BUILDING AN AUSTRALIAN USER COMMUNITY FOR VIVO

Presenters:
Robyn Rebollo
eResearch Senior Specialist
(Information Management)
Griffith University
r.rebollo@griffith.edu.au

Simon Porter
Information Manager, Research
University of Melbourne
simon.porter@unimelb.edu.au

Lance DeVine
Research Support Specialist
Queensland University of Technology
l.devine@qut.edu.au
BUILDING AN AUSTRALIAN USER COMMUNITY FOR VIVO

- VIVO Adopters Project Team Members: (not exhaustive..)
  - Malcolm Wolski, Project Stakeholder, Griffith Uni
  - Jo Morris, Project Mgr, Griffith Uni
  - Mark Fallu, Architecture, Griffith Uni
  - Gerhard Weis, Program Developer & Ontology, Griffith Uni
  - Arve Solland, Web Developer, Griffith Uni
  - Robyn Rebollo, Ontology and Metadata Dev, Griffith Uni
  - Lance DeVine, Architecture, Workflow, BA, QUT
  - Joe Young, Project Stakeholder, QUT
  - Simon Porter, Project Mgr / Architecture, Uni Melb
  - Cyrus Keong, Program Developer, Versi
About Us: Griffith University

- **Griffith University:**
  - Established in 1971 and officially opened in 1975.
  - We are home to about 40,000 students from all over the world.
  - Our 5 campuses span 3 cities
  - between Brisbane and the Gold Coast, each with distinct areas of teaching and research strength.
  - Our five campus locations: Gold Coast, Logan, Mt Gravatt, Nathan and South Bank.
Our research experts work in over 30 research centres developing new knowledge in the areas of medicine and healthcare, emerging technologies, social innovations, culture, learning and the arts, the environment, and governance and policy development.

Our research centres include externally supported research centres and collaborations with external research institutions.
ABOUT US: UNIVERSITY OF MELBOURNE

- University of Melbourne:
  - Established 1853
  - 11 Faculties/Graduate Schools
  - 45,569 Students (Aug 2009)
  - 7,326 Staff
ABOUT US: UNIVERSITY OF MELBOURNE

- Research: Australia’s leading research institution
- 2010 audited research income of $337 million
- National leader in publication outputs

Located within the Parkville Precinct: a world class biomedical research environment

Institution Collaboration Pattern: 2008 publications (Thomson ISI)
ABOUT US: QUEENSLAND UNIVERSITY OF TECHNOLOGY

- Queensland University of Technology:

QUT ...
40,000 students
1,700 researchers
Inner-city Brisbane, Queensland, Australia
What Brought Us Here

ANDS

“More Australian researchers reusing research data more often.”

The Australian National Data Service (ANDS) aims to:

- influence national policy in the area of data management in the Australian research community
- inform best practice for the curation of data
- transform the disparate collections of research data around Australia into a cohesive collection of research resources
Australian National Data Service

- Seeding the Commons – Projects that create infrastructure within institutions to collect and transform metadata about collections, then publish it to the Australian Research Data Commons (ARDC).

- Data Capture - Projects that create infrastructure within institutions to collect and manage data, and to improve the way metadata about it is managed.

- Metadata Stores (consolidating metadata about research data)

  - Uni Melbourne, QUT and Griffith all received funding from ANDS to participate in Seeding the Commons and building infrastructure for metadata capture at each university.

More information: [List of institutions with ANDS-funded ARDC projects]
INTERNAL DRIVERS

- Research Data Audit – Interviewing researchers (included a comprehensive questionnaire) helped us identify the kinds of research data that was out there, and whether it was available for discovery purposes. (Griffith / QUT)

- Profiling Researchers – The need to publically profile researchers via one gateway was evident after the NCRIS project. Also, there currently is no single staff profile system available that includes linked data between researchers, projects, collections and research publications. (Griffith)
**INTERNAL DRIVERS CONTINUEI**

- Research Data & Records Compliance (Uni Melb)

