

# The Key Role of Registries and Registry Standards in the Transition to a Federated Network of Repositories

Adrian Burton and Chris Blackall\*

Australian Partnership for  
Sustainable Repositories (APSR)

1





- **Registries and Registry Services**
- **Federated Networks of Repositories and Mass Data Storage Facilities**

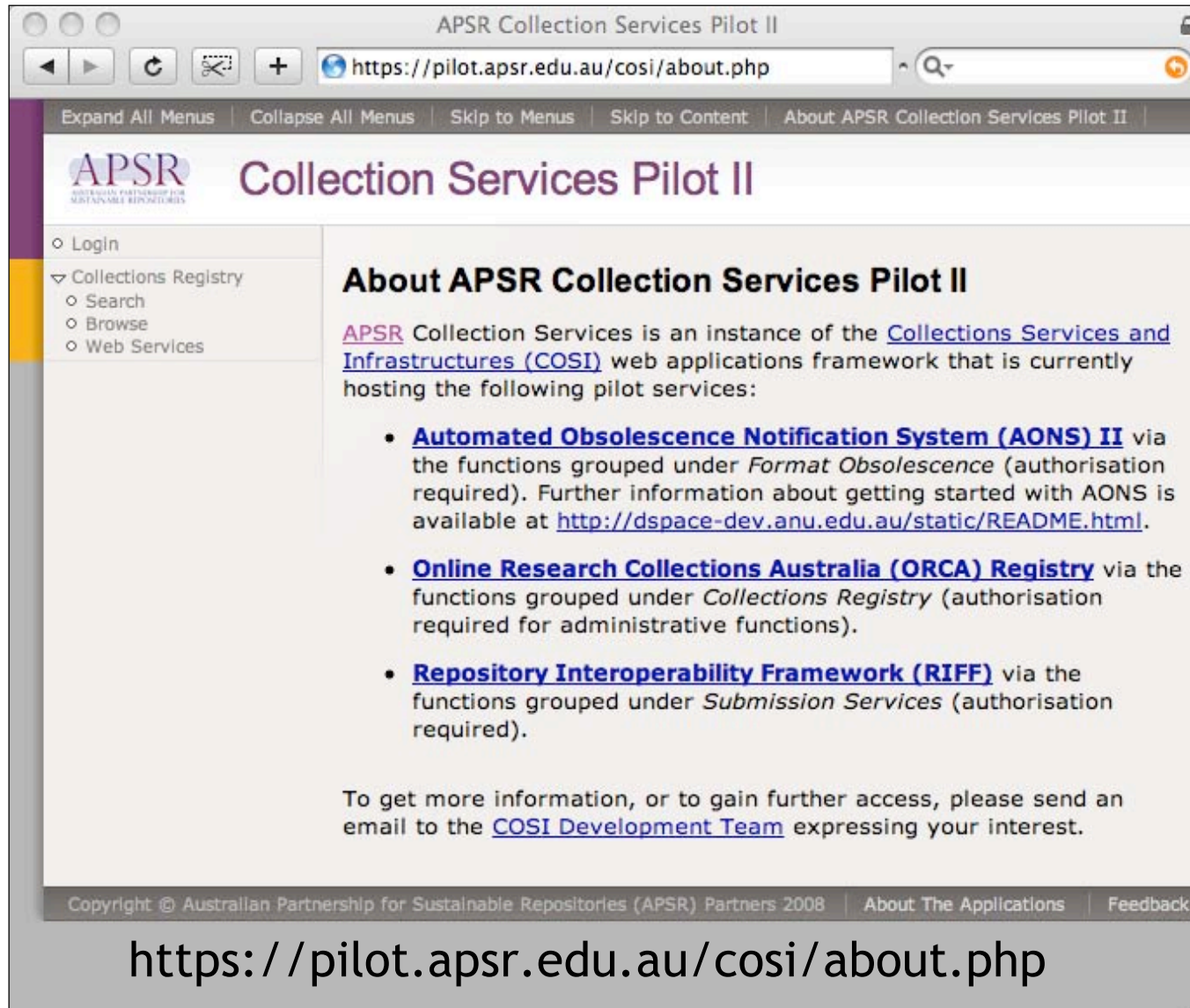
# Presentation Outline

1. Live presentation of the ORCA (Online Research Collections Australia) Registry
2. Discussion of the role and benefits of registries in the context of *federated* networks of repositories and mass data storage facilities
3. Current progress in developing global approaches to implementing registries

# ORCA Registry

- It's a pilot not a production service
- It's not a discovery service (it's for machine-to-machine transactions)
- When people interact, its repository/collections managers and upstream metadata aggregators
- All metadata records are automatically harvested
- It harvests collection-level metadata records, not item-level records
- It provides no capacity for record owners to edit records (a feature not a bug)
- It's pointless unless part of a federated architecture of repository services
- It works on the principle of 'register locally, discover globally'

# ORCA Registry Screen (I)



The screenshot shows a web browser window titled "APSR Collection Services Pilot II" with the URL <https://pilot.apsr.edu.au/cosi/about.php>. The page features a navigation menu with options like "Expand All Menus", "Collapse All Menus", "Skip to Menus", "Skip to Content", and "About APSR Collection Services Pilot II". The main content area is titled "Collection Services Pilot II" and includes a sidebar with a "Collections Registry" menu. The main text describes the pilot services, including the Automated Obsolescence Notification System (AONS) II, the Online Research Collections Australia (ORCA) Registry, and the Repository Interoperability Framework (RIFF). A footer contains copyright information and links to "About The Applications" and "Feedback".

