



# Best Practice in Data Management

Mojtaba Lotfaliany; MD, PhDc

PhD Student @ Non-Communicable Disease Control, School of Population and Global Health, University of Melbourne

Researcher @ Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences



# What this presentation is about

## Included:

- 1) Controlled Vocabulary
- 2) Good record keeping and meta data
- 3) Access controlling and security
- 4) Version controlling and tracking

## Not included:

- 1) Data Storage
- 2) Non digital data
- 3) Intellectual property, copyright and ownership
- 4) Data retention

# Controlled Vocabulary



# Controlled Vocabulary

- Controlled vocabularies ensure shared understanding of the terminologies used in taxonomies and classifications.
- Using established vocabularies promotes interoperability, discovery and re-use of data.
- Goal of Controlled vocabulary

# Controlled Vocabulary examples



Australian Doctor Command  
Language (DOCLE)



Australian Medicines  
Terminology

**SNOMED CT**  
The global  
language of  
healthcare

**AIHW** Authoritative information and statistics  
to promote better health and wellbeing



**OpenGALEN**

**MedDRA**



TENTH EDITION  
**ICD-10**  
INTERNATIONAL CLASSIFICATION OF DISEASES  
ICD-10 is a new code set for reporting  
medical diagnoses & inpatient procedures.

Victorian Admitted Episodes Dataset (VAED)

**R<sub>x</sub>N  
orm**

Cumulative Index of Nursing  
and Allied Health Literature  
(CINAHL)



# Controlled Vocabulary

## Controlled vocabularies in medicine and health

**Kate LeMay**

Senior Research Data Specialist

August 2016



# Controlled Vocabulary



[About ANDS](#) | [Events](#) | [Contact Us](#)



[Search for Research Data](#)

