



Complex licence requirements for the Bioregional Assessment Programme managed by provenance

Nicholas J Car & Matthew P Stenson

4th December 2015

The paper and this presentation

- The paper describes:
 - Attributions / licences separation
 - The licence information model
 - The Licence Tool
 - How the Licence Tool is used with other BA Linked Data tools
 - In outline, how transitive licence requirements are calculated via provenance
- This presentation will:
 - Overview relevant aspects of BA
 - Mention the original BA licencing issues
 - Skip most of the paper's content
 - Detail the transitive licence requirements calculations
 - Show how to calculate transitive licence requirements resolution

Bioregional Assessment

The Bioregional Assessment Programme (BAP), is a large, multidisciplinary program of work, assessing the impacts of coal and coal seam gas mining on water resources.

Several thousand datasets collected by and generated by the BAP will be stored long-term, enabling reuse over time.

All the 'Derived' datasets – those generated by the BAP – are licenced for open access and reuse – CC-BY 3.0.

Those 'Source' datasets from mining companies, community groups and government agencies have a range of licences and restrictions.



Original BAP licencing goals

- Collect licence information for each Derived dataset
 - Store the info in the Metadata Catalog
- Note conditions and handle them per dataset, including for descendent datasets
 - A licence class-based approach was to be taken: most restrictive wins
- Present a stacked attribution statement for each Dataset, based on its ancestors:
 - © CSIRO 2011,
 - © Commonwealth of Australia (Bureau of Meteorology) 2012
 - © Commonwealth of Australia (Bioregional Assessments) 2015

Original BAP licencing goals

- Collect licence information for each Derived dataset
 - Store the info in the Metadata Catalog
 - Note conditions and handle them per dataset, including for descendent datasets
 - A licence class-based approach was to be taken: most restrictive wins
 - Present a stacked attribution statement for each Dataset, based on its ancestors:
 - © CSIRO 2011,
 - © Commonwealth of Australia (Bureau of Meteorology) 2012
 - © Commonwealth of Australia (Bioregional Assessments) 2015
- Ensures much more restriction than necessary
- Only known from dataset provenance
- Special case of the above

Original BAP licencing goals

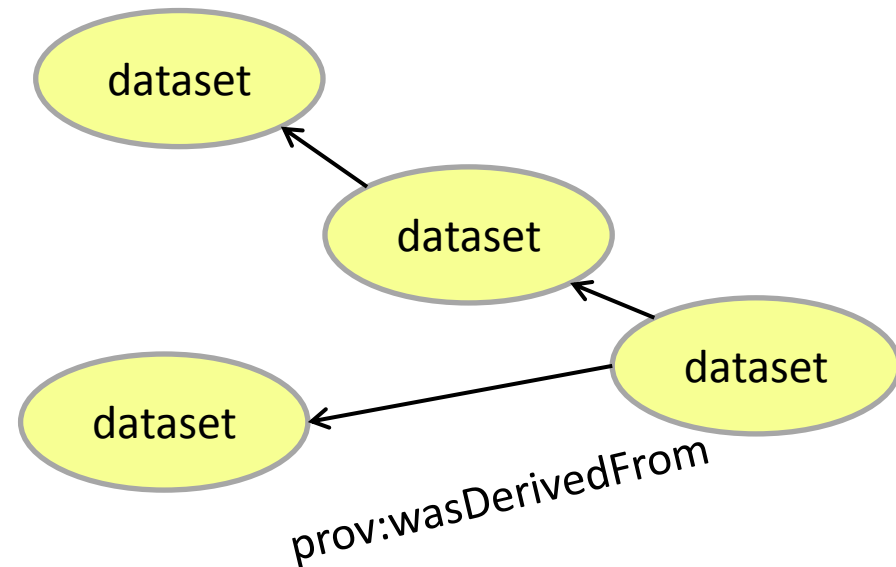
GUID	Dataset title	Bioregion	Source or D	Licence type (base licence)	Licence attribution 1	Licence attribution 2
4ffd342b-	2.5M Landmass Mask	BA_ALL	Derived	Creative Commons Attribution	© Commonwealth of Australia (Bioregiona	© Commonwealth of Austr
48f9247c-	9 Arc-second Hillshad	BA_ALL	Derived	Creative Commons Attribution	© Commonwealth of Australia (Bioregiona	© Commonwealth of Austr
8b65c3a4-	ABS Boundaries 2011	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Australian I	
65261aa9-	ABS Cooper LGA Work	LEB	Source	Creative Commons Attribution	© Commonwealth of Australia (Australian I	
ee39fa76-	ABS Mesh Block Popu	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Australian I	
c27fc127-	ABS Regional Populat	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Australian I	
d6b0c5c7-	ASRIS Australian Soil	BA_ALL	Source	Creative Commons Attribution	© CSIRO 2014	
e0a8bc96-	Asset database for th	GIP	Derived	Restricted licence for BA ONL	© Commonwealth of Australia (Bioregiona	
6968b11f-	Asset list for Galilee -	GAL	Derived	Restricted licence for BA ONL	© Commonwealth of Australia (Bioregiona	
b986c6c3-	Asset list for Galilee -	GAL	Source	Restricted licence for BA ONL	© Commonwealth of Australia (Bioregiona	
a5d8066e-	Asset list for Galilee -	GAL	Derived	Restricted licence for BA ONL	© Commonwealth of Australia (Bioregiona	
5d09d8f2-	Australia - Present Ma	BA_ALL	Source	Limited licence for BA ONLY t	© Commonwealth of Australia (Departmen	
e4a7604e-	Australian Average Ar	BA_ALL	Derived	Limited licence for BA ONLY t	© Commonwealth of Australia (Bioregiona	© Commonwealth of Austr
9d5a8d74-	Australian Coal Basins	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Geoscience	
d51a641a-	Australian Land Tenur	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Geoscience	
14e96462-	Australian Mineral De	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia 2012	
719a5433-	Australian Water Obs	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Geoscience	
d263c464-	Australian Water Res	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Bureau of M	
d4561e86-	Australia's River Basir	BA_ALL	Source	Creative Commons Attribution	© Commonwealth of Australia (Geoscience	
08933408-	BA All Regions BILO c	BA_ALL	Derived	Creative Commons Attribution	© Commonwealth of Australia (Bioregiona	© CSIRO

