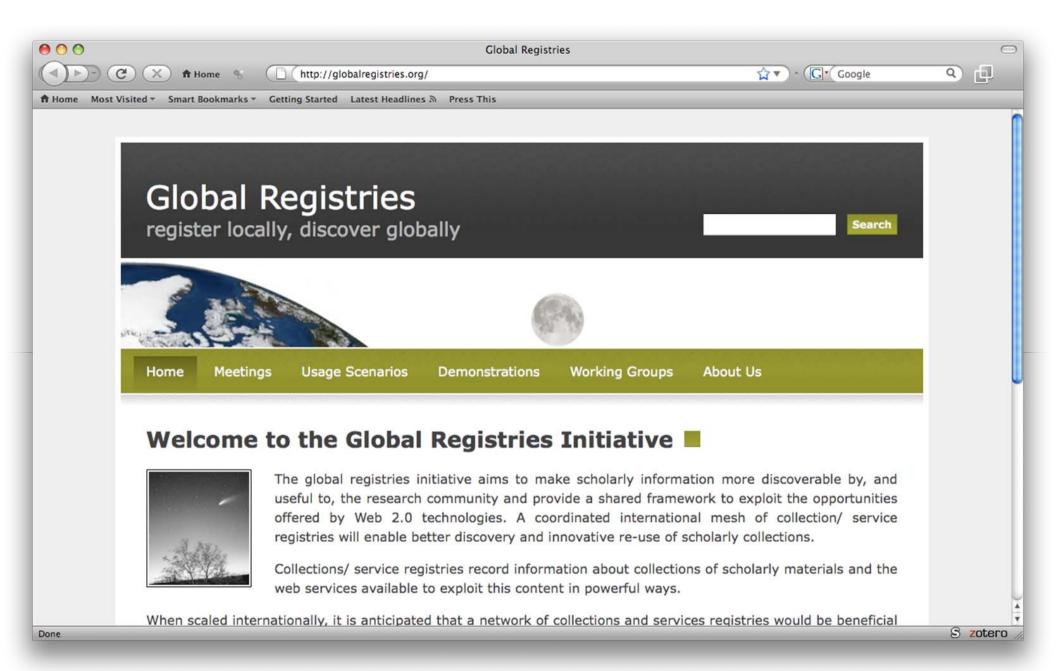
# The Global Registries Initiative: Progress Report and Software Demonstration

Chris Blackall, Australian National Data Service
Jeremy Frumkin, University of Arizona
Vic Lyte, University of Manchester

#### Who

- Jeremy Frumkin, Assistant Dean/Chief Technology Strategist, University of Arizona Libraries
  - Ockham Initiative http://www.ockham.org/
- Vic Lyte, Project Director, University of Manchester
  - Information Environment Service Registry http://iesr.ac.uk/
- Chris Blackall, Research Analyst, Australian National Data Service (ANDS)
  - ANDS Collections Registry https://services.ands.org.au/home/orca



#### What

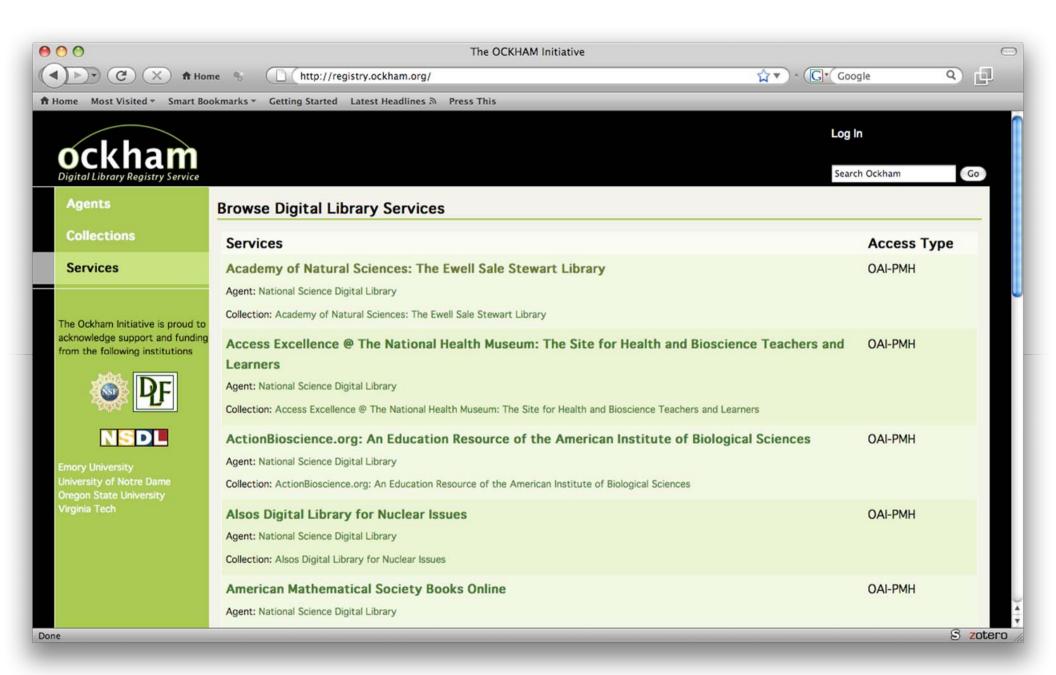
- http://globalregistries.org
- Collection/service registries record information about 'collections' of scholarly resources and related web services
- Collections? 'Any grouping of digital items'
- Global? 'register locally, discover globally'
- It works...stick around for the demo