Search the ANDS Site



About us



News & Events



Partners & Communities



Working with data



Online Services



Guides



Online Services

Research  
Vocabularies  
Australia

ANDS | Online Services

## Research Vocabularies Australia

# Controlled Vocabulary

## Vocabulary Creators/Providers

### Vocabulary Administrators



Ontologists

Librarians

Data managers

Vocab managers

### Applications



Vocabulary services

## Vocabulary Consumers

### People



Researchers

Librarians

Data managers

### Applications



Data centres

Portals

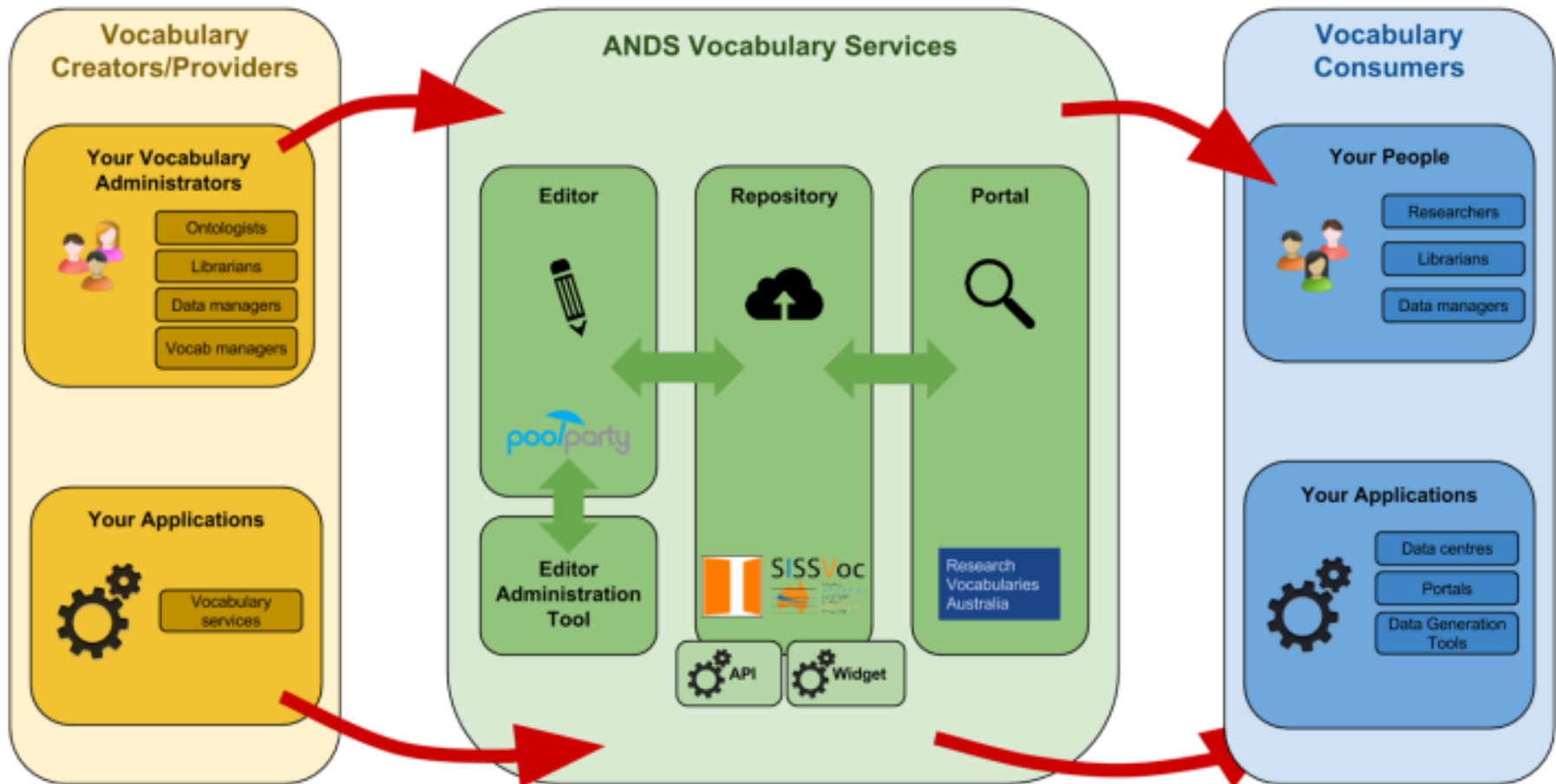
Data Generation Tools

Vocabulary services

How do we connect these people and systems?



# Controlled Vocabulary



# Controlled Vocabulary

1. **Find** and **learn** about controlled vocabularies relevant to research.
2. **Access** those vocabularies and **reuse** them in your community.
3. **Integrate** vocabularies into your local information systems at a technical level.
4. **Upload** and **describe** a vocabulary to **share** with others.
5. Make a vocabulary **machine readable** (more easily integrated into other's systems).
6. **Create** new or **import** existing vocabularies and **manage** them with your community's input.

