Renardus: Building an Academic Subject Gateway Service in Europe

Lesly Huxley

1 Introduction to the Renardus project

The aim of Renardus is to build a single service on the Web allowing users to search and browse from one site the quality-checked and Internet-accessible scientific and cultural resource collections distributed across existing European subject gateways. Renardus is a collaborative project of the EU's fifth framework 'User-friendly Information Society' programme. The project started in January 2000 with project partners drawn from national libraries, university research and technology centres and subject gateways across Europe. Renardus partners will investigate related technical, information and organisational issues, build a pilot system and finally develop a fully-operational service for release in June 2002. Full details are available from the project Web site at http://www.renardus.org/

2 Rationale

As the Internet continues to expand it is clear that no single, publicly-funded subject gateway or national gateway/digital library initiative can hope to identify, catalogue and organise all the Internet resources available to support the academic and research communities of Europe. Collaboration is needed to secure sustainable, high quality services. Several subject and national gateway services in Europe have already been developed on the basis of collaboration: the Finnish Virtual Library[1], the UK's Resource Discovery Network[2] and the DutchESS[3] service in The Netherlands operate on a national scale with the cooperation and contribution from libraries and universities in their respective countries. The Nordic countries have collaborated to develop an agriculture and forestry gateway, NOVAGate[4], whilst mathematicians can access relevant publications through the EULER[5] initiative which brings together information from across Europe.

Renardus aims to build on these collaborative initiatives: the fully-operational 'broker' service will allow extension of subject coverage to include a wide range of disciplines and involvement of partner organisations across Europe. One of the key features of the project is its collaborative and pragmatic approach: partner organisations from Denmark, Finland, France, Germany, the Netherlands, Sweden and the UK are working together to address the considerable implications for development of both technical and information standards as well as business and sustainability issues. The potential benefits from such collaboration include scale economies in the areas of metadata creation, abstracting and indexing by service providers and a more sustainable level of quality in mediated resource discovery. Overall, access to scientific and cultural resources in Europe should be improved through aggregation and improved consistency of collections and a common understanding of academic users'

---

[1] URL: <http://www.jyu.fi/library/virtuaalikirjasto/engvirli.htm>

2 URL:<http://www.rdn.ac.uk/>

[3] URL: <http://www.kb.nl/dutchess/>

[4] URL: <http://novagate.nova-university.org/>

[5] URL: <http://www.emis.de/projects/EULER/>

needs. Participating services should also benefit from the collaborative framework in terms of improved sustainability and a stronger position against international competition.

3 Data Gathering and Research to develop the Service Specification

Project work is distributed over three main phases: *Data Gathering and Research; Analysis and Design* and *Implementation.* A complementary element running through all phases involves *Dissemination and Support* to ensure full exploitation of - and continuing collaboration in - the fully-operational Renardus service. Much of the data gathering and research work was completed in the first six months of the project in order to feed into development of a functional specification for the Renardus service.

3.1 User and Functional Requirements

User requirements were collated at various levels. These included an investigation of target service provider (ie participating gateways) requirements in the areas of interoperability, interface and 'branding' issues, and collation of end users' evaluations of existing subject gateway services.

Respondents to the service providers' survey rejected the concept of a central repository to which all metadata is routinely copied in favour of a centralised subject index which would forward queries to relevant gateways. Service providers also specified the metadata schemas and protocols Renardus should support (Dublin Core[6] semantics and RDF/XML[7] syntax for metadata records, with Z39.50[8] and WHOIS++[9] used as protocols for communication between Renardus and the brokered gateways). Support for both browse and search functions was high and providers also felt a need for mappings between metadata formats to support a consistent presentation of search results. Mappings between classification schemes - even if only at the highest level - are also needed so that the cross-browsing functionality can be implemented in the pilot system.

Renardus partners also gathered data from a range of end user surveys previously undertaken by participating gateways. Generally, these showed that users are more at ease with finding their way around in gateways than is the case with the 'rest' of the Internet and that they appreciate the quality and currency of resources offered. The surveys also offered insights into user search and browse behaviours and preferences.

A third strand of research intended to inform the design of the Renardus broker service involved an initial analysis ('Use Cases') of the functional requirements of such a system from both end user and provider perspectives. Use cases are descriptions of how various 'players' will actually use the service, without specifying any technical solutions for how this functionality might be achieved. The individual use cases are intended to be quite narrow in their focus, and cover functions such as: performing a simple search, cross-browsing by subject, or displaying results. From the perspective of service administrators, other use cases cover requirements such as: maintaining metadata indexes, assuring data quality, or inserting data into Renardus. The use cases and activity diagrams are based on the Unified Modelling

---

[6] URL: <http://purl.oclc.org/dc/>

[7] URL: <http://www.w3.org/RDF/>
URL: <http://www.w3.org/XML/>

[8] URL: <http://lcWeb.loc.gov/z3950/agency/>

[9] URL: <http://src.doc.ic.ac.uk/computing/internet/rfc/rfc1913.txt>

Language (UML)[10]. This functional specification work is closely linked with the earlier user requirements work, the scoping of the pilot service and development of the data model and technical solutions.

3.2 Service Scope and Data Model

An early task for project partners was to agree the scope of the pilot system (and objectives for the fully-operational Renardus service). The resulting Scope Document addresses issues such as subject and geographical coverage, definitions and criteria for participating gateways. The terms and definitions of concepts used to scope the pilot system are based on those defined in the DESIRE Information Gateways Handbook[11]. A number of new concepts are also introduced to help communicate and promote Renardus' aims and objectives, in particular concepts such as “open subject gateways” and “resource discovery broker services”. Renardus partners have agreed an initial set of thirteen starting points for the pilot system: the intention is to refine these as the project progresses.