- To support the work of Seeding the Commons (Griffith / QUT) – Landing pages to display Research Data Australia records from each institution was required by ANDS.
Choosing a Technology: Vitro

Standards Based Semantic Web

Strategic Direction Strongly Aligned with ANDS agenda
RESEARCH ACTIVITY HUB

Griffith UNIVERSITY

QUT
The Education Investment Fund (EIF) project objective:

- The purpose of the data exchange hub is to collect appropriate metadata from research collections (at the content metadata level where possible) within the University with customised (and automated where possible) feeds from the various University content management systems. This hub will then act as a central University repository to feed information in a standard format to the Australian Research Data Commons, and university library discovery tools and other research federations where appropriate.

- The overall objective is to develop a sustainable solution to automate the collation of new research data held within the University and to populate Research Data Australia.

- Joint project implementation by QUT and Griffith Uni.
RESEARCH ACTIVITY HUB

- VIVO in Action! - Home Page
Research Activity Hub

- VIVO in Action! - Researchers in the Hub
RESEARCH ACTIVITY HUB

- VIVO in Action! - Researcher Profile
Research Activity Hub

- VIVO in Action! - Groups in the Hub

Groups

A group describes an organisational unit at Griffith University that has a relation to a collection, project, service or researcher in the Hub. Groups represented in the Hub include Griffith faculties, schools, centres and research centres.

Search groups
Group name search:

- Atmospheric Environment Research Centre
- Australian Centre for Intellectual Property in Agriculture
- Australian Institute for Suicide Research and Prevention
- Australian Rivers Institute

Total of 4 listings starting with A. Back to top

Total Listings: 73
**Research Activity Hub**

- **VIVO in Action! - Group Profile**

![VIVO in Action! - Group Profile](image-url)
VIVO in Action! - Projects in the Hub

Projects

A project is any piece of research work that is undertaken or attempted, with a start and end date and defined objectives. Projects represented in the Hub include, but are not limited to, Australian Research Council, National Health and Medical Research Council and other federal or state funded grants.

Search projects
Project name search:

- "Learning through (e-)sharing: a corpus-based approach to the pragmatics of English" Creating the Griffith Corpus of Spoken Australian English (GrCS Aust)
- "Understanding the mechanism of Wolbachia mediated antiviral protection"
- 2x Basler digital (DCAM) colour cameras (100 frames per second at 600x492 resolution) with lens, cables and tripod for use with existing Vicon motion capture system
- A Colonial and Conceptual History of Asymmetric Warfare and Security
- A Comparative Analysis of Gambling Regulation in Australian States and Territories
- A Comparative Study of an Education Intervention to Promote Family Involvement in Residential Dementia Care
- A Comparison of Anticipated Benefits and the Personal, Educational and Social Outcomes of Cochlear Implantation in Deaf Children
- A Cultural History of Information: Lessons from the Enlightenment
- A Focussed Ion Beam Scanning Electron Microscope for Advanced Analytical and Nanotechnology Research
- A High Resolution Analytical Scanning Electron Microscope for South-East Queensland
- A High-Throughput Neutron Spectrometer for the Study of Atomic and Molecular Motion at ANSTO
- A National Tray Facility
**Research Activity Hub**

- **VIVO in Action! - Project Record**

  ![Project Record Screenshot]

**PPBio Australasia - Karavatha**

**Details**

- **Project Name**: PPBio Australasia - Karavatha
- **Short Description**: 2010-07
- **Start Year-Month**: 2010-07
- **End Year-Month**: 2010-07
- **Work Email**: m-koz@griffith.edu.au
- **Location**: 170 Kassada Road

**Subject Areas**

- **Field Of Research**
  - Ecological Applications
  - Environmental Science and Management

**Managed by**

- **Mars, Jean-Marc**

**Relations**

- **Griffith School of Environment | Group**

**Participants**

- **Cassley, James Guy**

**Output**

- **Karavatha Forest Park - Terrestrial Pools**
**Research Activity Hub**

- **VIVO in Action!** - Research Collections in the Hub

---

**Collections**

A collection is a grouping of physical or data items that have some shared significance. Collections represented in the Hub are research focused and include datasets, collections, repositories or catalogue/indexes.