# Controlled Vocabulary

The screenshot displays the SKOS (Simple Knowledge Organization System) interface for the concept 'Immunology'. On the left, a tree view shows the hierarchy of research fields, with 'Immunology (10)' selected. The main area is titled 'Immunology' and includes a URL: <http://ands.poolparty.gov.au/anzsrc/773>. The interface features several tabs: 'Details', 'Notes', 'Documents', 'Linked Data', 'Triples', 'Visualization', 'Quality Report', and 'History'. The 'Details' tab is active, showing various metadata and relationships.

**SKOS +**

**Relations** (Add to Blacklist, Linguistics)

- Broader Concepts**  
[Medical and Health Sciences](#)
- Narrower Concepts**  
[Cellular Immunology](#)  
[Autoimmunity](#)  
[Applied Immunology \(incl. Antibody Engineering, Xenotransplantation and T-cell Therapies\)](#)  
[Immunology not elsewhere classified](#)  
[Allergy](#)  
[Tumour Immunology](#)  
[Transplantation Immunology](#)  
[Innate Immunity](#)  
[Immunogenetics \(incl. Genetic Immunology\)](#)  
[Humoural Immunology and Immunochemistry](#)
- Related Concepts**
- Exact Matching Concepts**
- Close Matching Concepts**
- Broader Matching Concepts**
- Narrower Matching Concepts**