Hypothesis

No large collection of datasets with diverse licences can, to date, guarantee all the conditions of all licences are met by the holding repository.

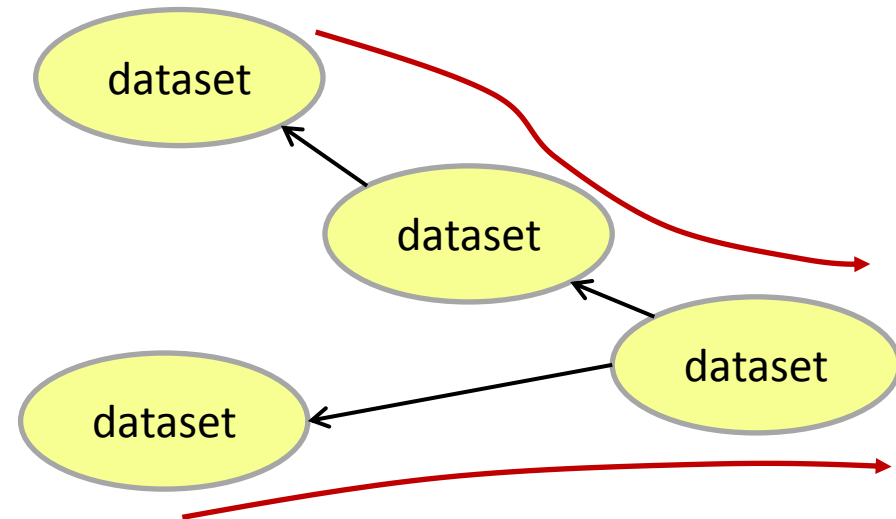
Licence calculation via provenance

- Provenance graph of dataset is the point of truth regarding dataset (and therefore licence) relations



