

Meeting Notes – 8th December 2008
CNI Fall Meeting (2009) – Global Registries Initiative
Renaissance Hotel - Washington DC

In attendance:

Adrian Burton - Deputy Director, Australian National Data Service / Project Leader, Australian Partnership for Sustainable Repositories;

James Farnhill - Programme Manager, Joint Information Systems Committee;

Jeremy Frumkin – Asst. Dean for Technology Strategy, University of Arizona;

Vic Lyte - Project Director, IESR & UK Repository Search, University of Manchester;

Lucy Nowell - Programme Director, Office of Cyberinfrastructure, National Science Foundation;

Andrew Treloar – Deputy Director, Australian National Data Service;

Pam Bjornsen - Director general / Directrice général CISTI / ICIST;

Neil Grindley - Programme Manager, Joint Information Systems Committee; and

Tim DiLauro – Digital Library Architect, John Hopkins University

Outcomes

The meeting resolved to:

1. Support the continuation and development of the Global Registries Demonstrator (<http://www.globalregistries.org/demonstrations.html>);
2. Establish a Global Registries Technical Working Group;
3. Establish a Global Registries Organisational Working Group; and
4. Build capacity and community engagement of individual registry services.

These initiatives will be coordinated by an interim steering group comprising of Adrian Burton, Jeremy Frumkin and Vic Lyte. Their first meeting will be Jan 7th. They will:

- Establish working groups by Feb 2009;
- Pursue US funding opportunities during Q1 2009;
- Liaise with NSF, JISC, NCRIS to promote coordination;
- Investigate involvement of other registry services (Canada, EU/ Driver, DataNet nodes, BAMBOO etc); and
- Present the concept and demonstrator at relevant venues (OpenRepositories2009).

The working groups will report by mid-year 2009. They will work by email and/or by targeted two day face-to-face meetings. Funding will be sought for meetings. The working groups will be charged to deliver the minimum lightweight requirements in their areas to allow the establishment of a global registries cooperative as soon as possible.

The Global Registries Technical Working Group will identify:

- A consensus on registry interchange format and procedures;
- Change management process to that consensus;
- A register of participating registries; and
- A roadmap for scaling up.

The Global Registries Organisational Working group will identify:

- The charter and scope of the global registries cooperative;
- A body to “own” and promote the interchange consensus;
- Data provider agreements and responsibilities;
- Methods to identify stakeholder and end-user demand;
- Develop use-cases to guide development and promotion of the Global Registries initiative based on user context and need;
- Quality control processes for the register of participating registries;
- A body to “own” the participant register; and
- A plan for scaling the organisation.

The Global Registries Demonstrator will pursue more sophisticated meta-search and aggregation services. It will also enlist participation of other registries into the demonstrator. The demonstrator will be developed in S1 2009 by the technical teams at the OCKHAM, IESR, and ANDS registries.

Notes form Discussion at the Meeting

General Context

This was a subsidiary discussion group held at the start of the CNI Fall 2009 Task Force Meeting.

Adrian Burton (AB) opened the meeting and offered a background context to the Global Registries initiative which had been promoted by the three present complementary US, UK & Australian pilots. The current consensus view was that, from a global registries perspective, there needed to be a collaborative synergy between partners. Due to current organizational and technical immaturity, it was not realistic for a single entity at this stage, to fulfil this role.

The collaborative vision and scope for the Global Registries initiative was to both to identify and maximise discovery and access to all available scholarly collections. However the context, definition and subsequent granularity of this resource were elaborated during the overall meeting.

There are four primary mission objectives:

1. Discovery & visibility of knowledge artefacts and means to access them;
2. Facilitating information brokerage;
3. Maximising knowledge federation; and
4. Promoting ongoing work supporting interoperability.

The shared (formative global) conceptual model relating to this initiative has five key attributes. These are:

- **A Collection** – of digital resources or possibly a knowledge or relationship artefact (what);
- A **machine-discoverable** or automatically connectable technical service entity comprising of SOAP, Representational state transfer ([REST](#)), or other evolutionary Open Knowledge Protocols and associated API's (where);
- **Agent** or Federated (Distributed Knowledge) Architectural protocol (where and how);
- Party, Security or **Community** Knowledge context (who); and
- **Activity** or Knowledge process / Work-flow staging model (when).

The collaborative discussions between the [Ockham](#), ORCA and [IESR](#) founders have matured into inputs into global specifications such as ISO2146. Workers such as Ann Apps have defined the [IESR](#) data model (schema) which has been used and extended as a baseline for this initiative and later generalized to a broader encompassing contextual framework.

As well as defined use-cases, a key challenge is to directly relate its usefulness both pragmatically and contextually into CRIS/ Subject and Context knowledge repository developments. The benefit of this will be to allow the combined registry hubs to be a partly-invisible 'Global Academic Yellow-Pages'. This will support knowledge providers as well as scalable search and discovery infrastructures/tools used by those researchers, teachers and practitioners.

Next Stage Challenges & Priorities

Recent experience from national subject-oriented repository initiatives indicates, from user acceptance testing, that researchers want **global** access to available knowledge and, if possible, associated and supplementary data, and consequently, learning & teaching artefacts.

Rather than building a monolithic component, it is important to utilise what is already out in the overall community. That is, to build the hub tier in both directions within the overall architecture, using open interfaces that both developers and users can easily utilize.

However a key change management challenge is to make the abstract benefits of the global registry network clearly visible and a 'first port of call' to the end-to-end communities it aims to serve.

We need to develop an open technology interface that can be applied across all library and distributed end-user knowledge communities.

Registries have 3 communities. These are:

- Direct human search & discovery layers be it Web 2.0 or individual or community search;
- By use of developers as part of a SOA-approach in emergent KM architectures or specific vector applications; and
- Outputs of 3rd-party Machine-driven (robot) gatherings such as Google.

There are three potential harvesting scenarios:

- Periodically acquire OAI PMH harvests on a single feed;
- Federated and dynamic search on a given query; and
- Academically-derived machine-driven (agent) gatherings.

There is an option to run the first two scenarios via a (new) registry repository.

However the fundamental question needs to be addressed as to why a 'library-style' approach is the best. For example, why can't we do a dynamic semantic & federated run-time search over a particular staging position?

In both scenarios, one would need a registry to reduce scaling impacts arising from user or machine-driven searches.

There needs to be current evaluation between the optimal use and choice of interface protocols such as SOAP – REST and beyond.

Presently, REST could be a good way of getting people on board, as participation is a key factor; other enterprise and scalable interface technologies such as SOAP etc. can be examined and implemented at a later date.

The key finding from recent internal and independent evaluations of the IESR UK was that one needed, as much as possible, to reduce technological, community and individual barriers to entry in uptake of these types of registries, bearing in mind their abstract nature within the overall knowledge landscape. Evidence of real & demonstrable buy-in by all stakeholder communities is essential despite the difficulties in immaturities of supportive and projective layers in the overall (global) Information Environment.

This mandates a change of management (over a technical) emphasis for this initiative. IESR has had notable success since it employed a person dedicated to outreach programmes and benefits/expectation activities within its top-down/bottom-up communities.