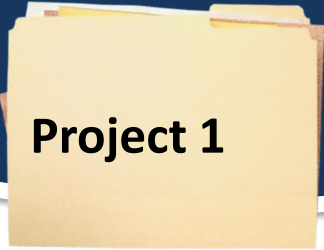
**Metadata:**

- Preferred Label:** Immunology
- Alternative Labels:**
- Hidden Labels:**
- Notation:** 1107
- Scope Notes:**
- Example:**
- Definitions:**

# Controlled Vocabulary

- Identifiers
- File naming
- Classification systems

# Organize Files Logically



Make sure your file system is logical and efficient



Project Name	Location	Experiment Name	Date	File Format
				Biodiv_H20_heatExp_2005_2008.csv
				Biodiv_H20_predatorExp_2001_2003.csv
				Biodiv_H20_planktonCount_start2001_active.csv
				Biodiv_H20_chla_profiles_2003.csv

# Choosing File Formats

- Accessible Data (in the future)
  - Non-proprietary (software formats)
  - Open, documented standard
  - Common, used by the research community
  - Standard representation (ASCII, Unicode)
  - Unencrypted & Uncompressed
  - Easy to track changes
  - Easy to be converted
  - Minimal human intervention

**Good record keeping and meta data**



# Good record keeping and meta data

**Metadata is data about data; information about information.**

Metadata is structured information associated with an object for purposes of discovery, description, use, management and preservation (NISO, 2007).

- A. **Descriptive metadata** (sufficient detail to uniquely identify it and enable its retrieval/discovery)
- B. **Structural metadata** (understanding how objects relate to each other and other entities)
- C. **Administrative metadata** (when and how it was created, file formats, technical details, and access rights.)
  - **Rights Management** (legal issues such as intellectual property rights, privacy and confidentiality;)
  - **Preservation metadata** (archive and preserve the resource.)



# Good record keeping and meta data

## **General and discipline based standards**

Metadata can be generated manually, or it can be created automatically.

Within data base or in separate files

**Examples:** *Dublin Core ; Data Documentation Initiative (DDI)*

# Good record keeping and meta data

KDPP\_master\_20160701

DTA File (encrypted)



Title: K-DPP Master File (Version 4)  
Subject: Raw Data  
Tags: #KDPP; #WideForm; #BothSex  
Categories: Stata data  
Comments: In comparison with previous version: 1) Biomarkers were cleaned 2) Doctor Diagn  
Rating: ★★★★★  
Authors: M. Lotfaliany; T. Sathish  
Revision number: 4

## File Meta 1.3

Rating: ★★★★★ Based on 5 ratings

Reviewed: 5 reviews

Downloads: 4115

Change Set: 83206

Released: Sep 10, 2015

Updated: Nov 27, 2015 by Dijji

Dev status: Stable ?

### DOWNLOADS

 [FileMeta\(x64\).msi](#)

application, 1124K, uploaded Sep 10, 2015 - 3075 downloads

 [FileMeta\(x86\).msi](#)

application, 1100K, uploaded Sep 10, 2015 - 667 downloads

 [FileMeta 1.3 sources.zip](#)

source code, 1223K, uploaded Sep 10, 2015 - 373 downloads

# Data Dictionary

## Project Documentation

- Context of data collection
- Data collection methods
- Structure, organization of data files
- Data sources used
- Data validation, quality assurance
- Transformations of data from the raw data through analysis
- Information on confidentiality, access and use conditions

## Dataset Documentation

- Variable names and descriptions
- Explanation of codes and schemas used
- Algorithms used to transform data
- File format and software (including version) used

# Access controlling and security



# Access controlling and security

- Keep master copy to an assigned team member
- Restrict write access to specific members (developers)
- Record changes with version control
- Keep confidential data off internet servers (or behind firewalls), put sensitive materials on computers not connected to the internet

# Access controlling and security

- Physical security... who has access to your office.
- Computer: Keep virus protection up to date, does your computer have a login password, not sending personal or confidential data via e-mail or FTP, transmit via encrypted data, imposing confidentiality agreements for data users

