The project will suggest new ways in which digital libraries can be used and interact with other systems.	Medium
Other SDI projects	The project will advance the use of semantics in SDIs.	High
Semantic web projects	The project will provide methods and techniques to extend the semantic web.	Medium
OGC Catalogue and Geosemantics working groups	The project will develop standards and methods to extend the work of these groups.	High
Other digital discovery services (e.g. Google Scholar)	The project will present an alternative approach that may be adopted by such services.	Medium
Knowledge scientists	The project will develop a model to formalise scientific knowledge.	Medium
Students	The project will provide a framework for a more advanced tool to help students do their work.	High
Wider GI community	The project represents a new development in geospatial interoperability, and will affect future work in the area.	Medium

The most important stakeholder groups identified in this table are having input into the project through the steering committee and project consortium. A number of marine researchers and scientists are involved in the steering committee, as is a knowledge infrastructure expert. The project consortium includes a number of geospatial, geosemantics and standard experts.

7. Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing				
Departure of key staff.	4	2	10	▪ Ensure knowledge is shared among staff members and documents added to wiki. ▪ Ensure critical project areas are covered by more than one person.
Staff without required skills.	2	2	4	▪ Include time and resources to train staff where needed.
Organisational				
Variation in ideas about the objective and focus of the project between partners (research focus vs. practical implementation focus).	3	3	9	▪ Spend time at the start of the project creating a common vision. ▪ Clearly define and agree on responsibilities and requirements for different project components.

Departure of an institution from the consortium	1	3	3	▪ Secure by contract lock-in. ▪ Ensure knowledge is shared among staff members and documents added to wiki. ▪ Include people in the team who can take on the tasks of others if necessary.
Insufficient resources to complete tasks as required.	3	4	12	▪ Ensure project is well scoped and scope creep is avoided. ▪ Include flexibility in the project to allow components to be implemented in different ways or less important components distinguished from essential components. ▪ Ensure partner sign-off on the deliverables and timetable. ▪ Carefully track progress against the Gantt chart.
Technical				
Software components unable to provide required functionality or insufficiently open.	2	4	8	▪ Carefully evaluate software before selection. ▪ Ensure flexible architecture to allow substitution.
Excessive customisation required to some software components.	2	4	8	▪ Carefully evaluate software before selection. ▪ Ensure flexible architecture to allow substitution.
Digital library content unable to be efficiently and appropriately accessed because of access constraints.	2	4	8	▪ Ensure flexible architecture to adopt creative solutions. ▪ Ensure that the knowledge infrastructure architecture is not dependent on the digital library interface.
Web services and data to match the selected use case and domain ontology unavailable.	2	5	10	▪ Select the use case and ontology with the available web services and data in mind. ▪ Select a use case and ontology that has a wide range of possible areas of applicability.
Digital library standards insufficient to handle semantic content.	5	2	10	▪ Extend existing standards or create a new standard.
External suppliers				
Vendors for software components do not supply as required.	1	3	3	▪ Ensure flexible architecture to allow substitution.
Legal				
Disputes arise over IPR.	1	4	4	▪ Ensure IPR is clearly covered in

				the consortium agreement.
Security issues cause data use problems.	2	3	6	▪ Select data where security issues are not of concern.