- Australian Institute for Suicide Research and Prevention Collection
- Australian Rivers Institute Collection
- Clinical and Community Practice Innovation Data Collection
- Discussion Paper Series in Economics and Business Statistics
- Discussion Paper Series in Finance
- Discussion Papers Series in Accounting and Business Law
- Griffith Corpus of Spoken Australian English (GCSAuse) Collection
- Griffith Research Online (GRO)
- Karawatha Forest Park - Terrestrial Plots
- Queensland Collection of Art Research Collection
- Research Data Management Centre
- Research Survey Centre Data Collection

Total: 12
RESEARCH ACTIVITY HUB

• VIVO in Action! - Research Collection Record

Karawatha Forest Park - Terrestrial Plots

Collection Name: Karawatha Forest Park - Terrestrial Plots
Collection Type: Dataset
Short Description: PESIO UTER program: Terrestrial plot data for Karawatha Forest Park, southeast Queensland. PESIO research grid with 33 plots has been established in Karawatha Forest Park, Brisbane. Data collected includes: (i) mesoscale variation of forest and fauna communities in response to factors such as relief, topography and fire history, and (ii) associations between fauna species composition and vegetation.

Collection Rights: Access to this dataset is supplied on condition that the principal investigators responsible for collecting data in the dataset are credited in any publications that use the data. It is recommended that persons interested in using the data contact the collection owner.

Webpage: http://lom.effects.griffith.edu.au/research/items/38927d06-7f19-1963-34d8-9204c23e650d

Coverage: [Map of Karawatha Forest Park]

Subject Areas
Field Of Research:
- Ecological Applications
- Environmental Science and Management

Managed by:
- Here-Jean Marc
- Environmental Futures Centre

9/9/2010 VIVO Conference August 2010
Visualizing Collaborations using Javascript InfoVis Toolkit

Connected Sunburst

Dynamic SPARQL query is used as an input for this visualization.

This example shows how researchers within a group collaborates through projects.

Left click to select a node (researcher) and show its relations.
Multiple goals for the establishment of a research data registry

Goals:

- Participation in the national research data commons
- The need to play a key role in helping the University communicate its research identity both internally and externally
- The need to enable university research data and records policy compliance
BACKGROUND: THE NEED FOR A RESEARCH DATA REGISTRY

- **Policy – 2005, review 2009**
  - Good start but not enough

- **Audit – 2007, 2009**
  - Policy and compliance is off the radar for most
  - Compliance – patchy
  - Storage – ad hoc
  - Local registers
  - Risk management
ENGAGING WITH RESEARCHERS

Central resources are the missing key to enabling compliance

- **Central storage** – a petabyte-scale Research Data Centre for Parkville Precinct

- A Central research data registry
WHAT IS A CENTRAL RESEARCH DATA REGISTRY?

Definition: This is an institution-wide record of all research data and records stored in the university.

At a minimum:

- description of the research data and records, associated researchers and projects,
- the location of the data (digital and analogue),
- access restrictions,
- and relocation and disposal schedules.

(requirements based on policy)
**Scale**

- Intended scale (biggish) – definitely enterprise scale

- Over time, it should be as important to the University as publications reporting and grants management
A central research data registry happens in the context of richly described research:

- Grants, Publications, Research Program, External and Internal Researchers, and departments
- All of this information exists and should be reused in the registry...
- Should not be an information silo
B U I L D I N G  T H E  R E G I S T R Y

- Piggybacking on our public researcher profile system - ‘find an expert’ dataset...

http://www.findanexpert.unimelb.edu.au/
DATA ENTRY PROCESS

(1) Pre seed Data Registry with Find an Expert data

(2) Scan for references to data in publications

(3) Create Skeleton Research Data Record

(4) Infer as many of the details about the research data record from the publication
research data for a consumer-oriented technique for planned residential developments

The empirical study is based on innovation theory where end-users are primary stakeholders in the innovation process, and even innovators themselves. A case study of a new participation method based on a two-phase Internet questionnaire is used to research practical solutions in integrating end-users into urban planning process.