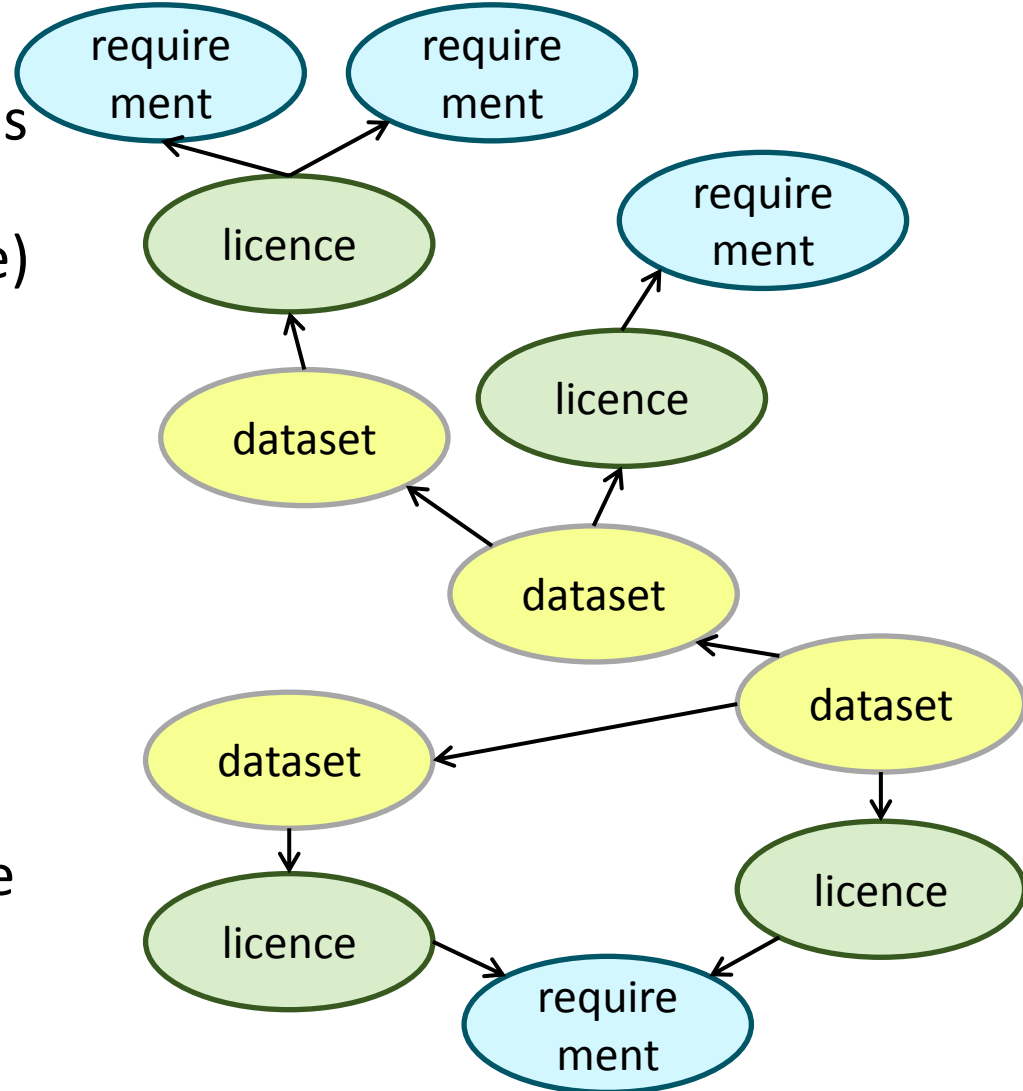
Licence calculation via provenance

- Provenance graph of dataset is the point of truth regarding dataset (and therefore licence) relations
- Use the provenance graph to determine all information flows

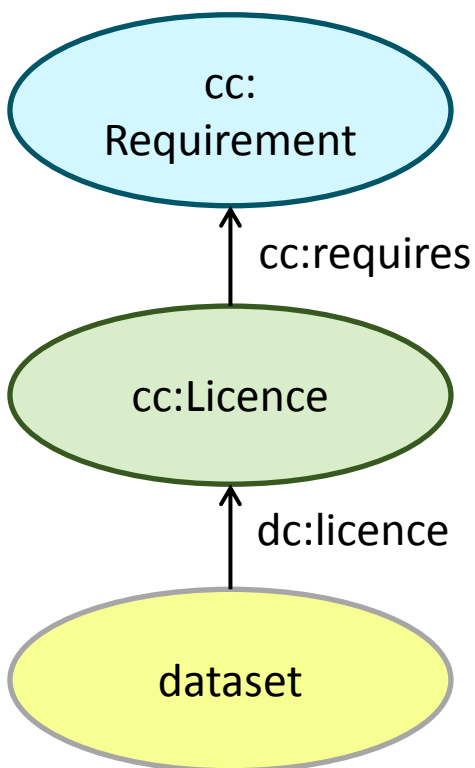


Licence calculation via provenance

- Provenance graph of dataset is the point of truth regarding dataset (and therefore licence) relations
- Use the provenance graph to determine all information flows
- Add metadata from various project sources to provenance graph as needed



The CC Licence Model

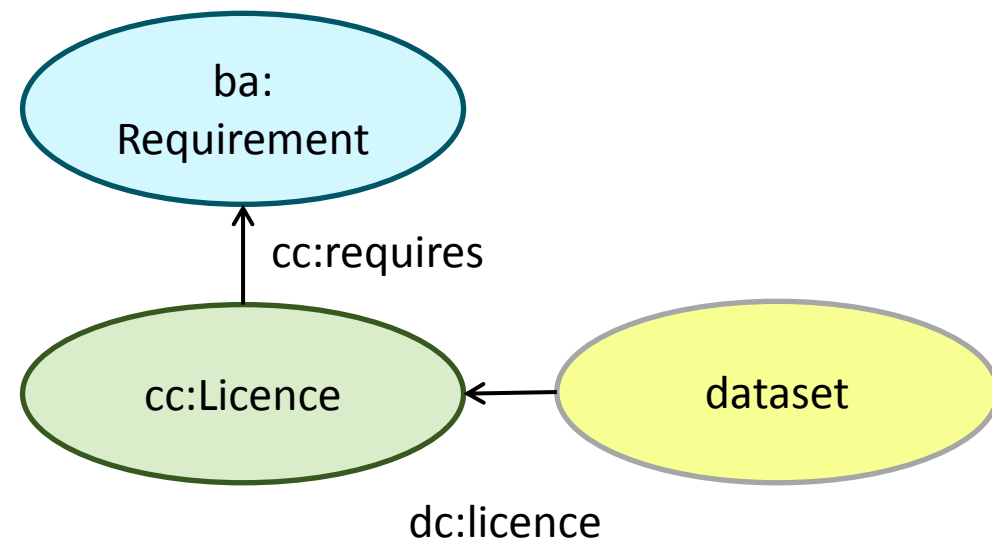


- The Creative Commons licence information model¹ has a “Requirement” class that a licence can point to
- Requirement class is “an action that may or may not be requested of you”
- Requirement not a well defined Semantic Web class