# Access controlling and security

- Quick Links**
- Dashboard
  - Schedule Manager
  - View Alert
  - Add Alert
  - Pending Email
  - Pending SMS
  - Patient Search
  - Test Search
  - Sample Collection
  - Rejected Sample
  - Unpaid Bill
  - Help Text
  - Laboratory Support

**Login Status**

Login As: Admin Admin  
 Since: 20/04/2011 19:16  
 Designation: Admin  
 UserType: Super Administrator  
 Logout  
 Help

## ACCESS CONTROL

### Access Control

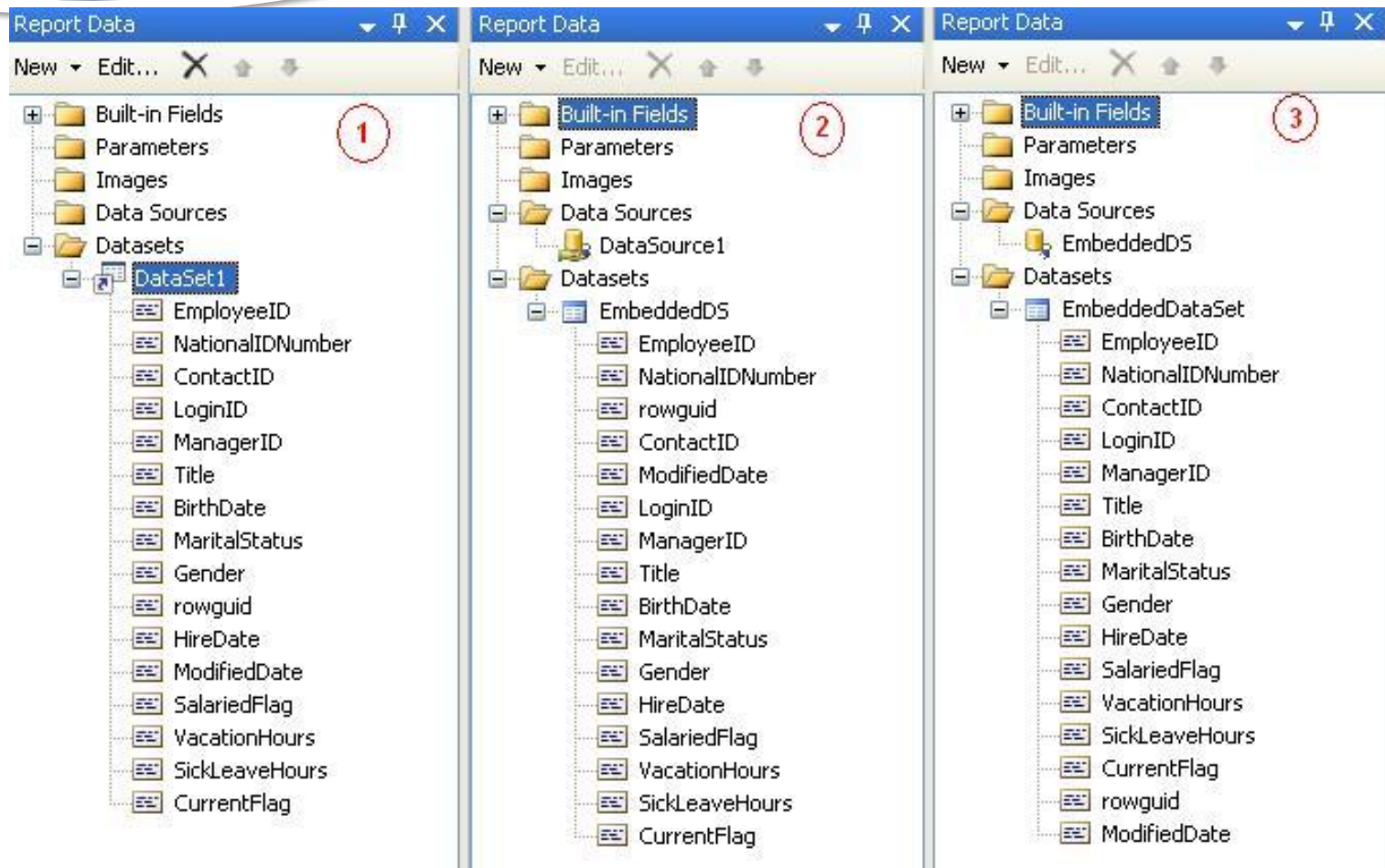
G1 Doctor	G2 Receptionist
G3 Technician	G4 Lab Coordinator
G5 Bio Chemist	G6 Administrator
G7 Super Admin	G8 Accountant
G9 Collection Center	