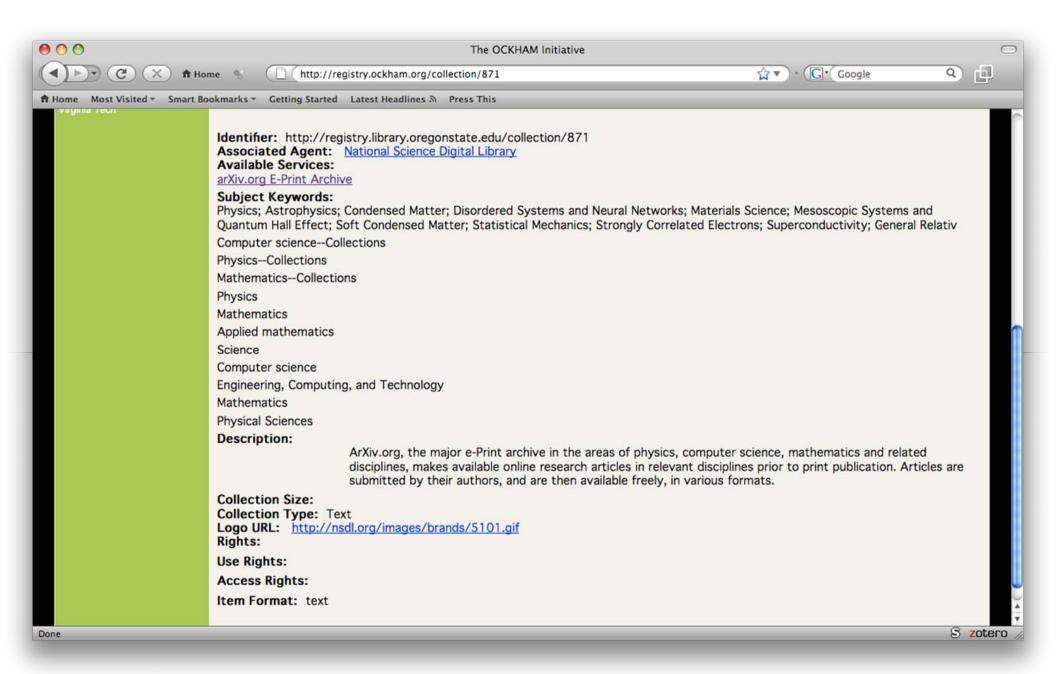
### Why

- Managing the data deluge
- Supporting data-intensive science (escience)
- Federating of research data collections
- Some questions:
  - Why not leave it to Google?
  - Why not use Semantic Web/Linked Data http://linkeddata.org/

#### How

- Collections and services records are harvested locally and aggregated nationally/globally
  - 'Federated' registry architecture
  - Metadata harvest protocol
  - Registry interchange formats
- Policies and metadata sharing agreements





#### UK / Information Environment Service Registry

- The Information Environment Service Registry (IESR) is the UK's free catalogue of electronic resources;
- Sponsored by the Joint Information Systems Committee (JISC) and part of shared services layer in overall IE landscape;
- The IESR aims to provide a reliable source of descriptive information about electronic resources and technical details about how they can be accessed via web or machine interfaces;
- http://iesr.ac.uk

#### IESR – Main Challenges

- Initial concept too technically abstract had to have a way to engage nontechnical stakeholder community – What are the benefits;
- High access barrier to easy and ubiquitous use detailed metadata schema required too much effort for collection and resource contributors;
- Restricting underlying technology needed to be more scalable, adaptive and ubiquitous;
- Wide range of collection / resource / artifact types potentially in scope;
- Need to optimize sister and allied search and discovery services:
  - www.intute.ac.uk/irs/demonstrator
- Need end-to-end clear (non-technical) use cases.
- Need stakeholder and organisational involvement / steerage / buy-in What are the benefits and how does it improve end-user experience?