¹ <http://creativecommons.org/ns>

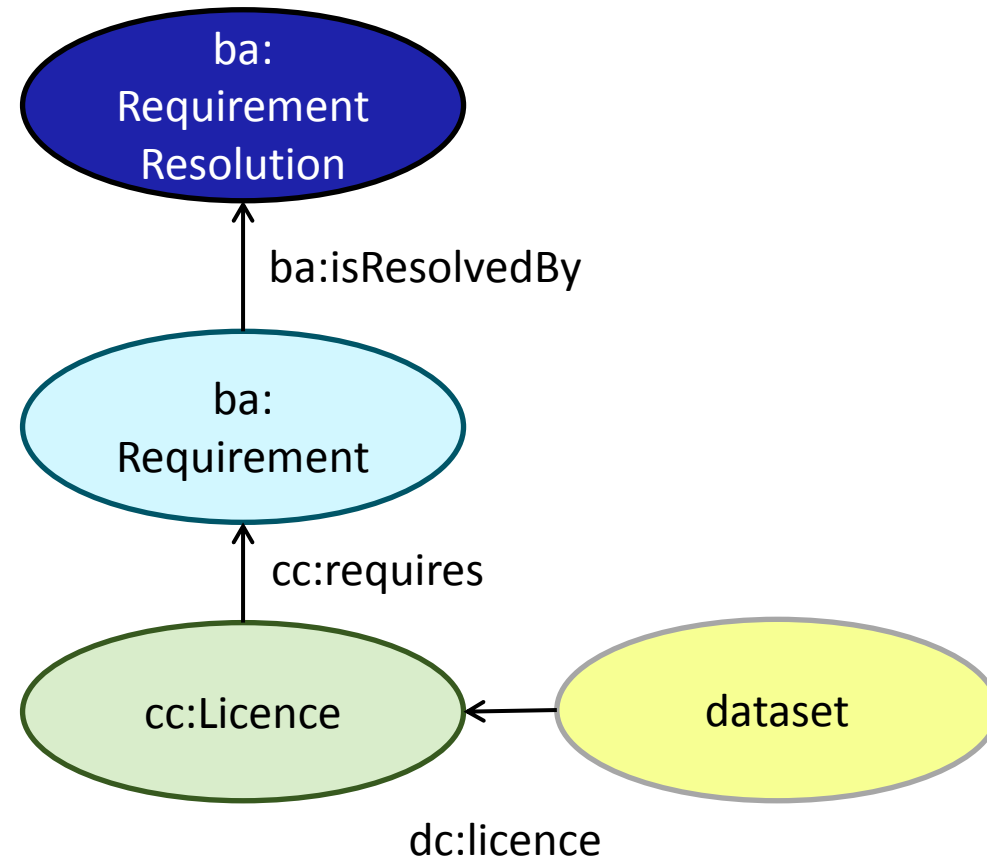
The BA Licence Model

- ba:Requirement things are OWL things
- Demo list at <http://data.bioregionalassessments.gov.au/id/requirement/>