APSR Collection Services Pilot II

Expand All Menus | Collapse All Menus | Skip to Menus | Skip to Content | About APSR Collection Services Pilot II

APSR  
AUSTRALIAN PARTNERSHIP FOR SUSTAINABLE REPOSITORIES

Collection Services Pilot II

o Login

o Collections Registry

- o Search
- o Browse
- o Web Services

## About APSR Collection Services Pilot II

APSR Collection Services is an instance of the [Collections Services and Infrastructures \(COSI\)](#) web applications framework that is currently hosting the following pilot services:

- [Automated Obsolescence Notification System \(AONS\) II](#) via the functions grouped under *Format Obsolescence* (authorisation required). Further information about getting started with AONS is available at <http://dspace-dev.anu.edu.au/static/README.html>.
- [Online Research Collections Australia \(ORCA\) Registry](#) via the functions grouped under *Collections Registry* (authorisation required for administrative functions).
- [Repository Interoperability Framework \(RIFF\)](#) via the functions grouped under *Submission Services* (authorisation required).

To get more information, or to gain further access, please send an email to the [COSI Development Team](#) expressing your interest.

Copyright © Australian Partnership for Sustainable Repositories (APSR) Partners 2008 | [About The Applications](#) | [Feedback](#)

<https://pilot.apsr.edu.au/cosi/about.php>



# ORCA Registry Screen (II)

The screenshot shows a web browser window titled "APSR Collection Services Pilot II". The address bar contains the URL "https://pilot.apsr.edu.au/cosi/orca/browse.php" and a Google search bar. Below the browser window, the website header includes the APSR logo (Australian Partnership for Sustainable Repositories) and the title "Collection Services Pilot II". A navigation menu on the left lists "Login", "Collections Registry" (expanded to show "Search", "Browse", and "Web Services"), and "Web Services". The main content area is titled "Browse the Collections Registry" and features four links: "Collections", "Services", "Parties", and "Activities". The footer contains copyright information for APSR Partners 2008 and links for "About The Applications" and "Feedback".

# ORCA Registry Screen (III)

APSR Collection Services Pilot II

Expand All Menus | Collapse All Menus | Skip to Menus | Skip to Content | About APSR Collection Services Pilot II

APSR AUSTRALIAN PARTNERSHIP FOR SUSTAINABLE REPOSITORIES

Collection Services Pilot II

- Login
- Collections Registry
  - Search
  - Browse**
  - Web Services

### Browse the Collections Registry

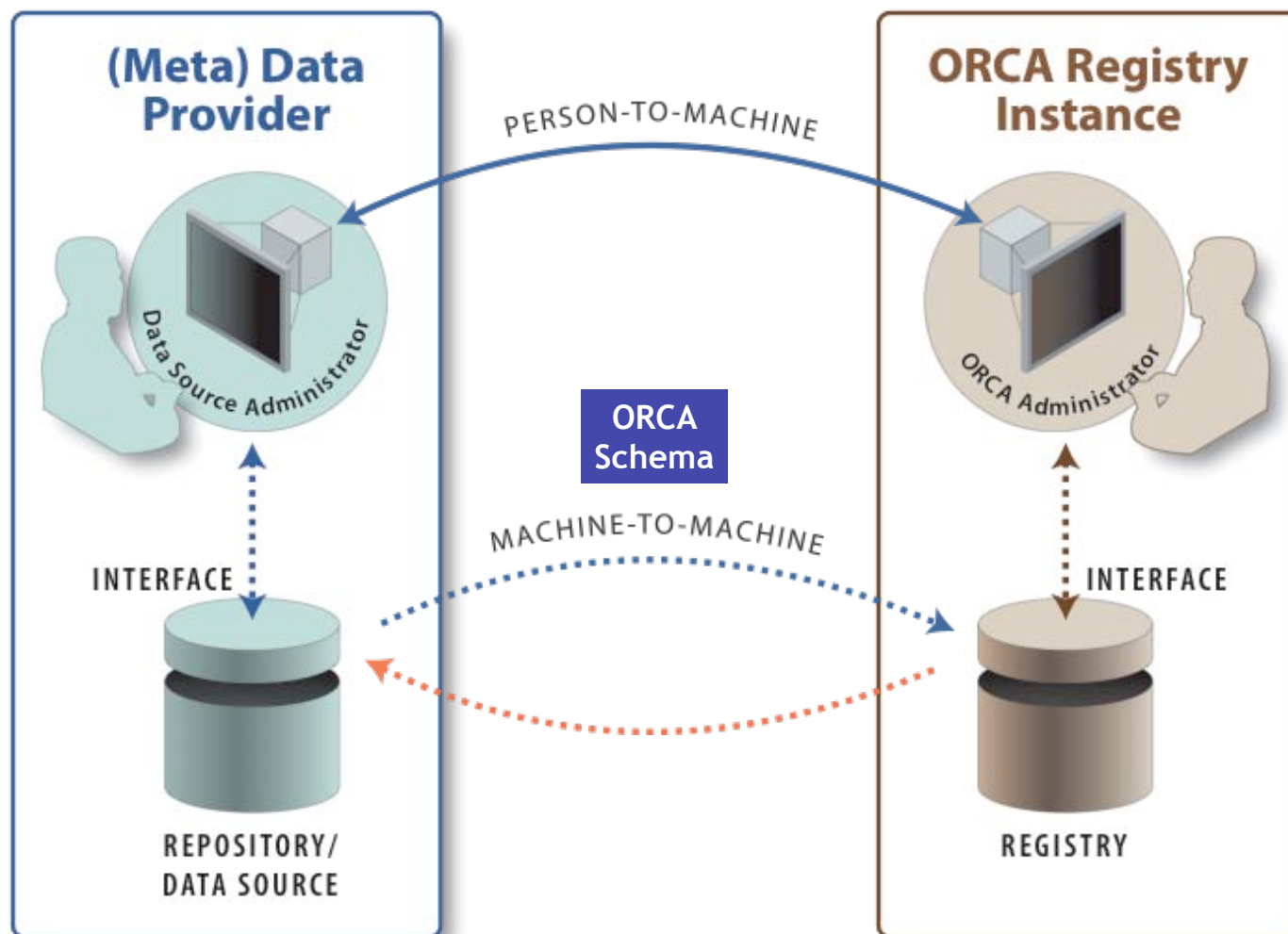
[Collections](#) [Services](#) [Parties](#) [Activities](#)