Renardus partners have also developed a minimum common set of metadata elements supported by participating gateways as the first steps towards defining the data model for the service. Gateway service providers' responses to a questionnaire were collated, providing information about collection descriptions, target groups, resource categories, quality criteria, controlled vocabularies etc., as well as descriptions of their respective metadata sets. A metadata mapping has been drawn up as the basis of a minimum common set of metadata elements. Four of the gateways have metadata schemes based on IAFA/ROADS templates, two use individual schemes, and two use schemes based on Dublin Core. As most of the partners use ROADS/IAFA as an exchange a simple exchange of metadata should be possible. The data model will continue to be developed and refined as the pilot system is built and tested. The following elements will form the basis of the minimum set: DC.Title; DC.Creator, DC.Description, DC.Identifier, DC.Subject, DC.Publisher, DC.Language, DC.Type, with further refinement based on newly-announced DC qualifiers.

3.3 Review of Broker Models

Renardus partners are facing one of the biggest challenges for those who are currently developing digital libraries: how to provide integrated access to the wide range of distributed and heterogeneous Internet information resources and services available. The success of this integration is seen as beneficial both to libraries and their end users. As a first step in developing the architecture that will underpin the Renardus broker system, Renardus partners undertook an extensive and comprehensive review of 18 existing broker models that have been developed for a variety of existing services, projects and initiatives. The models were chosen because they were perceived to be to the digital library context of Renardus and were 'mapped' against the generic model known as the MODELS Information Architecture (MIA)[12]. The review groups the 18 models into four broad categories on a continuum from simple to complex. A tabular summary of the review of all protocols used makes it clear that a majority of the broker models reviewed used two protocols: HTTP and Z39.50. The review also lists the wide variety of software that has been applied or developed to support the implementation of the broker models described.

---

[10] See description in: UML Distilled Second Edition: A Brief Guide to the Standard Object Modelling Language,(2000), Martin Fowler, with Kendall Scott. Addison-Wesley: IBSN 0-201-65783-X

[11] http://www.desire.org/handbook/

[12] MODELS - MOving to Distributed Environments for Library Services
URL: <http://www.ukoln.ac.uk/dlis/models/>

3.4 Review of Standards and Technologies

The architectural design and technical implementation of the pilot system will be based on existing and emerging standards and technologies to ensure that the functionality of the operational broker service is extensible. Interaction of the broker with existing resource collections is to be carried out according to open standards, for example Z39.50. A range of existing models will be accommodated within the broker service. Renardus partners have undertaken an initial review focusing on ten standards and technologies. The review provides a brief overview of functionality and strengths and weaknesses in the Renardus context.

3.5 Business and Organisational Models

Business issues are an important - but often neglected - part of the development of sustainable gateway services, for cooperation between gateways and for the development of broker services such as that proposed by Renardus. Gateway services are increasingly in transition from short-term funded 'research and development project' status to fully-operational 'service' status with the requirement for longer-term sustainability. Within Renardus, work is currently underway on a wide-ranging review of business models in operation in gateways *outside* the Renardus project (a detailed review of participating gateways' business models will be the focus of a later deliverable). Business and legal issues could be seen as an additional strand of the interoperability problem faced by collaborative 'broker' services such as Renardus: one that may ultimately be more difficult to solve than the technical and standards-based interoperability problems that formed part of the focus of 'first-generation' gateway-related projects like ROADS[13] and DESIRE[14].

Preliminary work has also been undertaken on issues relating to the organisational structure that might be adopted by the fully-operational Renardus service. Partners are considering an overview of the existing 'landscape', that is: which collaborative models are already in place; what kind of organisational models have existing services adopted; what is the role of national institutions. In further discussion, project partners will need to address questions such as "should national institutions play a key role in coordinating activities?"; "should the organisational framework be based directly on collaboration between the gateways themselves?". The answers to these and other questions will inform the development of the organisational model to be adopted by Renardus, and will play a key part in development of guidelines for gateways that may want to participate in the initiative.

4 Subsequent Phases

The results of the early work described above are already being analysed and will inform the design of the pilot system and implementation of a testbed environment in which issues of data interoperability and multilinguality can be explored. The pilot system will be available from mid-2001 and tested with the addition of at least one more gateway service. The fully-operational system will be implemented in June 2002.

5 Dissemination and Support for Opportunities for Participation

The findings of the project's research and development work is being disseminated through the Renardus Web site; a regular email newsletter - the Renardus News Digest[15]; and conference papers and presentations. Full reports of project findings are offered at the Web

---

[13] URL: <http://www.ilrt.bristol.ac.uk/roads/>

[14] URL: <http://www.desire.org/>

[15] URL: <http://www.renardus.org/news/>

site as soon as they are available, together with descriptions of related projects and services worldwide.

User guidelines on data interoperability will be developed to facilitate participation in the Renardus broker service. These requirements will be drawn up based on some of the early work of the project described above and continuing work on technical standards and solutions, business and organisational issues. We shall be presenting a workshop in September 2001 to support services interested in participating in the service. Identification of potential participants is underway and Renardus partners are particularly keen to enter dialogue with other initiatives who may want to participate in the verification of the pilot service (through addition of at least one other service) or in the fully-operational broker service.

For further information about Renardus, visit the Web site at:

http://www.renardus.org/

or contact the Dissemination and Support team at:

The Institute for Learning & Research Technology, University of Bristol
8-10 Berkeley Square, Bristol BS8 1HH, UK

Email: lesly.huxley@bristol.ac.uk