Module <small>(Click on the green button to show page list)</small>	G1	G2	G3	G4	G5	G6	G7	G8	G9
Schedule Manager	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>
cm pre visit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
cm update visit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
cm view schedule	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
common script schedule	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
lab worksheet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
pending schedule	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
schedule	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
schedule1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
common script sms	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Management	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>
Referral Manager	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>
Sample Collection	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>
Accounts	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>
Out Source Manager	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>
Inventory	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>
Test Manager	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>
Search Manager	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>
Collection Center	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>	<input checked="" type="checkbox"/> <a href="#">Click Here</a>

Save

Go To Dashboard

# Access controlling and security





# Version controlling and tracking



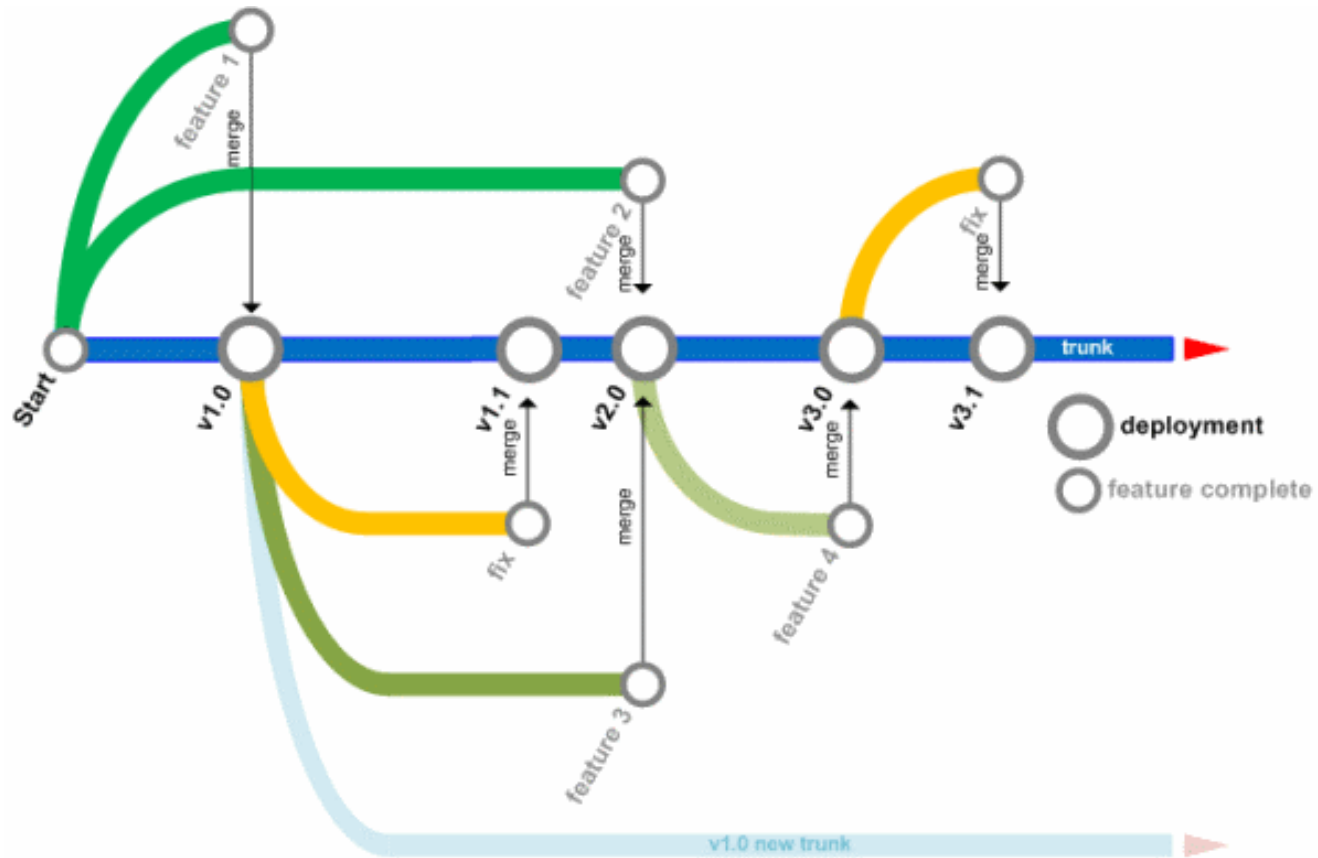
# Version controlling and tracking

**Table 3. Most Cited Retracted Articles**

First author	Journal	Year published	Year retracted	Times cited*	Reason for retraction
Wakefield	<i>Lancet</i>	1998	2004; 2010	758	Fraud
Reyes	<i>Blood</i>	2001	2009	740	Error
Fukuhara	<i>Science</i>	2005	2007	686	Error
Nakao	<i>Lancet</i>	2003	2009	626	Fraud
Chang	<i>Science</i>	2001	2006	512	Error
Kugler	<i>Nature Medicine</i>	2000	2003	494	Fraud
Rubio	<i>Cancer Research</i>	2005	2010	457	Error
Gowen	<i>Science</i>	1998	2003	395	Fraud
Makarova	<i>Nature</i>	2001	2006	375	Error
Hwang	<i>Science</i>	2004	2006	368	Fraud
Potti	<i>The New England Journal of Medicine</i>	2006	2011	361	Fraud
Brugger	<i>The New England Journal of Medicine</i>	1995	2001	336	Fraud
Van Parijs	<i>Immunity</i>	1999	2009	330	Fraud
Potti	<i>Nature Medicine</i>	2006	2011	328	Fraud
Schön	<i>Science</i>	2000	2002	297	Fraud
Chiu	<i>Nature</i>	2005	2010	281	Error
Cooper	<i>Science</i>	1997	2005	264	Fraud
Le Page	<i>Cell</i>	2000	2005	262	Error
Kawasaki	<i>Nature</i>	2004	2006	243	Fraud
Hwang	<i>Science</i>	2005	2006	234	Fraud