8. Standards

Name of standard or specification	Version	Notes
[ebRIM] OASIS ebXML Registry Information Model	3.0	ebRIM will be used for the registry information model.
[CSW] OGC 04-021r3 OGC® Catalogue Services Specification	2.0.2	CSW is an OGC standard for catalogues, and will be used as a basis for the registry.
[WRS] OGC 05-025r3 OGC® ebRIM (ISO/TS 15000-3) Profile of CSW	2.0	WRS is an ebRIM profile of CSW, bringing the two together, and will be used as an interface to the registry and extended to include ontologies as part of this project.
[WFS] OGC 04-094 Web Feature Service Implementation Specification.	1.1.0	WFS is a specification for an interface to allow access and querying of geographic information. WRS web services may be included as resources in the knowledge infrastructure.
[WPS] OGC 05-007r4 Web Processing Service Implementation Specification.	1.0.0	WPS is a specification for an interface to allow processing of geographic information. WPS web services may be included as resources in the knowledge infrastructure.
[WMS] OGC 03-109r1 Web Map Services Implementation Specification.	1.3.0	WMS is a specification for an interface to allow generation of maps. WMS web services may be included as resources in the knowledge infrastructure.
[WCS] OGC 06-083r8 Web Coverage Services Implementation Specification.	1.1.0	WCS is a specification for an interface to allow access to geographic information in the form of coverages. WCS web services may be included as resources in the knowledge infrastructure.
Z39.50	2003	Z39.50 is a protocol for searching and retrieving information and may be used for the digital libraries interface.
SRW (SRU via HTTP SOAP)	1.2	SRW is a protocol for searching and retrieving information and may be used for the digital libraries interface.
OAI-PMH	2.0	OAI-PMH is a protocol for harvesting metadata descriptions in an archive and may be used for the digital libraries interface.

9. Technical Development

The project will use Rapid Application Development (RAD) techniques. The objective is to rapidly deliver prototype software. The nature of the project means that RAD is both suitable and practical, allowing for iterative design and efficient reuse without the lead times associated with other development paradigms.

Software components will be tested individually as part of their configuration and customisation, and then overall testing of the entire knowledge infrastructure including all components and the interactions between them will be performed as shown in the project Gantt chart.

Software will be stored and versions managed using Subversion. UML diagrams will be used to document software where possible.

10. Intellectual Property Rights

For the avoidance of doubt, all background information and know-how used in connection with the deliverables shall remain the property of the Party introducing the same.

Any Results or materials created and or used for the deliverables and all rights therein shall be owned, in the first instance, by the party creating those Results. In the event that any Results are jointly created the parties shall have joint ownership of those results. The project will attempt to ensure that all deliverables are clearly marked identifying ownership of the intellectual property in them

Each consortia member will grant to the other(s) a non-exclusive free licence to use the deliverables, including the software tools, for the purposes of carrying out the Project and for teaching and research purposes during the Project and after the end of the Project. The Parties have no obligation or liabilities to maintain or support the deliverables beyond the timescales agreed with JISC.

The Parties hereby acknowledge that they are required at the end of the Project to provide to those institutions defined by Section 65(5) of the Further and Higher Education Act 1992 a non-exclusive licence to use the Deliverables for non-commercial purposes. The Parties hereby agree to supply such copies and to grant such licences as may be required by the JISC.

Project Resources

11. Project Partners

The project will consist of a consortia led by **EDINA,** with primary partners **Social Change Online UK**; the **Institute for Geoinformatics at the University of Muenster**, Germany and the **Digital Enterprise Research Institute at the National University of Ireland, Galway,** and associate partner the **School of Geosciences at the University of Edinburgh**.

The consortium agreement had not been signed on the date of lodgement of this document, but signing is expected within 2 weeks.

The following table lists all project partners, key staff and their roles.

Organisation	Individual	Role (in brief)
EDINA National Data Centre, University of Edinburgh, UK	Dr David Medyckyj-Scott	Project director and sponsor.
	Anne Robertson	Project manager from 31 March 2008.
	Chris Higgins	Technical advisor, particularly in relation to registries and web services, coordinator of programming activities performed by EDINA.
	Mark Small	Software engineer. Responsible for installation, customisation and development of the knowledge infrastructure architecture including the registry, knowledge management middle-