A B C D E **F** G H I J K L M N O P Q R S T U V W X Y Z

Results 1 to 10 of 31 (0.176 seconds).

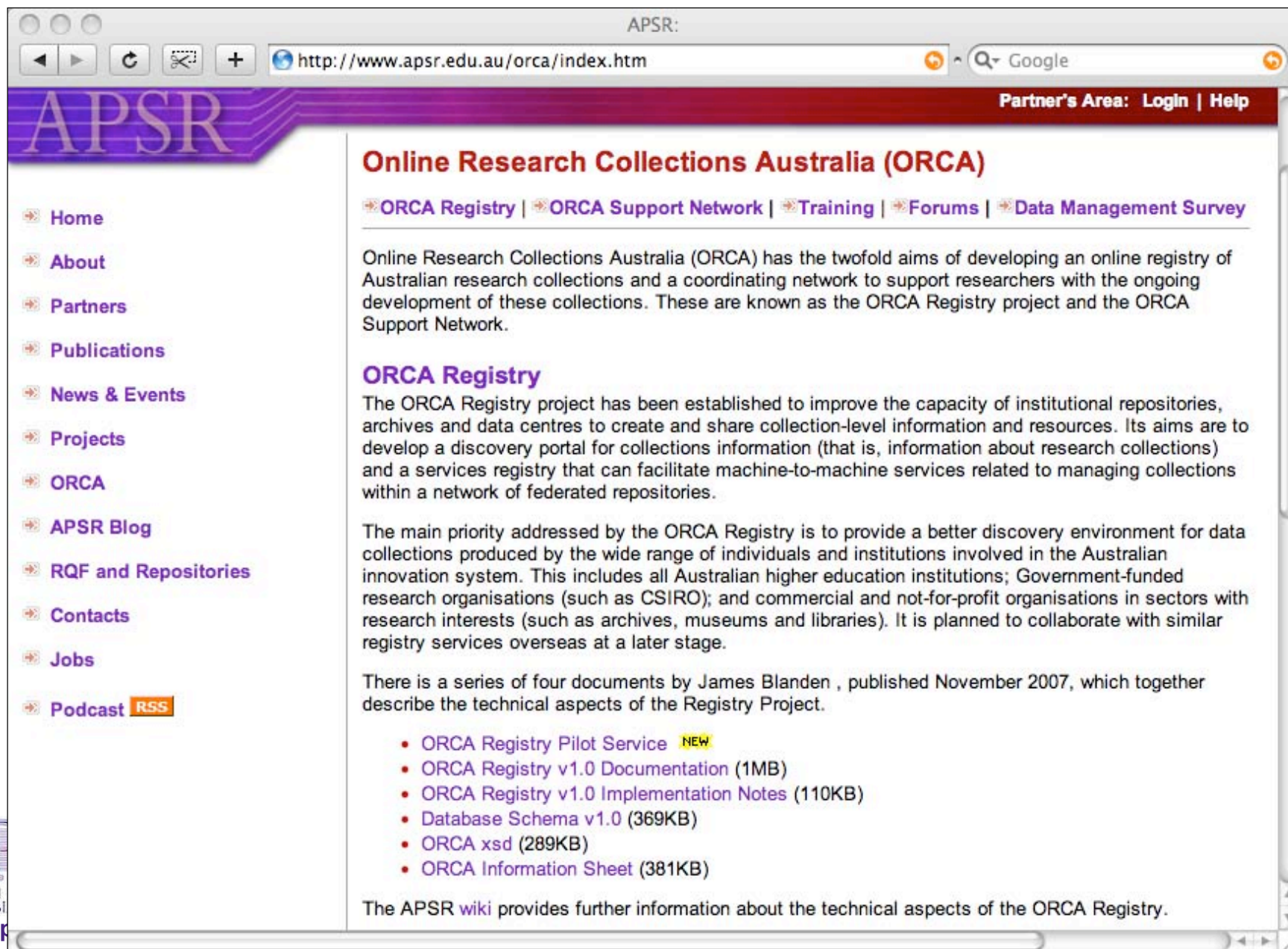
[First Step in a Virtual Observatory; Calibrating the MaCHO Images](#)  
**Collection—Key:** au.edu.anu.anusf.x60  
**Source:** ANU Supercomputer Facility  
**Relations:** Has Principle Investigator: [Schmidt · B P](#)  
**Subjects:** RFGD: 240101  
**Description:** The Recently Destroyed Great Melbourne Telescope Acquired nearly 10 Terabytes of data from 1992-2002, with the data currently not easily accessible to the wider community. In an effort to set up a node of a virtual observatory, we need to calibrate the positions of this dataset,...