Research Descriptions

results published in
A consumer-oriented technique for planned residential developments

Custodians

Associated Principal Investigator
HEYWOOD, CHRISTOPHER (DR)

Custodian Department
Architecture, Building and Planning
Feedback so far...

- Feedback from the Faculty (so far)
STANDARDS TECHNOLOGY AND DEVELOPMENT
STANDARDS

- **RIF-CS**
  - The Registry Interchange Format - Collections and Services (RIF-CS) Schema was developed as a data interchange format for supporting the submission of metadata to a collections service registry.
  - It is based on ISO2146 but only includes elements needed for a collection service registry and so is not a full binding to the standard.

- **OAI-PMH**
  - The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability. **Data Providers** are repositories that expose structured metadata via OAI-PMH. **Service Providers** then make OAI-PMH service requests to harvest that metadata.
ISO2146

Registry Object

Collection
an aggregation of physical or digital objects

Parties collect or own collections

Party
a person or group

Parties organise or participate in activities

Activity
Something occurring over time that generates one or more outputs

Service
a system that provides one or more functions of value to an end user

Collections are accessed through services

Activities are accessed through services

Access policy
a means of defining who may access a service and under what conditions

Protocol Information
a set of rules for ensuring interoperability between systems

Services have access policies

Services may be delivered through protocols
Registry Interchange Format - Collections and Services (RIF-CS)

Registry Object

- Collection
  - Catalogue or Index
  - Collection
  - Registry
  - Repository
  - Dataset
- Party
  - Group
  - Person
- Activity
  - Award
  - Course
  - Event
  - Program
  - Project
- Service
  - Harvest – OAI-PMH
  - Search – Z3950
  - Syndicate – RSS ...

## RIF-CS: A Typical Object

<table>
<thead>
<tr>
<th>Type</th>
<th>A value taken from a controlled vocabulary indicating the type of object being described.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier</td>
<td>Primary and alternative identifiers for the object.</td>
</tr>
<tr>
<td>Name</td>
<td>The name of the object in either a simple or compound form.</td>
</tr>
<tr>
<td>Description</td>
<td>A textual description or URI resolving to a description relevant to the object.</td>
</tr>
<tr>
<td>Subjects</td>
<td>A subject category into which the object falls or the object is related.</td>
</tr>
<tr>
<td>Location</td>
<td>Location(s) relevant to the collection. A location element should contain information about a single location (e.g. collection website).</td>
</tr>
<tr>
<td>Related Objects</td>
<td>Information about a related registry object.</td>
</tr>
<tr>
<td>Related Info</td>
<td>A URI pointing to information related to the collection.</td>
</tr>
</tbody>
</table>
RIF-CS – LINKED DATA IN DISGUISE

Collection B
Type: Collection

Collection A
Type: Dataset

Is Part Of
Is Part Of

Is Owner Of
Is Managed By

Is Participant In
Is Participant In

Has Collector

Party:
Type: Person
Name: Joe Blogs

Party:
Type: Group
Name: Institute of Sustainable Resources

Party:
Type: Group
Name: Queensland State Government

Activity:
Type: Project
Name: Solar Power Generation

Collection C
Type: Dataset

Is Part Of
Is Part Of

Is Owner Of
Is Managed By

Is Participant In
Is Participant In

Is Funded By
Is Manager Of

Is Participant In
Is Participant In

Is Owner Of
Is Managed By

Is Participant In
FROM COLLECTIONS REGISTRY TO DISCOVERY
Research Data Australia Web Site

Home

Records by Classification
- Collections (1122)
  Where a collection is a useful grouping of physical or digital items.
- Parties (343)
  Where a party is a person or organisation that has some relationship to a collection, service, activity, or party.
- Services (2)
  Where a service is a mechanism for gaining some kind of access to or information about a collection (or items within a collection).
- Activities (13)
  Where an activity is an undertaking or process related to the creation, update, or maintenance of a collection.