¹ <http://creativecommons.org/ns>

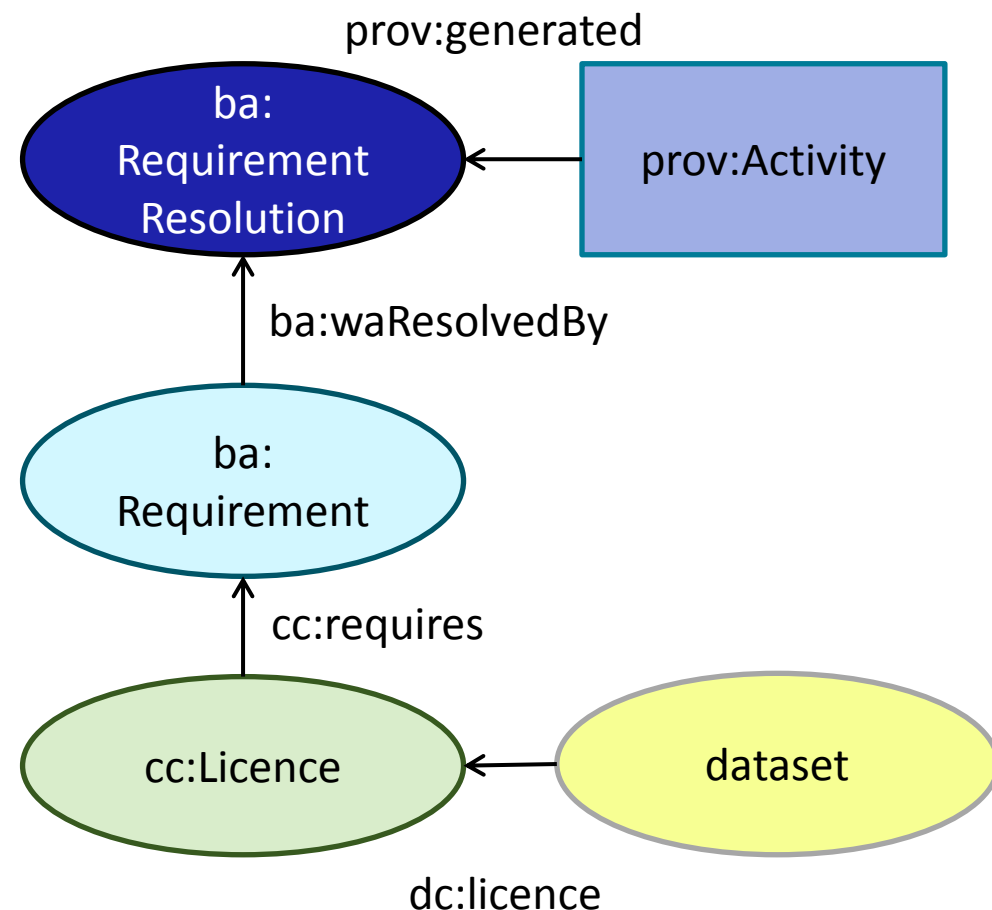
The BA Licence Model



- **ba:Requirement** things are OWL things
- Demo list at <http://data.bioregionalassessments.gov.au/id/requirement/>
- Requirements are resolved by **ba:RequirementResolution** objects