# Online Research Collections Australia (ORCA) Registry





# For more ORCA info...



APSR: <http://www.apsr.edu.au/orca/index.htm> Google

**APSR** Partner's Area: [Login](#) | [Help](#)

- Home
- About
- Partners
- Publications
- News & Events
- Projects
- ORCA
- APSR Blog
- RQF and Repositories
- Contacts
- Jobs
- Podcast [RSS](#)

## Online Research Collections Australia (ORCA)

[ORCA Registry](#) | [ORCA Support Network](#) | [Training](#) | [Forums](#) | [Data Management Survey](#)

Online Research Collections Australia (ORCA) has the twofold aims of developing an online registry of Australian research collections and a coordinating network to support researchers with the ongoing development of these collections. These are known as the ORCA Registry project and the ORCA Support Network.

### ORCA Registry

The ORCA Registry project has been established to improve the capacity of institutional repositories, archives and data centres to create and share collection-level information and resources. Its aims are to develop a discovery portal for collections information (that is, information about research collections) and a services registry that can facilitate machine-to-machine services related to managing collections within a network of federated repositories.

The main priority addressed by the ORCA Registry is to provide a better discovery environment for data collections produced by the wide range of individuals and institutions involved in the Australian innovation system. This includes all Australian higher education institutions; Government-funded research organisations (such as CSIRO); and commercial and not-for-profit organisations in sectors with research interests (such as archives, museums and libraries). It is planned to collaborate with similar registry services overseas at a later stage.

There is a series of four documents by James Blanden , published November 2007, which together describe the technical aspects of the Registry Project.

- ORCA Registry Pilot Service **NEW**
- ORCA Registry v1.0 Documentation (1MB)
- ORCA Registry v1.0 Implementation Notes (110KB)
- Database Schema v1.0 (369KB)
- ORCA xsd (289KB)
- ORCA Information Sheet (381KB)

The APSR [wiki](#) provides further information about the technical aspects of the ORCA Registry.

**AI**  
AUSTRALIAN  
SUSTAINABLE  
[www.apsr.edu.au](http://www.apsr.edu.au)

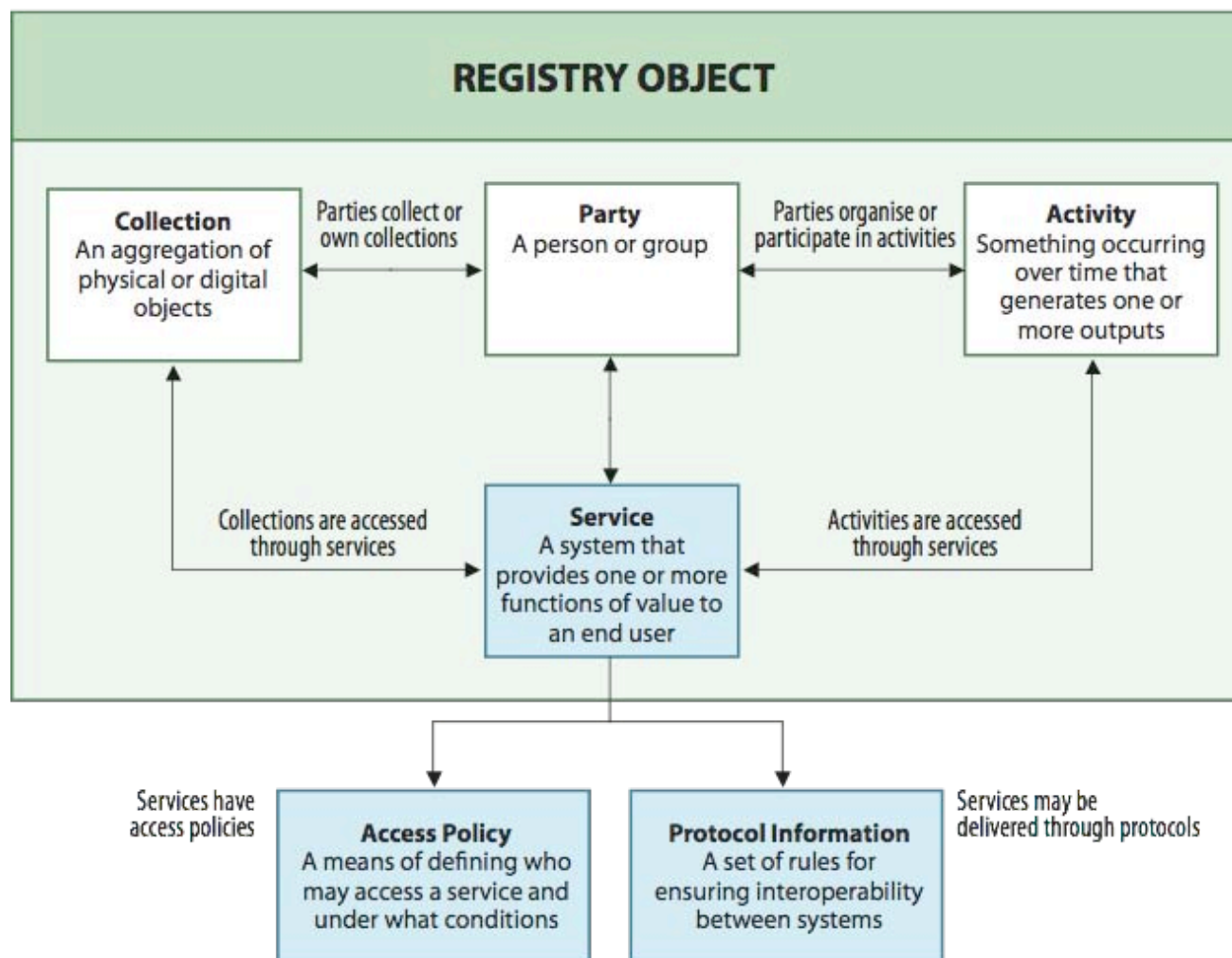
## ORCA Registry technical details (I)