Records by Grouping
- Australian Institute of Marine Science (455)
- iVEC MEST - Western Australian Marine Data and Projects (898)
- Monash University (10)
- Polar Information Commons (32)
- Publish My Data (51)
- Queensland University of Technology (34)
**How to Store RIF-CS in VIVO?**

- Try to use as much of the existing VIVO ontology and associated ontologies as possible.

- However, need to extend the Core VIVO Ontology
  - Admin extensions – new properties
  - Vocab extensions – new classes
  - Relationship extensions – object properties
  - Classifications – new separate ontologies
ANDSHarvest Ontology

- Name: ANDSHarvest
- Namespace: http://www.ands.org.au/ontologies/ns/0.1/VITRO-ANDS.owl#
- Some new classes:
  - OAIHarvest
  - Research Data
    - Research Catalog, Research Collection, Research Records Collection, Research Registry, Research Dataset
- Subject area Keyword

<table>
<thead>
<tr>
<th># class</th>
<th>shortdef</th>
<th>example</th>
<th>group</th>
<th>ontology</th>
<th>display level</th>
<th>update level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ANDSHarvest:OAIHarvest</td>
<td>ANDS HARVEST DETAILS</td>
<td></td>
<td>OAIHarvest</td>
<td>ANDSHarvest</td>
<td>editor</td>
<td>editor</td>
</tr>
<tr>
<td>2 ANDSHarvest:Research Data</td>
<td>Research Data</td>
<td>Survey results for project X</td>
<td>Information Resource</td>
<td>ANDSHarvest</td>
<td>public</td>
<td>self</td>
</tr>
<tr>
<td>ANDSHarvest:Research Registry</td>
<td>Research Registry</td>
<td></td>
<td>Information Resource</td>
<td>ANDSHarvest</td>
<td>unspecified</td>
<td>unspecified</td>
</tr>
<tr>
<td>ANDSHarvest:Research Repository</td>
<td>Research Repository</td>
<td></td>
<td>Information Resource</td>
<td>ANDSHarvest</td>
<td>unspecified</td>
<td>unspecified</td>
</tr>
<tr>
<td>ANDSHarvest:Research Dataset</td>
<td>Research Dataset</td>
<td></td>
<td>Information Resource</td>
<td>ANDSHarvest</td>
<td>unspecified</td>
<td>unspecified</td>
</tr>
<tr>
<td>3 ANDSHarvest:Subject area Keyword</td>
<td>Subject Area</td>
<td></td>
<td>Information Resource</td>
<td>ANDSHarvest</td>
<td>unspecified</td>
<td>unspecified</td>
</tr>
</tbody>
</table>
Many new object properties to map required relationships: (examples)

- ANDSHarvest:isDescribedBy
- ANDSHarvest:isManagedBy
- ANDSHarvest:isOwnedBy
- ANDSHarvest:isSupportedBy
- ANDSHarvest:isCollectorOf
- ANDSHarvest:isPointOfContactFor
- ANDSHarvest:publishedIn

Many new data properties: (examples)

- ANDSHarvest:arc
- ANDSHarvest:dateResearchCommenced
- ANDSHarvest:dateResearchCompleted
- ANDSHarvest:groupDescription
- ANDSHarvest:locationOfDataManagementPlan
- ANDSHarvest:nonDigitalDataLocation
- ANDSHarvest:retentionPeriod
RESEARCH CLASSIFICATION ONTOLOGIES

- Maintained by the Australian Bureau of Statistics
- Classifications:
  - Socio-economic Objective classifications (SEO)
  - Fields of Research (FOR)
  - Type of Activity (TOA)
- Hierarchical Classifications using 2, 4 and 6 digit codes
- Example
  - Division (86) Manufacturing
  - Group (8607) Agricultural Chemicals
  - Objective (860702) Chemical Fertilisers

- Classification modelled using:
  - rdfs:subClassOf
  - skos:broader
  - skos:narrower
  - skos:exactMatch
Metadata Exchange Hub