¹ <http://creativecommons.org/ns>

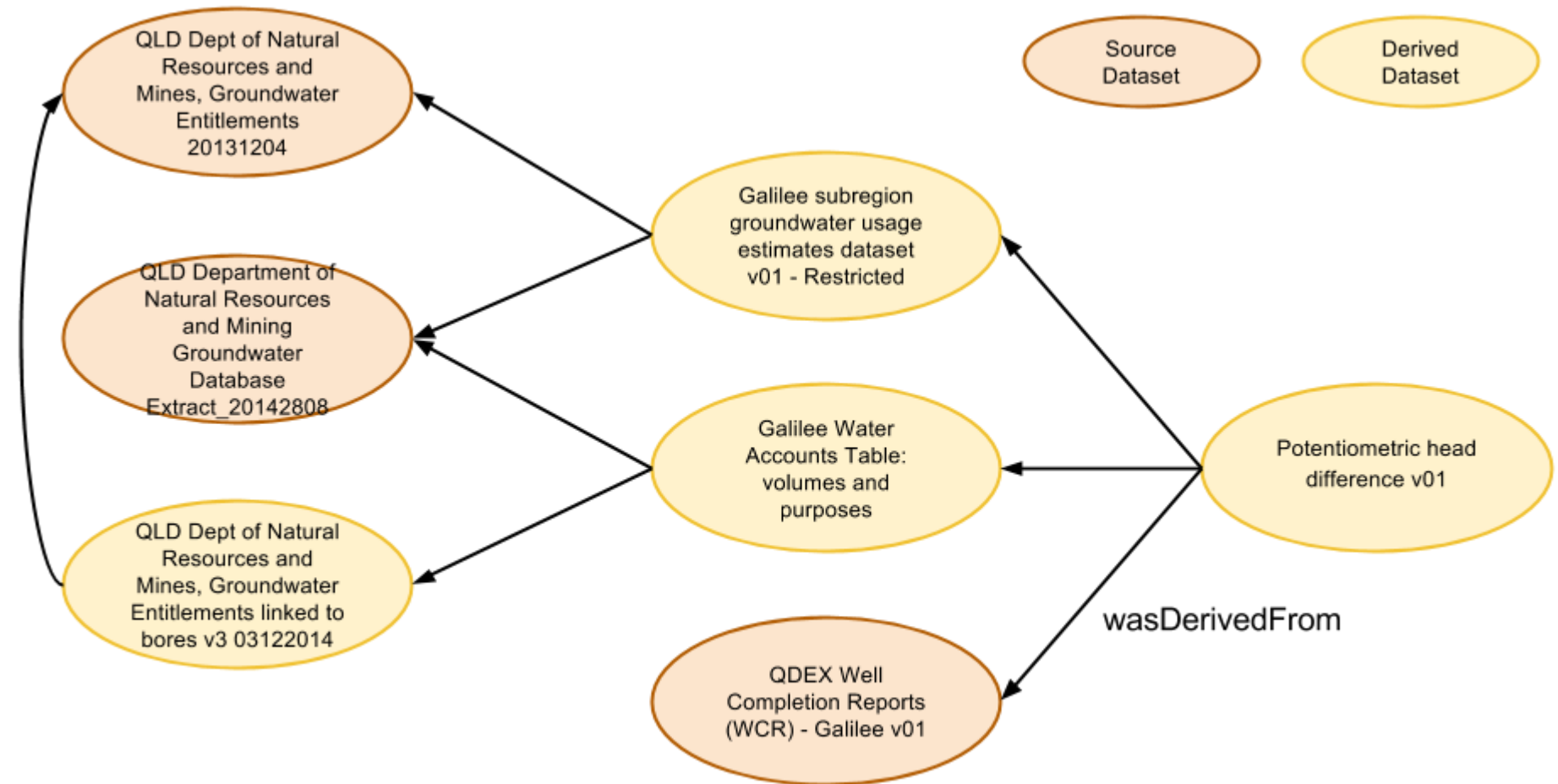
The BA Licence Model



- **ba:Requirement** things are OWL things
- Demo list at <http://data.bioregionalassessments.gov.au/id/requirement/>
- Requirements are resolved by **ba:RequirementResolution** objects
- These are generated by actions
 - Also have a demo list [/id/requirementresolution/](#)