- PHP/PostgreSQL Web application (Apache lic.)
- Part of our own Web application framework (COSI- Collections and Services Interface)
  - Supports REST-type service calls and messaging (i.e. not WS-I standards and protocols)
- An ORCA Registry instance can be configured for local and/or a national/global network topology
  - ORCA Registry instances have REST service interfaces for metadata harvesting and providing

## ORCA Registry technical details (II)

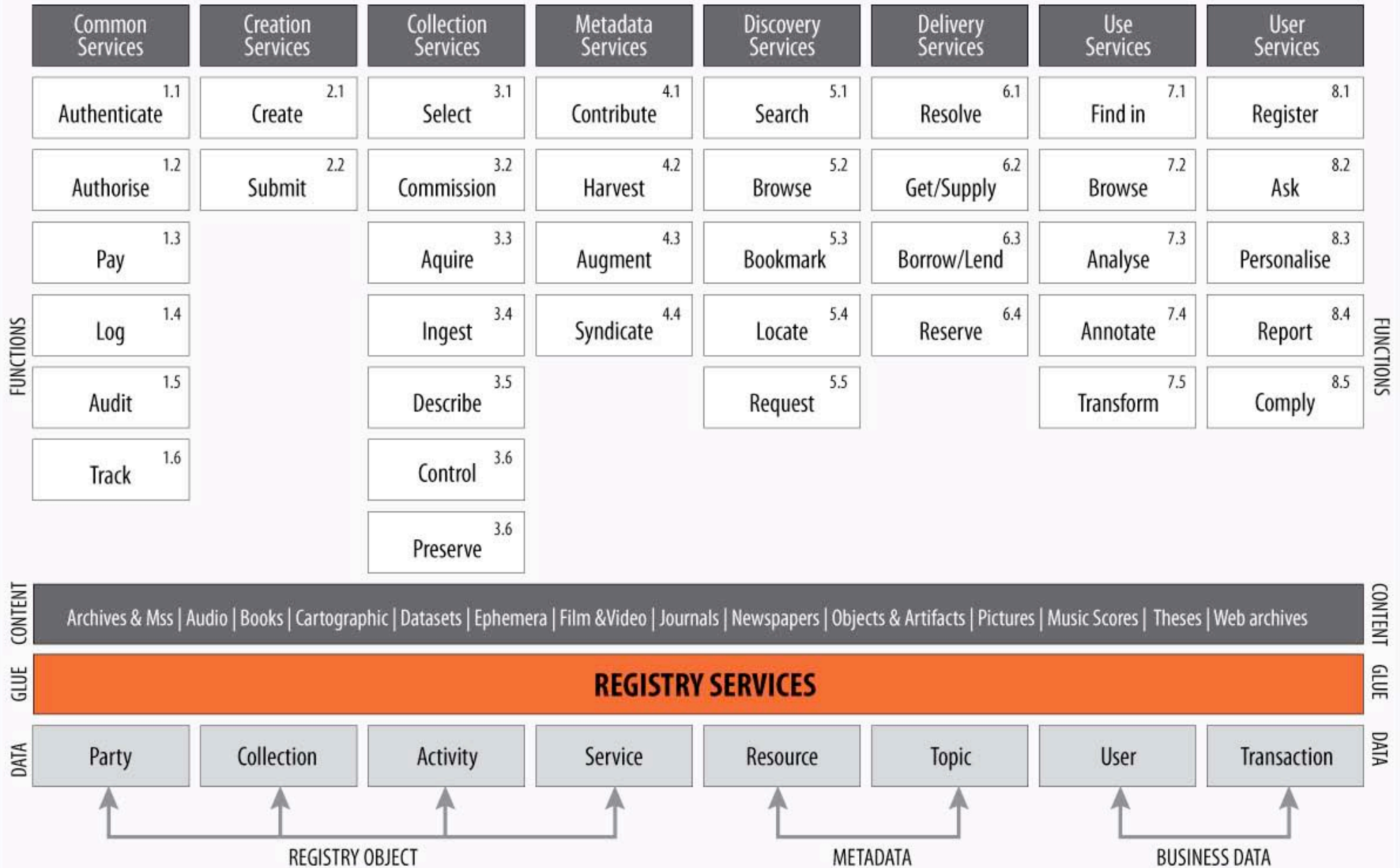
- First implementation of the draft ISO2146 (Registry Services for Libraries and Related Organizations)
- ORCA Registry uses ISO2146 XML binding for the basis of the DBMS data model and the metadata interchange profile (expressed as a .xsd schema)
  - Working on a ORCA Schema profile for OAI-PMH (June 2008)
- ISO2146 is compatible with the Dublin Core Collections Application Profile
- Interoperable with IESR and OCKHAM metadata
- Latest draft has input from APSR to align with current research information systems (i.e. CERIF compatibility)
- Credits: James Blanden, Scott Yeadon, Judith Pearce, Adrian Burton and Chris Blackall