\*As of June 22, 2012

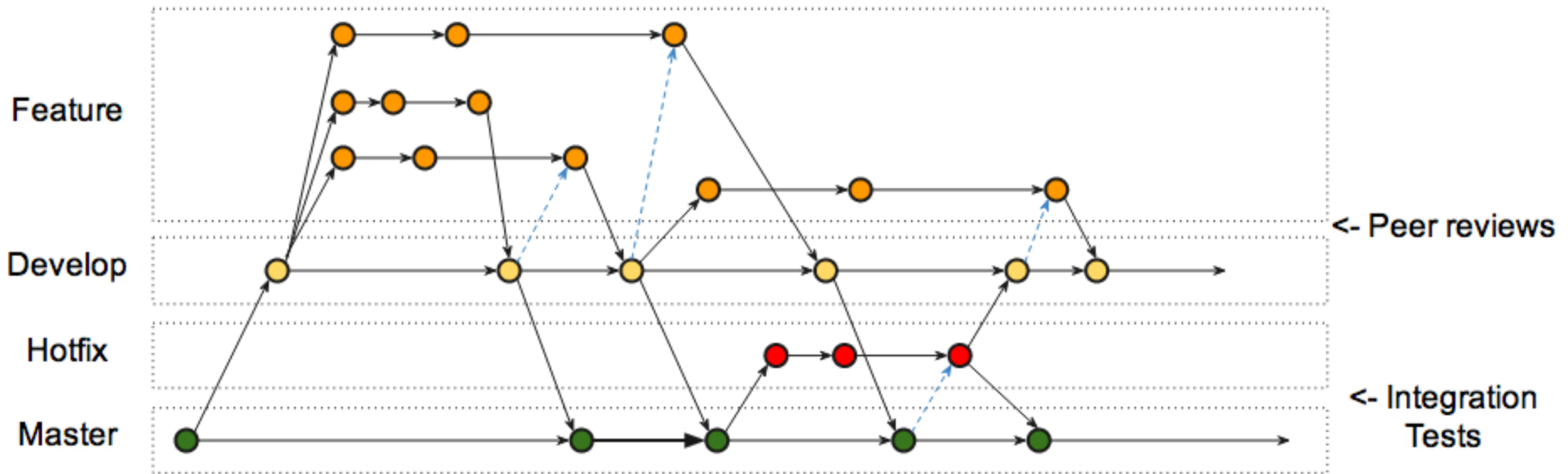
# Version controlling and tracking



History

Revision		Comment	Stamp	Date	Author
4112:3e045e7ad7 ...	master	changed about dialog wording	<span>2.0.0.5119</span>	Jun 21 at 3:34pm	rosey@example.com
4111:eaed86a2aa ...	H4332	Merged.		Jun 21 at 3:47pm	astro@example.com
4110:63d80d8808 ...		Minor changes to the lock command and lock dialog.	<span>2.0.0.5117</span> <span>2.0.0.5118</span>	Jun 21 at 3:15pm	spacely@example.com
4109:cc511381f6 ...		sanitize commit/wit relation text displayed in wit history	<span>2.0.0.5116</span>	Jun 21 at 1:24pm	gabbo@example.com
4108:9907a1b899 ...		force inputwrapper divs to contain their floated innards	<span>2.0.0.5114</span> <span>2.0.0.5115</span>	Jun 21 at 11:26am	gabbo@example.com
4107:69357b717c ...		restored lost IE style fixes	<span>2.0.0.5113</span> <span>goodBuild</span> <span>goodTests</span> <span>2.0.0.10791</span> <span>goodBuild</span> <span>goodTests</span>	Jun 19 at 4:25pm	gabbo@example.com
4106:3cee16604c ...		merge	<span>2.0.0.5111</span> <span>2.0.0.5112</span>	Jun 19 at 4:02pm	rosey@example.com
4105:c27125a78f ...		added secondary table styles		Jun 19 at 4:01pm	rosey@example.com
4104:aa12c46c43 ...		mrg	<span>2.0.0.5109</span> <span>2.0.0.5110</span>	Jun 19 at 1:28pm	eric@sourcegear.com
4103:4fdae8622c ...		mrg		Jun 19 at 1:26pm	eric@sourcegear.com
4102:dd38408728 ...		change dagnum for the hooks dag. fix db_deltas verify to skip abandoned dags.		Jun 19 at 1:23pm	eric@sourcegear.com
4101:6f5878f9e6 ...		merge	<span>2.0.0.5108</span>	Jun 19 at 12:27pm	rosey@example.com
4100:b654991700 ...		added gray up triangle		Jun 19 at 12:26pm	rosey@example.com
4099:ba00cecc3a ...		simplest better-than-it-was fix for new-build-status controls	<span>2.0.0.5107</span>	Jun 19 at 12:11pm	gabbo@example.com
4098:bf4738c365 ...		Merged.		Jun 19 at 11:10am	astro@example.com
4097:29bfcaba8b ...		Set CRULOPT_TIMEOUT to 0 for good measure... though 0 should be the		Jun 19 at 11:00am	astro@example.com