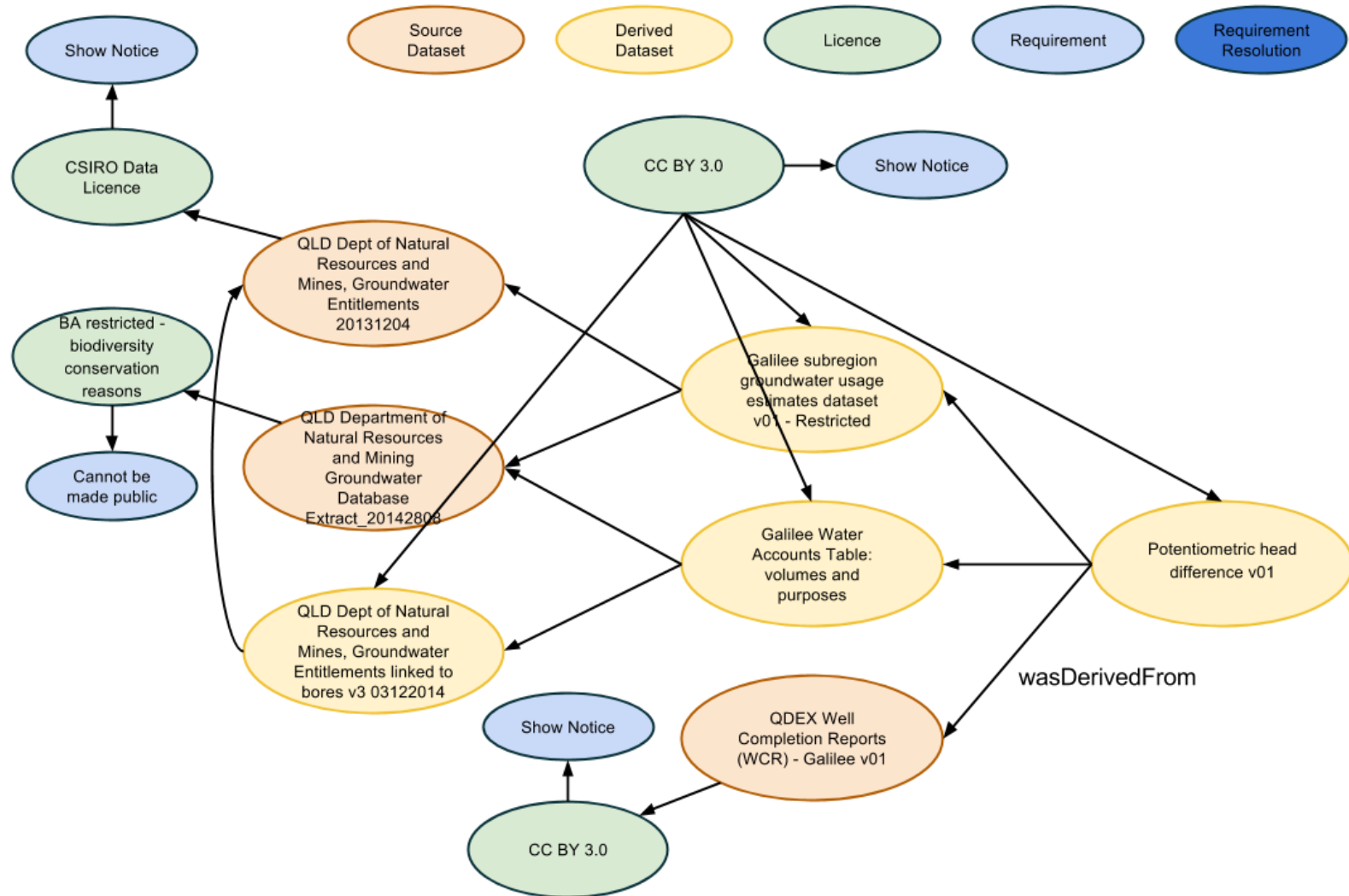
¹ <http://creativecommons.org/ns>

Real BA Case Study



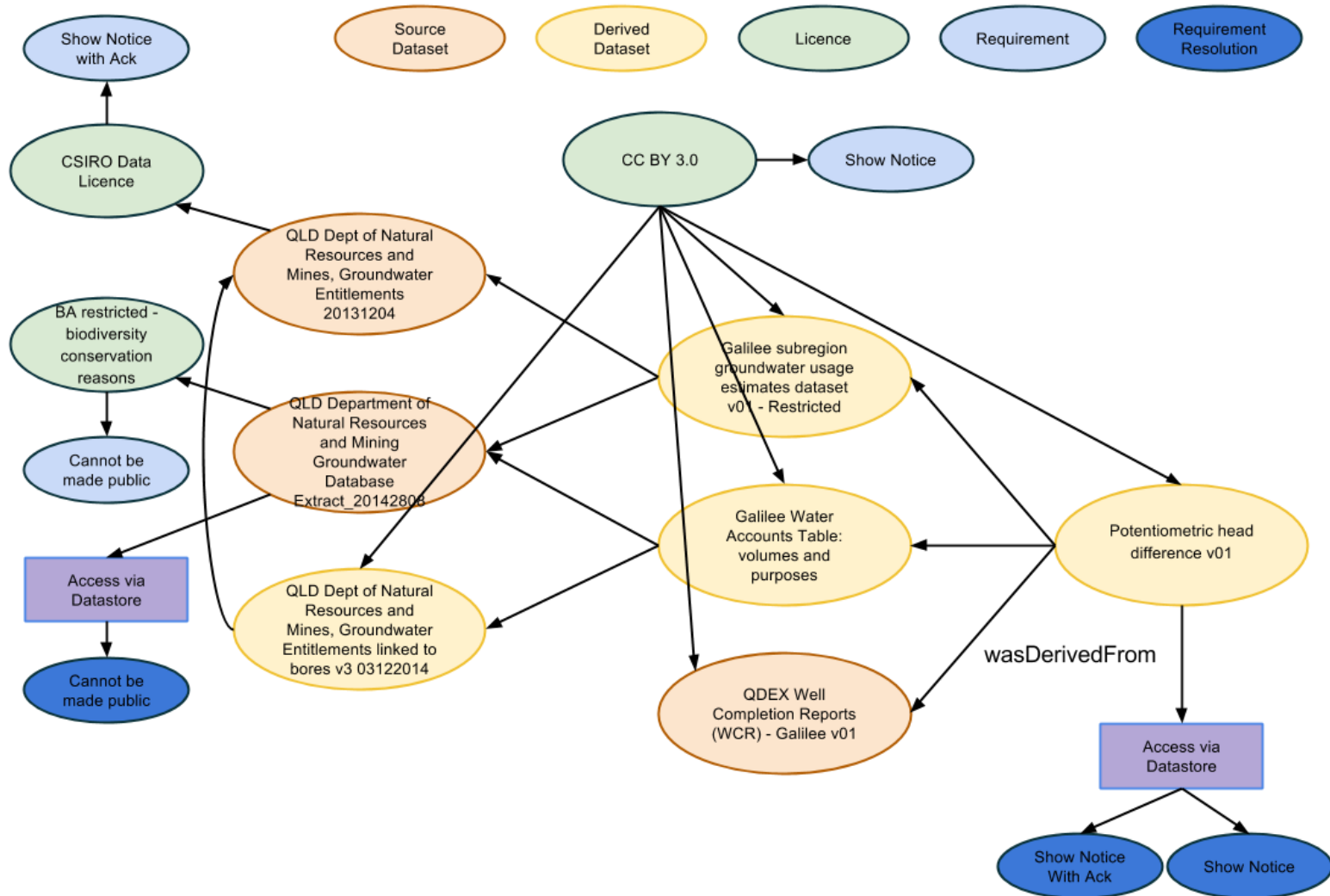
From: http://data.bioregionalassessments.gov.au/function/metadataexporter/6c212a1a-658c-41de-92a7-6054349a848b?_view=provenance