# ISO2146 Registry Services for Libraries and Related Organizations





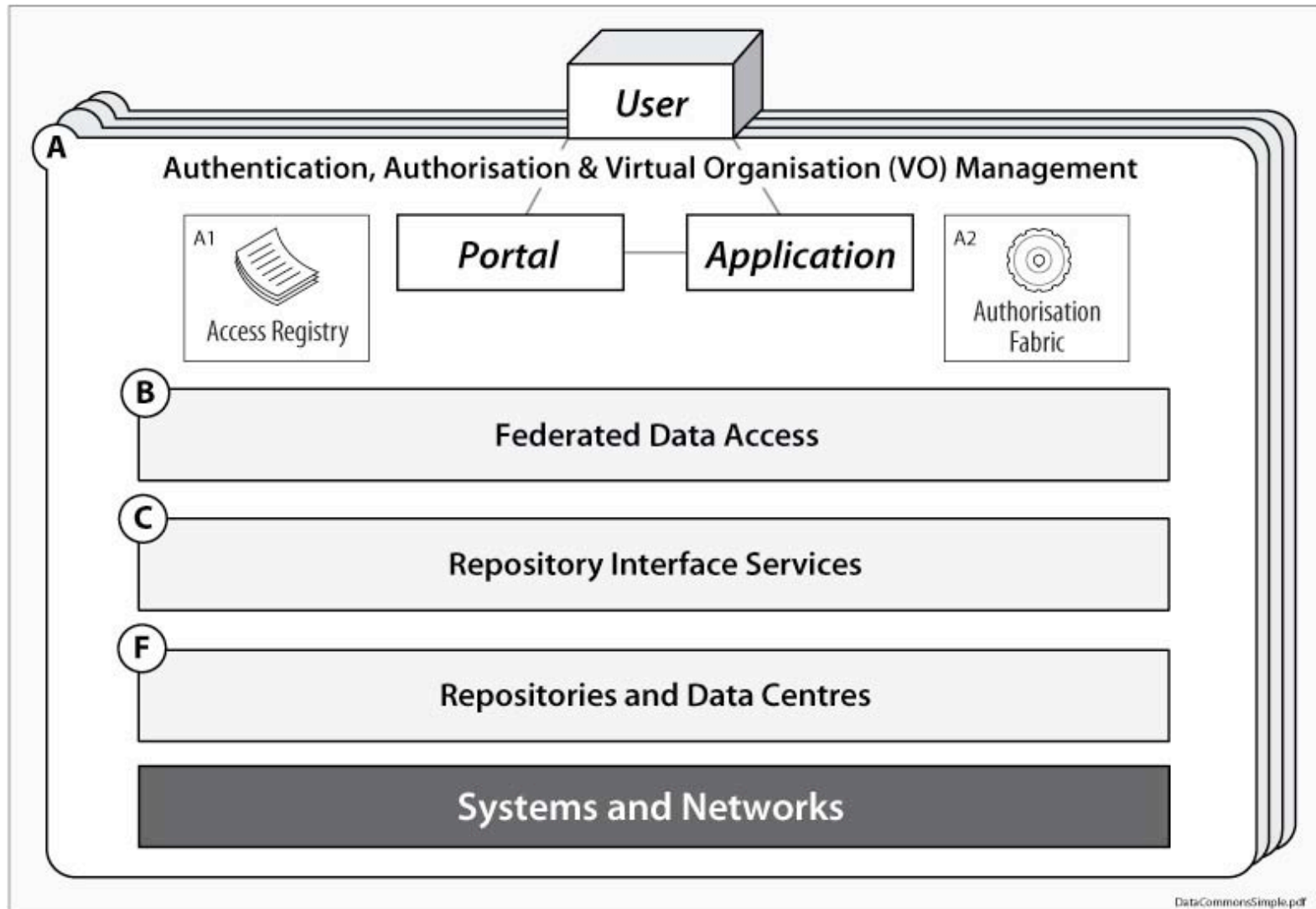
# DIGITAL LIBRARY SERVICE FRAMEWORK



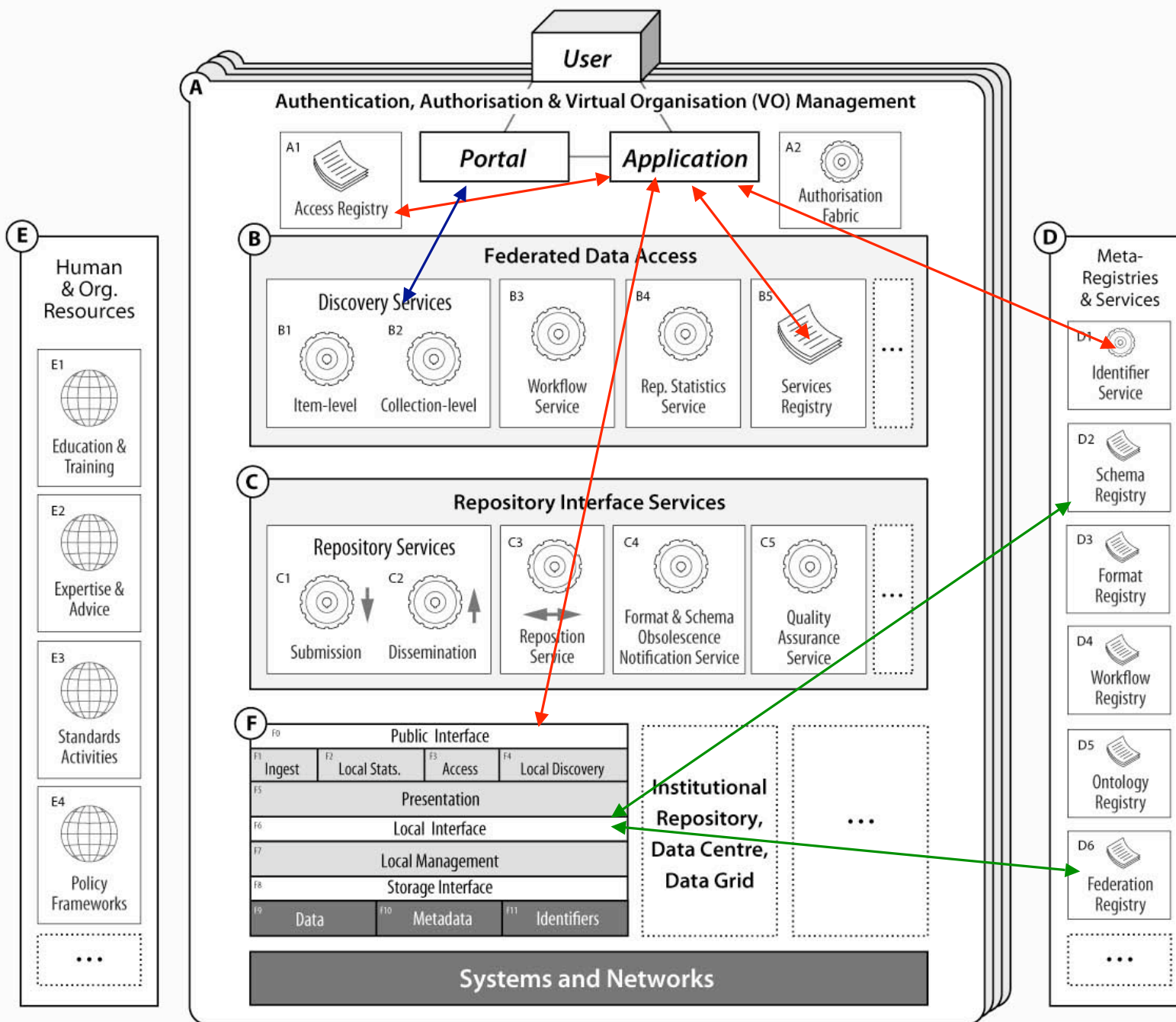
[DLSF\_services.pdf]-[v.1]-[12/02/08]-[Chris Blackall]-[Redrawn from Digital Library Service Framework, National Library of Australia, 8 Feb 2008, ver 0.3, p.4]