# Version controlling and tracking



# Conclusion

## MDM Maturity Model



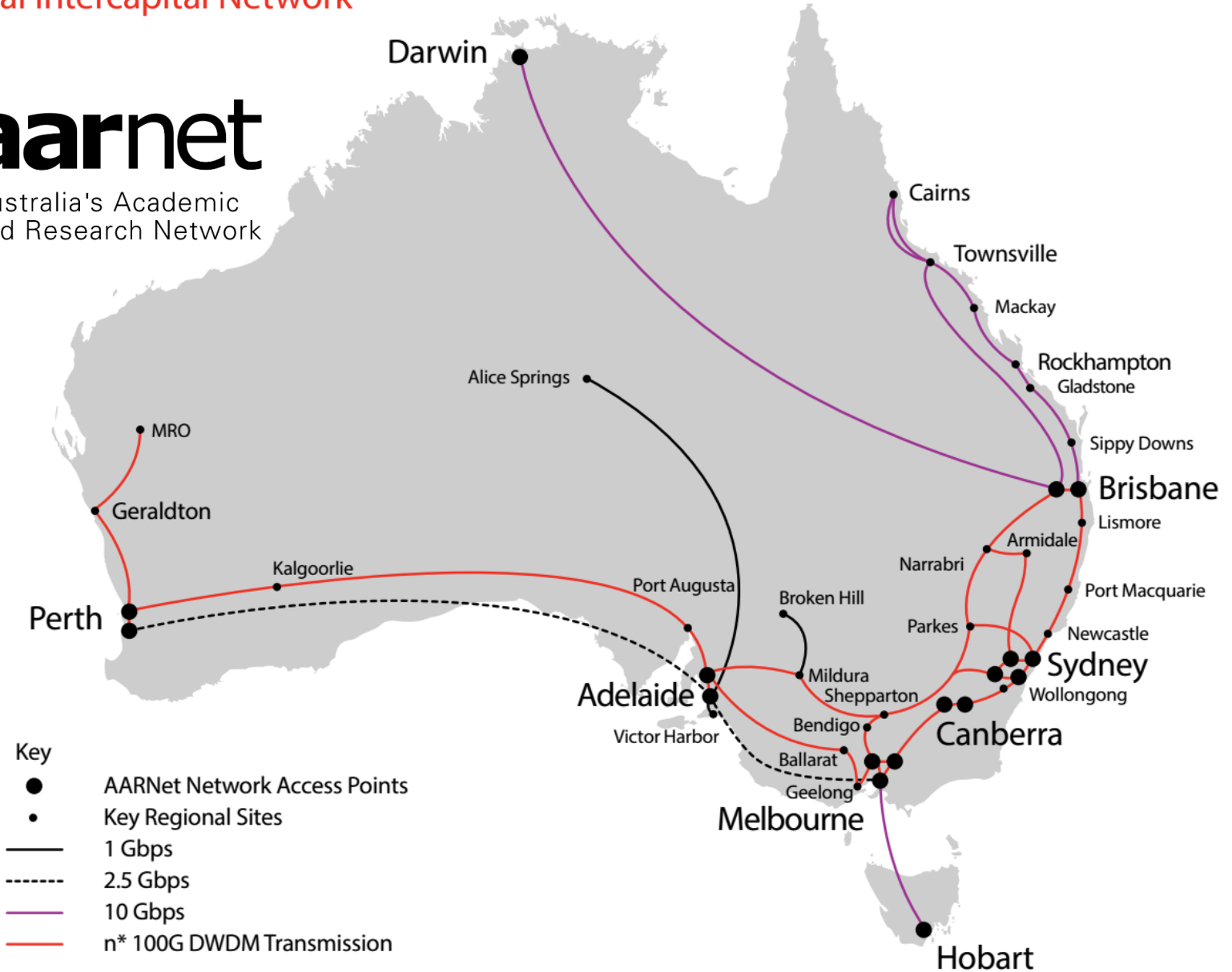


# AARNet National Intercapital Network



# aarnet

Australia's Academic  
and Research Network





# Acknowledgment , References



UNIVERSITY  
*of* VIRGINIA  
LIBRARY



University of Melbourne Research Platforms

We empower researchers to do great things with IT



# Acknowledgment , References



**NCRIS**  
National Research  
Infrastructure for Australia  

---

An Australian Government Initiative

# Contact info

- Email: [mloftfaliany@student.unimelb.edu.ac](mailto:mloftfaliany@student.unimelb.edu.ac)
- Tel: +61 450 55 1367
- Please don't hesitate to contact me!