		ware, discovery interface and part of the digital libraries interface and creation of web services.
Social Change Online UK	Dr Kristin Stock	Project manager until 31 March 2008 (subcontracted through EDINA). Technical lead for the project (subcontracted through EDINA). Responsible for overall architecture of the knowledge infrastructure, information architecture for the registry and input into ontology development.
University of Edinburgh, School of Geosciences, Institute of Geography UK	Dr Femke Reitsma	Responsible for development of the domain ontologies, coordination of stakeholders, design and management of the experts workshop and input into development of the scientific knowledge and source ontologies and selection of content.
	Yang Ou	Assisting Dr Reitsma at start of the project.
Institute for Geoinformatics, University of Münster Germany	Professor Werner Kuhn	Leader and adviser to the Muenster Team.
	Florian Probst	Responsible for development of the scientific knowledge and source ontologies, ontology engineering, providing input into the overall architecture and selection and implementation of the ontology management tools.
	Mohamed Bishr	Responsible for ontology engineering, providing input into the overall architecture and selection and implementation of the ontology management tools and knowledge management middleware.
	Jens Ortmann	Responsible for ontology engineering, assisting with creation of the domain ontology, providing input into the overall architecture and selection and implementation of the ontology management tools.
Digital Enterprise Research Institute, National University of Ireland, Galway	Sebastian Ryszard Kruk	Leader and scientific advisor to the DERI team.
	Maciej Dabrowski	Responsible for design and implementation of the digital library interface.

12. Project Management

EDINA will act as the lead partner for the purposes of project administration and finance. Overall responsibility for the project will rest with senior staff at EDINA. The project will be co-ordinated by a Project Director, Dr David Medyckyj-Scott. Project management for the beginning of the project (until 31 March 2008) will be subcontracted to Dr Kristin Stock (0.25 FTE), and thereafter will be carried out by Anne Robertson (0.4 FTE) based at EDINA.

A work package focusing on project management will be produced as part of the overall project plan. A series of full project meetings will take place at project kick-off, mid-term and close. To limit the T&S outlay, the majority of communications will be conducted virtually by video/teleconferencing. Day-to-day communications amongst project members shall be conducted by email and phone.

Where problems arise, these will be dealt with in the first instance at each site. If a problem persists, this will be reported to the Project Manager and the Project Director. If the problem cannot be resolved within the project, it will be referred to the JISC office.

The following table lists the Project Team members and their contact details.

Organisation	Individual	Contact
EDINA JISC National Datacentre, University of Edinburgh, UK	Dr David Medyckyj-Scott	d.medyckyj-scott@ed.ac.uk
	Anne Robertson	a.m.robertson@ed.ac.uk
	Chris Higgins	chris.higgins@ed.ac.uk
	Mark Small	m.small@ed.ac.uk
Social Change Online UK	Dr Kristin Stock	kristin.stock@nottingham.ac.uk
University of Edinburgh, School of Geosciences, Institute of Geography UK	Dr Femke Reitsma	femke.reitsma@ed.ac.uk
	Yang Ou	adamsiera@googlemail.com
Institute for Geoinformatics, University of Münster Germany	Professor Werner Kuhn	kuhn@uni-muenster.de
	Florian Probst	f.probst@uni-muenster.de
	Mohamed Bishr	m.bishr@uni-muenster.de
	Jens Ortmann	jens_ortmann@hotmail.com
Digital Enterprise Research Institute, National University of Ireland, Galway	Sebastian Ryszard Kruk	sebastian.kruk@deri.org
	Maciej Dabrowski	maciej.dabrowski@deri.org

The project management structure will also include a Steering Group. The roles of this group are to oversee the project, review progress, represent the best interests of the communities, and provide advice and guidance to the project team. In addition to the project director and project manager, the steering committee consists of the following people:

- Vacancy, semantics or knowledge infrastructure expert;
- Dr Roy Lowry, Technical Director, British Oceanographic Data Centre;
- Alex Midlen, CoastNet;
- Dr David Green, Aberdeen Institute for Coastal Science and Management;
- Jane Stephenson, Head of Information Services, National Oceanography Centre;
- Dave Cotton, Marine Data and Information Partnership;
- Vince Gaffney, Chair in Landscape Archaeology and Geomatics, University of Birmingham and
- Alison Turner, Programme Manager, Joint Information Systems Committee.