#### IESR – Current priorities & activities

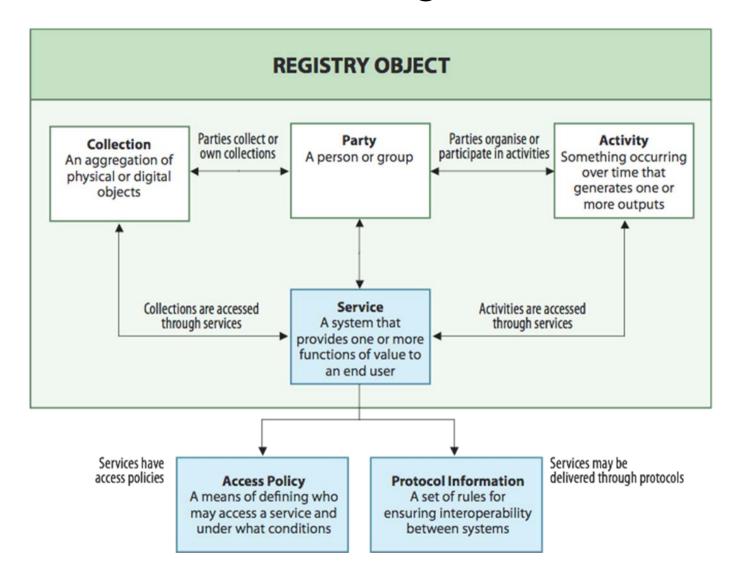
- Development of an outreach service and re-branding how can we add benefit to you;
- Developing a minimal metadata description set, based on user requirements, supported by in-house cataloguing resource;
- Embedding advanced technologies 'under the hood' with open interface on top

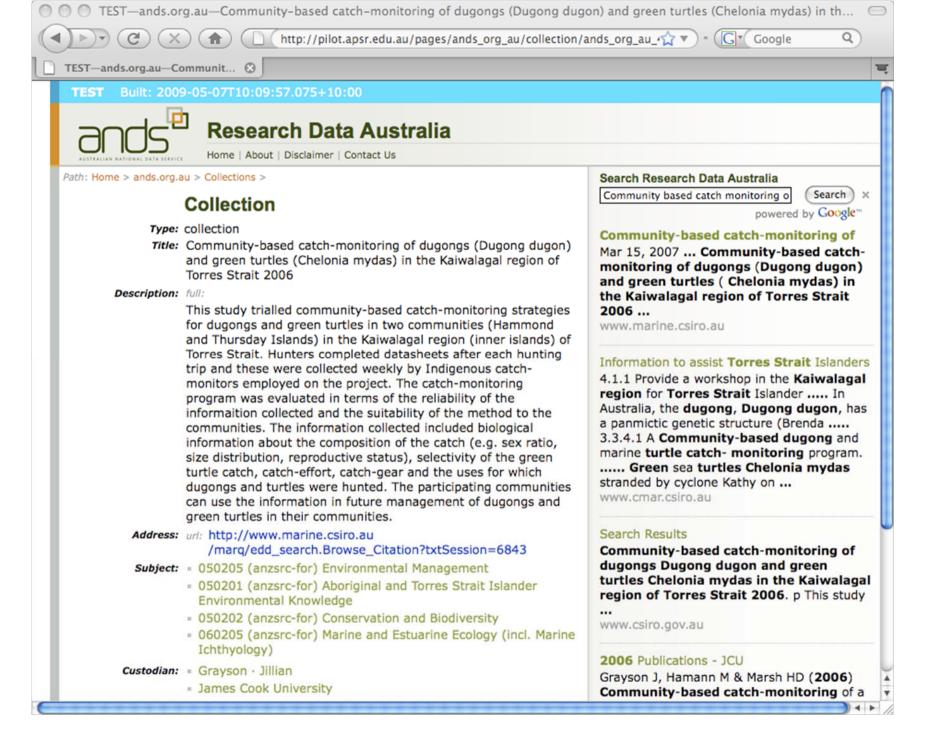
   Autonomy IDOL & Lucene for automatic metadata generation, dynamic discovery (clustering, concept matching) and unstructured search;
- Establishing early and visible wins:
  - Search engine optimization;
  - National Health Service;
  - Regional / National life-long learning initiatives;
  - Cross-sector knowledge initiatives;
  - Academic-oriented 'Yellow-Pages' for electronic resources.
- ➤ Academics want Global search scope ⇒ need for GRI topology to expose related assets to search services

#### Australia

- Collection/Service Registry initiative is a central part of national research infrastructure funded by the Australian Government
- Focus on building a Research Data Commons: data discovery and citation services
- Drivers: Innovation and research evaluation

## ISO 2146 Registry Services for Libraries and Related Organizations





## Next Steps

- Formation of working groups
  - Technical
  - Strategic / Organizational / Policy
- Set of experiments testing GRI

#### Demonstrator 1: ANDS Pilot

- The ANDS Collection/Service Registry global pilot: This demonstrator aggregates records from three collection/service registries into a combined registry run by the Australian National Data Service (ANDS)
- The aggregated pool of records is from Australia, UK, and USA, and can be searched and browsed
- Aggregation is performed using the OAI-PMH (Protocol for Metadata Harvesting) with the RIF-CS Interchange format
- https://devel.apsr.edu.au/cosi/orca/search.php

## Demonstrator 2: LibraryFind

- This demonstrates the usage of metasearch, or the querying of registries distributed over three different continents
- It supports distributed search protocols such as z39.50, SRU/SRW, and Open Search, as well as OAI-PMH aggregation
- http://rs.library.oregonstate.edu