**Research Data Australia**

**OAI-PMH Provider**

- AND S PIDS
- NLA PIDS
- Other

**Persistent ID Generator Interface**

- RDF Model

**Existing Vitro**

- Landing Pages
- SPARQL Endpoint

**Presentation**

**Controllers and Logic**

**RDF Individuals**

**Ontologies**

**Data Transformer and Integrator**

- RDF Model

**Persistence Layer (eg. MySQL)**

**Metadata Exchange Hub**

**RDB**

**XML**

**CSV**
**Persistent Identifiers**

- Every RIF-CS object requires a globally unique identifier – or Persistent Identifier (PID).

- ANDS provides a web-service for minting PIDs based on the Handle System ([http://www.handle.net/](http://www.handle.net/)).

- The National Library of Australia is also involved with providing PIDS for people (party) through its People Australia project.

- Large national research bodies such as the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC) also have unique IDs for researchers.

- Exactly what PID to use for researchers and other parties is still to be resolved.
Providing Metadata for Harvesting

- Institutions expose RIF-CS metadata to Research Data Australia via OAI-PMH.

- To provide such metadata using VIVO we:
  1) perform a SPARQL Select query on VIVO to extract relevant records.
  2) apply a XSLT to transform the previous results into RIF-CS compliant XML.
  3) store the RIF-CS XML in a directory monitored by our OAI-PMH Provider software. The software encapsulates the XML within a compliant OAI-PMH feed ready for harvesting.

- The slight complication in this process is using a non-DC metadata schema with OAI-PMH.
WORKFLOWS

- Using Kepler 2.0 software for data workflow and transformation.
- Free Java software (BSD) which incorporates GUI and workflow engine with ability to execute workflows via GUI or command-line.
- Workflows can be nested and modularized for re-use with inbuilt mechanisms for packaging and sharing.
- Comes pre-built with over 300 actors and can be extended using Java API and other JVM languages.
- Has a module system that allows the construction of customised distributions for easy deployment.
WORKFLOWS

- An example test workflow.
WORKFLOWS

Some of the existing actors include such functionality as:
- file reading/writing and other file operations
- web service execution
- RDB connection and querying
- record and string manipulation
- mathematical functions
- Control flow

For our project we have added actors for:
- Jena DB connection
- SPARQL Queries, Select, Construct, Describe
- RDF Model operations
- CSV to XML
- XQuery
WORKFLOWS

- Custom Kepler actors for working with VIVO
WORKFLOWS

- Read CSV file, transform to RDF, merge with existing model and save to DB.
WORKFLOWS
WORKFLOWS

- SOLR / Jena
**WORKFLOWS**

- Workflows are saved as XML

- Command line execution:

```
Java -cp ./kepler.jar org.kepler.build.runner.Kepler -runwf -nogui hello.xml
```
WORKFLOWS

- To Do:
  - Actors for easier manipulation of RDF and XML
  - More documentation and examples.
  - Think about how to combine efforts with the VIVO Harvester.
Index bibliometric data using Lucene.

- Retrieve Wikipedia section labels from DBPedia. (RDF)
  - Other vocabularies/taxonomies could be used, eg. MeSH
- Clean wiki labels and keep 2-grams and 3-grams (“Concepts”)
  - Higher order n-grams may be used as well.
- Create “semantic representations” of the concepts found within publication abstracts using contextual features within the abstracts.
- Use heuristics to construct networks through the semantic space of the abstracts.
An automatically generated view onto a portion of the conceptual landscape of QUT’s ePrints repository.
- Can be used to generate semantic signatures/profiles for researchers.
- Can be used to discover researchers with similar interests.
- Can be used to explore the research landscape via associative links.
OUTCOMES

- Griffith’s implementation of VIVO for Research Data Australia has created a new, expansive project for an inclusive Griffith Profile system, making research data accessible via one entry point. It will expose the ‘linked’ nature of relationships between a researcher and their research outputs.
Outcomes Continued

• Most researchers are receptive to having information about them and their research activities published to Research Data Australia and displayed publically via a web semantic discovery solution.