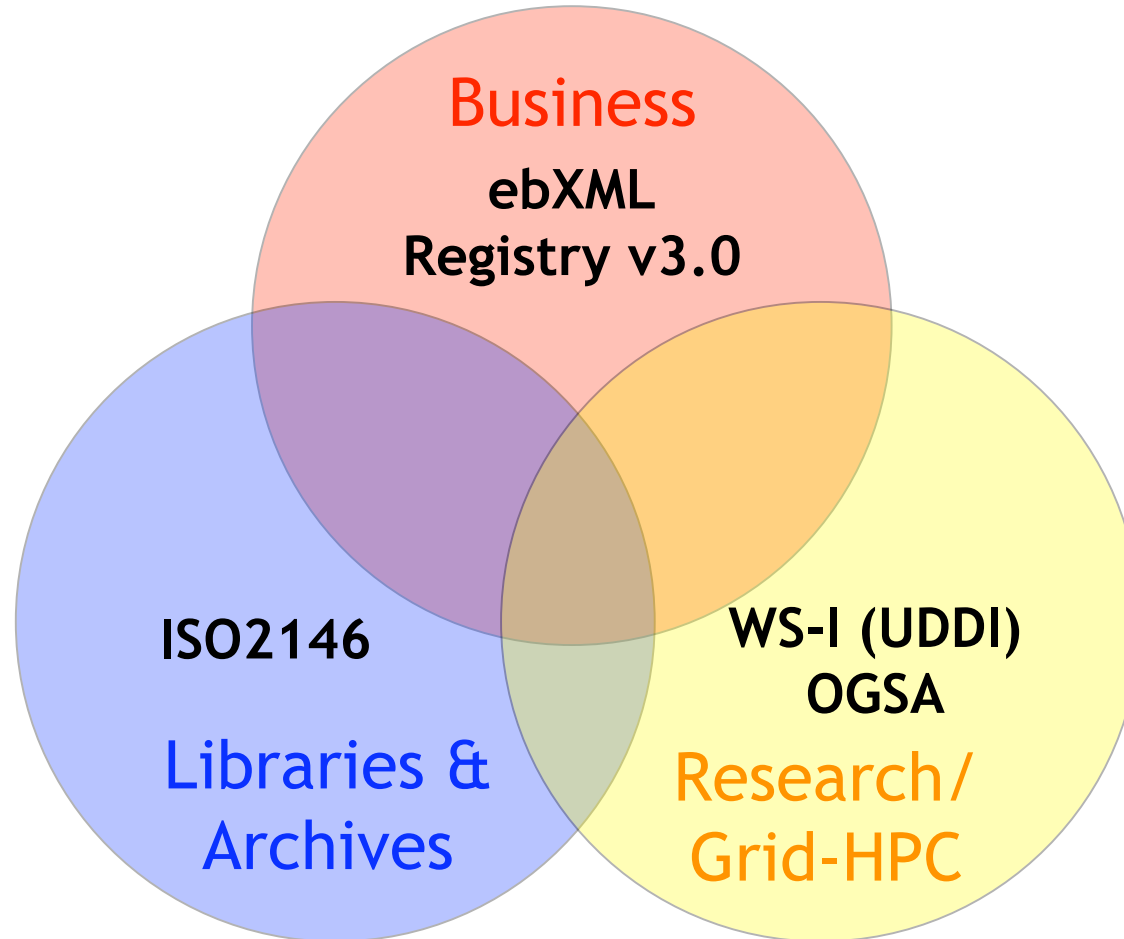
# Australian Data Commons Architecture (I)



# AUSTRALIAN DATA COMMONS ARCHITECTURE



# Registry technology/standards options



WS-I (UDDI): Web Services-Interoperability (Universal Description, Discovery and Integration)  
OGSA: Open Grid Services Architecture

## Andy Powell (eduserv foundation)

Instead of developing *institutional* repositories, “we need to focus on building and/or using global scholarly social networks based on global repository services.”

Andy Powell, *Repositories thru the looking glass*, 2008.

<http://efoundations.typepad.com/efoundations/2008/02/repositories-th.html>

## Lorcan Dempsey, OCLC

“Registry services will need to get more common if we are going to have efficient interactions within networks of library providers and consumers”

Lorcan Dempsey. *Registries: the intelligence in the network*. August 20, 2006.

<http://orweblog.oclc.org/archives/001105.html>



# What are Global Repository Services? (I)

- Global repository services are **not** equivalent to networked repositories
  - IRs are typically implemented as enterprise applications that are focused on institutional needs rather than researcher needs
  - Low levels of interoperability between leading repository platforms
  - Service interfaces are available in IRs, but little in the way of shared services infrastructure
  - Commercial Web 2.0/Cloud Computing providers offering services

# What are Global Repository Services? (II)

- Federated Networks of Repositories (and Mass Data Storage Facilities)
  - **Shared authentication and authorization services** (that enable users to log on to anywhere within the Federation and access Federation content)
  - **Shared discovery services** (that enable users to discover content unique to the Federation)
  - **Shared content creation and workflow services** (that enable users to create new content and content metadata that adds extra value to the Federation)
  - **Shared current research and statistics services** about research outputs, impact and quality (that enable users to make strategic decisions about current and future research)
  - ...

# Caveats

- In order the benefits to be achieved, we (repository developers) must develop systems that:
  - enable users of low to moderate IT skills to easily submit new content into repositories/data centers
  - reduce the time and effort required for users to capture and describe content metadata (preferably within the application that the content was originally created)
  - provide usage feedback services, such of usage statistics, impact measures and so on
  - Provide user-friendly metadata editing services to improve the quality of metadata records over time.

# Federated Networks of Repositories

- “Towards an Australian Data Commons” (2007)
  - Australian Access Federation (AAF)
  - Australian National Data Service (ANDS)
  - Australian Research Collaboration Services (ARCS)
  - National Computing Infrastructure Services (NCIS)
- DRIVER (European Union)
- JISC Information Environment (UK)
  - Shared Infrastructure Services (SIS) Projects (2008)
  - Information Environment Services Registry (IESR)
- Datanet Projects (National Science Foundation-USA 2008)
- Global Registries Initiative (APSR, IESR, OCKHAM and others)

“The simpler the user-experience, the greater the scope, complexity and abstraction of the underlying information infrastructures”

**ATM Machine  
(3D Model)**