13. Programme Support

Specific requirements for support are not known at this time.

14. Budget

The budget is included in Appendix A

Detailed Project Planning

15. Workpackages

The project consists of the following work packages:

- WPx: Project Management;
- WP0: Use Cases;
- WP1: Coastal and Resource Ontologies;
- WP2: Architecture;
- WP3: Discovery Interface and
- WP4: Digital Library Interfaces.

Detailed project planning and scheduling has been undertaken using a Gantt Chart, included in Appendix B (separate document). The Gantt Chart shows the milestones and deliverables and identifies the work packages that each task or group of tasks falls within. The Gantt Chart has been used instead of the usual work package template with the agreement of the Programme Manager, as the Gantt Chart will be the live document used to manage the project.

The Gantt chart is detailed and changes are possible through the life of the project. For this reason, the most important milestones have been defined as controlled milestones. Controlled milestones are shown in red in the Gantt Chart and will be used to allow JISC to monitor project progress. Any changes to the Gantt Chart that affect controlled milestones and any circumstances that indicate that controlled milestones may not be met will be discussed with JISC. Changes to the Gantt Chart that do not affect controlled milestones will remain internal to the project. This approach will allow the Gantt Chart to be used as a living document, but ensure that JISC can still monitor project progress satisfactorily.

Internal deliverables are distinguished from external deliverables. Deliverables that are to be provided to JISC are shown in red, while internal project deliverables are shown in blue. Any changes to the Gantt Chart that change the delivery date of an external deliverable will also be discussed with JISC, while changes that only affect internal deliverables will not.

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Early 2008	The domain experts' workshop.	Was the workshop effective in collecting domain expert knowledge?	Observation and reflection.	The creation of a useful and effective ontology (see above) that reflects experts' knowledge.
Mid 2008	The marine domain ontology.	Does the marine domain ontology correctly reflect marine experts' knowledge? Is the marine domain ontology useful in supporting the KI?	Feedback to domain experts to ensure validity of the ontology. Feedback from the wider research community through publication of the ontology.	Domain expert feedback confirms the ontology's validity. Peer-review publication is accepted.

Early 2009	The semantic registry design.	Is the approach used to store and query ontologies in the registry appropriate, valid and practical?	An extension to the existing standard will be created and submitted to the Open Geospatial Consortium for ratification.	The extension is ratified by the Open Geospatial Consortium.
End of project	The ontology for scientific resources.	Does the scientific resource ontology support the KI in meeting project objectives? Is the structure of knowledge in the KS useful and intuitive to users?	Feedback from the wider research community through publication of the ontology. Appropriate operation of the KI. Observation and questionnaire.	Peer-review publication is accepted. The KI works and allows users to find scientific resources using various discovery search mechanisms. Observation and questionnaire results show support from 75% of users.
End of project	The demand for and interest in the KI as a concept.	Is the research community interested in the idea of knowledge infrastructures?	Project web site hits.	The web site has more than 1000 hits over the life of the project.
End of project	The KI architecture.	Is the KI architecture appropriate?	Peer-reviewed paper submitted to journal. Effectiveness of KI when complete.	The KI works and allows users to find scientific resources using various discovery search mechanisms. Peer-review publication is accepted.
End of project.	The digital library interface.	Is the approach used to include semantic information in the digital library content appropriate, valid and practical?	Implementation and publication.	Digital library content is successfully included in the KI and peer reviewed publication of the work is accepted.
End of project	The discovery and invocation user interface.	Is the discovery interface intuitive and does it provide a useful result for users?	Observation and questionnaire.	Observation and questionnaire results show support from 75% of users.
End of project	The usefulness of the KI to users.	Is the KI useful to scientists?	Observation and questionnaire.	Observation and questionnaire results show support from 75% of users.
End of project	Resource tagging.	Is resource tagging useful to scientists? Do users of the KI use the tag cloud to help them	Observation and questionnaire.	Observation and questionnaire results show support from 50% of users.

		search?		