• Participating in the Research Data Commons (aka Research Data Australia) has enabled each university to identify additional registry objects that ISO 2146 & RIF-CS will be able to accommodate for categorising and discoverability purposes (expansion of Service types and Commercial Activities [enterprises] con researchers).
OUTCOMES: CULTURAL CHANGE

1. Linking research activities, projects/grants, people, publications and datasets. Researchers using the central storage service will inevitably doing some data-centric research where the data is directly related to research work, publications, etc. Capturing information about the data and research activities/groups using the storage service gives the University (and ITS) some rich information about what’s sitting on the storage services, rather than just offering it up as a blank slate.

2. Data is the new intellectual property of the university. The University needs to maintain an awareness of all the data that is stored, used and how it is used. Different types of data are associated with different retention and destruction policies, and those policies need to be able to be linked not just to research activities, but the resources (e.g. storage) they are using.

3. ITS needs to capture information about how its resources are being used by the research community. Ultimately we need information as to how providing research services supports the Melbourne and Parkville research community. If we wish to attract more funding to provide research services from the University (and elsewhere) we need to collect the necessary data to ensure that cases for funding can be made in future.

- We mandate that on an annual basis the Storage Authorisation Contact tells us who’s using the storage service, what research group(s) are using it, and how much space they’re using (a fairly minimal requirement).
- We mandate that on an annual basis the Storage Authorisation Contact Review information we provide them on publications, people, projects/grants and activities, and help us clarify the links between say, a publication and a dataset which is stored on the storage service.

By reviewing that data — instead of mandating they provide us with new data — it means we can (1) verify the publications, services, people etc. that are stored centrally already and (2) clarify the links between those services, datasets, people, etc. This process is nothing new, as to a large extent it is already being done through the Research Data Registry project. It is just a case of ensuring that we can appropriately link the storage service and the research data registry activities.

Doing this will ensure that we build services into the research fabric of the university, rather than just offering it as a disconnected service.

Any thoughts & response on this would be appreciated. I’ll work with Rob to make sure we build this into the SLA.
CHALLENGES

- Early adopters of the technology – Little knowledge about VIVO in Australasia when our projects began.

- In kind contributions, specifically project support, exceeded our estimations for a comprehensive implementation that delivered on all requirements (ontology expansion for RIF-CS, landing pages, OAI Cat implementation with successful harvest tests from Research Data Australia).

- Sustainability after project / knowledge transfer – Identifying reliable service support for VIVO after project closure. Possible solutions include outsourcing to eResearch services & solutions (VeRSI, Intersect).
CHALLENGES CONTINUED

• IP issues of research data – Who owns the data? The University in most cases, however, there are a few exceptions to the rule.

- Griffith Intellectual Property Policy:

  - The intellectual property of which the University claims ownership under sections 2.1 and 2.2 of this IP Policy includes (without limitation) intellectual property in relation to:
  - (a) patentable and non-patentable inventions, together with field and laboratory notebooks relating to such inventions;
  - (b) registered and unregistered designs, plant varieties and topographies;
  - (c) databases, computer software, firmware, courseware, and related material;
  - (d) works generated by and/or with computer equipment or software, except where specifically exempted under other provisions in this IP Policy;
  - (e) University-commissioned works;
  - (f) Confidential information associated with each and every item listed in this section 3.0;
  - (g) Copyright in works and materials except for those specifically excluded by this IP Policy;
  - (h) Educational materials except for those specifically excluded by this IP Policy.
CHALLENGES CONTINUED

- **DATA Related:**
  - Privacy issues – Data that needs to be kept private and confidential due to the nature of the research vs. data that can be sanitised if permissible.
  
  - Data management practices by researchers is rudimentary. Getting access to the metadata/data is a challenge. Researchers are storing data in a variety of solutions, and a number of them are unreliable.
  
  - Data Storage – Lack of available funding for data storage has caused problems for Universities. Researchers are not allocating enough funding for storage in their grant applications.
CONCLUSION

- Questions?