Real BA Case Study



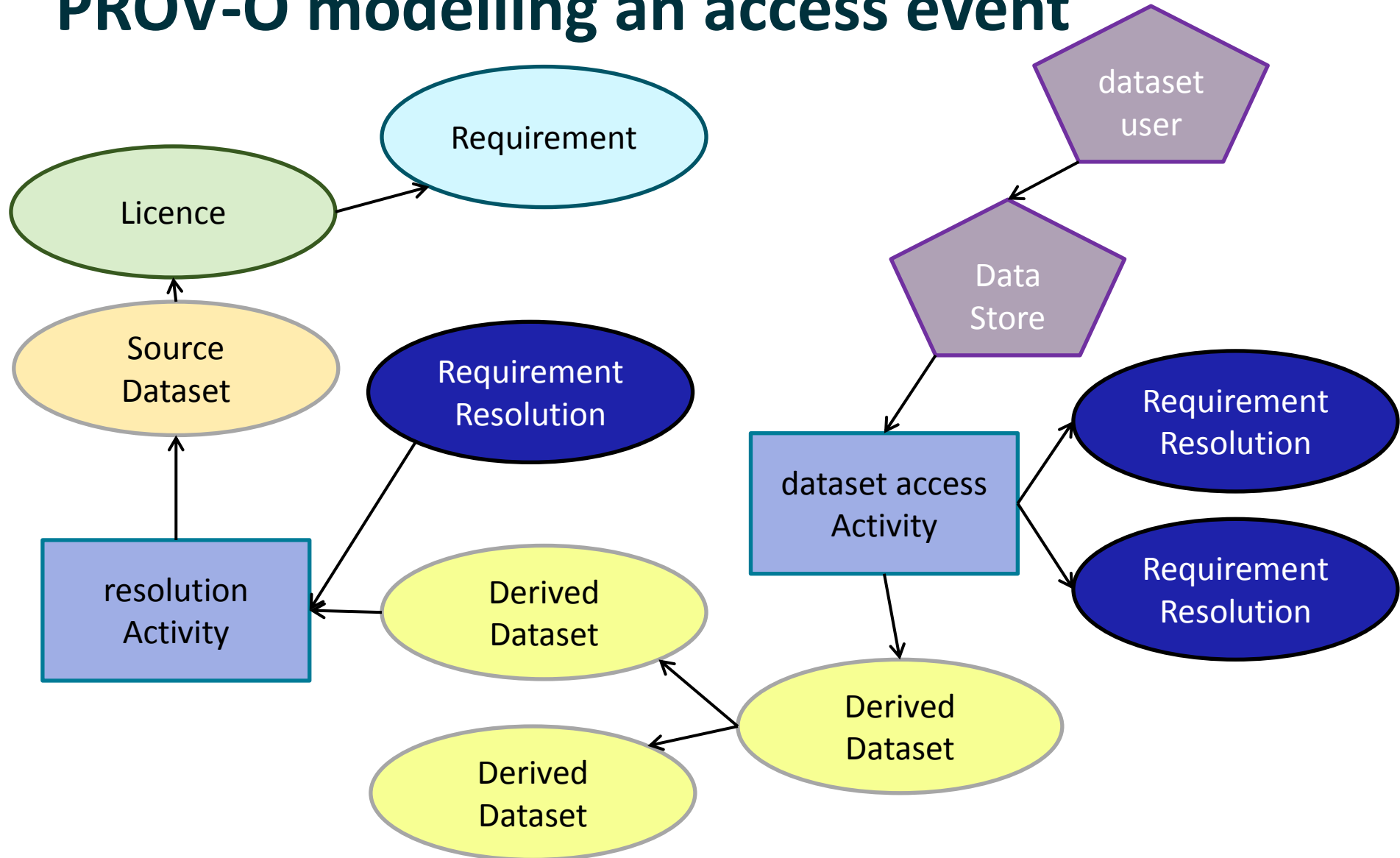
From: http://data.bioregionalassessments.gov.au/function/metadataexporter/6c212a1a-658c-41de-92a7-6054349a848b?_view=provenance

Real BA Case Study



From: http://data.bioregionalassessments.gov.au/function/metadataexporter/6c212a1a-658c-41de-92a7-6054349a848b?_view=provenance

PROV-O modelling an access event



Findings

- There are very few distinct Requirements
 - Perhaps 10
- Most Requirements can be resolved by data store actions
 - Show notices
 - Limit access
- Some Requirements can be resolved only by one-time deliberate action
 - Removing/lumping/de-identifying/averaging
- Resolution status of all Requirements for all datasets can be calculated easily using a provenance graph + metadata

Future Work

- Further normalise Requirements to cater for identical actions, different data (e.g. attribution)
- Make a project-independent licencing ontology
 - Extension to PROV & CC
- Publish real Requirement & RequirementResolution objects
 - As demonstrators
 - For others to use directly (e.g. CC “show notice”)
- Implement this system for the Foundational Spatial Data Frameworks¹ project
- Propose this for use with the data.gov.au ontology
- Have this reviewed by lawyers

¹ http://www.anzlic.gov.au/foundation_spatial_data_framework

Thank you

Geoscience Australia

Nicholas Car
Data Architect

t +61 2 6249 9093
e nicholas.car@ga.gov.au

CSIRO Land and Water

Matthew Stenson
Research Team Leader

t +61 7 3833 5590
e matthew.stenson@csiro.au

LAND AND WATER
www.csiro.au