17. Quality Plan

Output	Domain Experts' Workshop				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
Feb 2008	Use of best practice methods and techniques.	Adoption of methods used in the SWING project.	Participation of SWING project staff. Completion of paper evaluating the process.	Femke Reitsma	
March 2008	Validity of outcome (ontology).	Feedback to domain experts.	Ontology is accepted by 80% of participating domain experts.	Femke Reitsma	

Output	Project Web Site				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
End of project.	Usability	Monitoring, evaluation and feedback.	Lack of negative feedback from users.	Mark Small	
End of project.	Accessibility	Monitoring.	Web site is available and accessible (up) by 90% of users 90% of the time over the course of the project.	Mark Small	

Output	Scenario and Use Case Description				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
Feb 2008	Validity	Input from domain experts.	Support from domain experts.	Femke Reitsma	

Output	Knowledge Infrastructure Design				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
End of project.	Use of best practice methods and techniques.	Research and study of existing methods.	Acceptance of submitted paper. Lack of negative feedback from peer review.	Kristin Stock	
End of project.	Adherence to	Research and study of	Acceptance by standards	Kristin Stock, Maciej	

	standards	appropriate standards.	organisations of any necessary standards extensions or new standards.	Dabrowksi	
End of project.	Fitness for purpose	Prototyping to ensure that system resulting from the design meets requirements.	Positive feedback from users testing the KI.	Kristin Stock, David Medyckyj-Scott	

Output	**Implemented Knowledge Infrastructure and Discovery Interface**				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
End of project.	Use of best practice methods and techniques.	Use of Rapid Application Development (RAD) techniques. Documentation of design using UML.	An operating system with sound architecture and implementation.	Chris Higgins, Mark Small	
End of project.	Adherence to standards	Research and study of appropriate standards.	Use of appropriate UML, OGC and DL standards.	Chris Higgins, Mark Small	
End of project.	Fitness for purpose	Prototyping to ensure that system resulting from the design meets requirements.	Positive feedback from users testing the KI.	Chris Higgins, Mark Small	
End of project.	Usability	Monitoring, evaluation and feedback.	Positive feedback from users.	Chris Higgins, Mark Small	

Output	**Digital Libraries Interface**				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
End of project.	Use of best practice methods and techniques.	Research and study of existing methods.	Acceptance of submitted paper. Lack of negative feedback.	Maciej Dabrowksi	
End of project.	Adherence to standards	Research and study of appropriate standards.	Proposals made to relevant bodies regarding standards requirements for semantic content.	Maciej Dabrowksi	

Output	**Ontologies**

Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
End of project	Use of best practice methods and techniques.	Research and study of existing methods.	Acceptance of paper evaluating the process.	Werner Kuhn	
End of project	Validity.	Feedback by domain experts and users.	Acceptance of papers describing the ontologies.	Werner Kuhn	

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
Ongoing	Web site.	JISC and GI communities, domain experts, scientists	To raise awareness and inform.	The purpose of the project, the benefits that it will offer to scientists and the scientific contribution it will make.
Feb 2008	Stakeholders workshop.	Marine domain experts	To raise awareness and engage experts in the project.	The benefits of a semantically-enabled knowledge infrastructure. The importance of domain experts in the process and our desire for ongoing involvement.
Ongoing	Conferences and workshops.	JISC and GI communities, domain experts, scientists	To inform.	The benefits of a semantically-enabled knowledge infrastructure.
Ongoing	Scientific journal papers.	The scientific research community.	To inform and extend scientific knowledge.	The approach adopted for the KI, the results of evaluation.
Ongoing	Reports.	JISC, GI and DL communities.	To inform.	Lessons learnt from various parts of the project.
End of project	Standards specifications.	Standards community, GI community	To inform, to provide specifications for future use and extension.	The standards used and created for the project and how they can be used to provide interoperability.
End of project	Demonstration application.	JISC and GI communities, DL community,	To promote the results and inform.	The benefits of a semantically-enabled

		domain experts, scientists		knowledge infrastructure. How a KI might be used.

19. Exit and Sustainability Plans

Project Outputs	Action for Take-up & Embedding	Action for Exit
Project reports and other written outputs	Available for wider dissemination.	Archived and available for access via project web site and JISC.
Prototype application	Persist as prototype application.	EDINA committed to making available for 6 months post project.
Scientific papers	Published in scientific journals.	Revisions after peer review as required after the project is complete.
Standards specifications.	Submitted for ratification to standards organisations.	
Ontologies	Published in scientific journals and on web site.	EDINA committed to making available for 6 months post project.

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Ontologies	Other researchers and developers of systems may use or extend the ontologies created in the project.	Researchers and ontology engineers will adopt and extend the ontologies provided they are known about and accessible.	The ontologies should be published, provided on a web site for download long after the project is completed (for example on the personal web site of the person responsible for publication), and donated to MMI.
Prototype application	The prototype is an operational system, and may be extended and have further content added to provide a more complete product.	Seek funding for later projects to extend the work. Publish on the web and allow users to add content. Consider possibilities for commercialisation.	The prototype may not be industrial-strength, so use as is may be limited.
Scientific knowledge	The project involves the creation of new scientific knowledge in a number of areas: the modelling of scientific resources, the architecture of knowledge infrastructures and the use of semantics to support discovery	The research may be adopted and extended by other researchers, provided it is published in reputable scientific journals.	

	of scientific resources.		
Standards specifications	Other developers of systems for discovery may be interested in the use of semantics for discovery of resources.	The standards may be adopted and extended by others working in the area.	

Appendixes

Appendix A. Project Budget

Directly Incurred Staff	**April 07–March 08**	**April 08–March 09**	**TOTAL £**
Project manager [redacted]	[redacted]	[redacted]	[redacted]
Technical Advisor [redacted]	[redacted]	[redacted]	[redacted]
Senior Academic Liaison [redacted]	[redacted]	[redacted]	[redacted]
University of Edinburgh Computing Support [redacted]	[redacted]	[redacted]	[redacted]
Software Engineer [redacted]	[redacted]	[redacted]	[redacted]
Project RA [redacted]	[redacted]	[redacted]	[redacted]
Total Directly Incurred Staff (A)	**£50,123**	**£106,188**	**£156,310**
Non-Staff	**April 07–March 08**	**April 08–March 09**	**TOTAL £**
Travel and expenses	[redacted]	[redacted]	[redacted]
Hardware/software	[redacted]	[redacted]	[redacted]
Dissemination	[redacted]	[redacted]	[redacted]
Other – subcontracting to MUSIL	[redacted]	[redacted]	[redacted]
Other – subcontracting to SCO	[redacted]	[redacted]	[redacted]
Other – subcontracting to DERI	[redacted]	[redacted]	[redacted]
Total Directly Incurred Non-Staff (B)	**£56,533**	**£115,567**	**£172,100**
Directly Incurred Total (A+B=C) (C)	**£106,656**	**£221,754**	**£328,410**
Directly Allocated	**April 07–March 08**	**April 08–March 09**	**TOTAL £**
Directly Allocated Total (D)	**£0**	**£0**	**£0**
Indirect Costs (E)	**£56,243**	**£112,487**	**£168,730**
Total Project Cost (C+D+E)	**£162,899**	**£334,241**	**£497,140**
Amount Requested from JISC	**£134,777**	**£277,997**	**£412,775**
Institutional Contributions	**£28,122**	**£56,243**	**£84,365**
Percentage Contributions over the life of the project	**JISC** 83 %	**Partners** 17 %	**Total** 100%

Appendix B. Gantt Chart

See separate document.