



Australian Government



PROVIDING SCIENTIFIC WATER RESOURCE
INFORMATION ASSOCIATED WITH COAL
SEAM GAS AND LARGE COAL MINES

Leveraging QA/QC processes to deliver provenance as a first-order scientific output in large-scale environmental assessments

William Francis, Nicholas J Car, **Becky Schmidt** and Simon Gallant
CSIRO Land and Water, Canberra, Australia

A scientific collaboration between the Department of the Environment, Bureau of Meteorology,
CSIRO and Geoscience Australia



Australian Government



PROVIDING SCIENTIFIC WATER RESOURCE
INFORMATION ASSOCIATED WITH COAL
SEAM GAS AND LARGE COAL MINES

~~Leveraging~~ **Trapezing from** QA/QC processes to deliver provenance as a first-order scientific output in large-scale environmental assessments

William Francis, Nicholas J Car, **Becky Schmidt** and Simon Gallant
CSIRO Land and Water, Canberra, Australia

A scientific collaboration between the Department of the Environment, Bureau of Meteorology,
CSIRO and Geoscience Australia

Trapezing

Publishing integrated scientific information for policy makers

- Use existing embedded QA/QC process

Trapezing into...

- A checking system

That will eventually trapeze into...

- A provenance system



CSIRO Land and Water

Environmental Information Systems Program

Dr Becky Schmidt



Ms Sue Cuddy



Mr Matt Stenson



Dr David Lemon



Knowledge
Delivery

Knowledge
Integration

Systems
Implementation

Interoperable
Systems
data

information

knowledge

Discovery

Access

Integrate

Integrated
Modelling

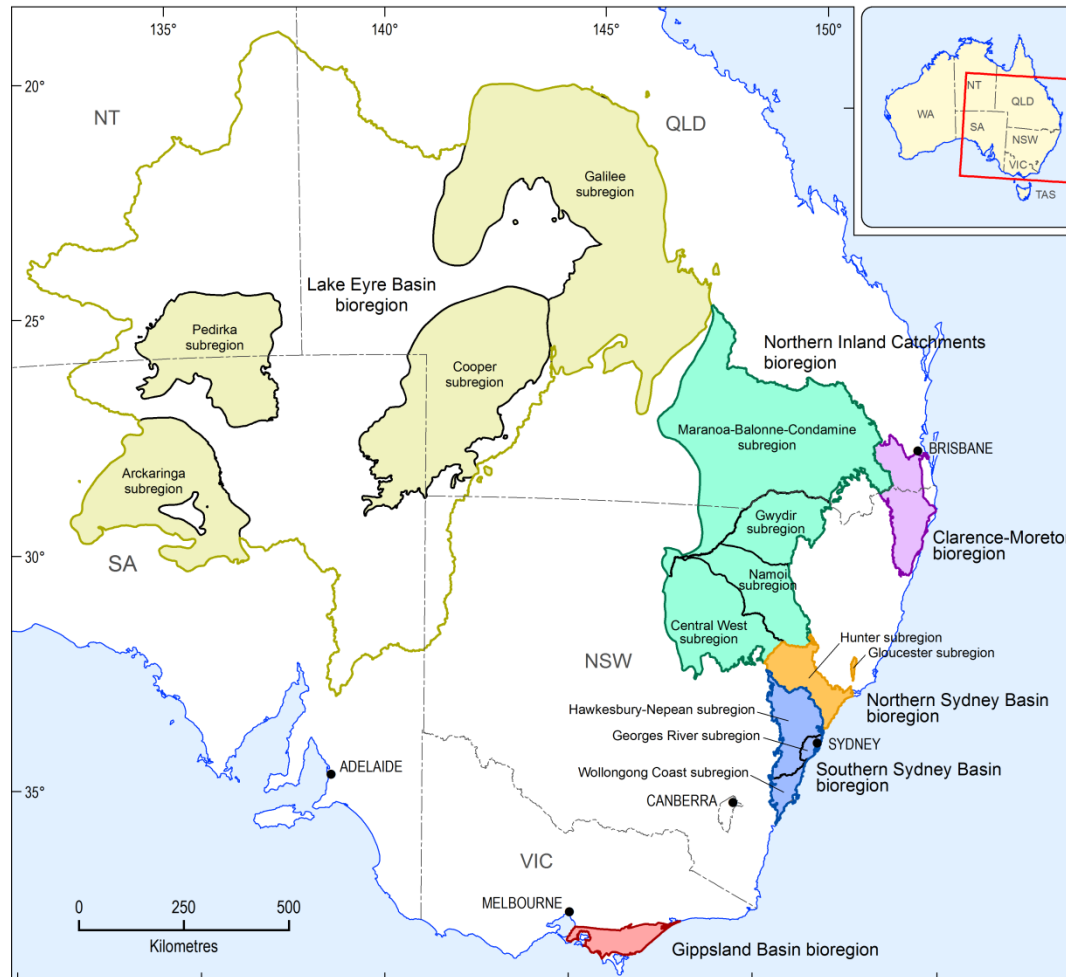
Integration
Assessments

Deliver

Action

The Bioregional Assessment

What are the impacts of coal mining and coal seam gas extraction on water resources?
Programme



What are we delivering?

Scientific advice

- easily digestible and searchable for a policy and public audience
- the BA methodology is the ultimate authority
- customised for each bioregion
- transparent: report uncertainty and provenance

Proponent
applies for
development
approval



OWS Advice
Team gathers
and
synthesises
**supporting
information,**
writes draft
advice



IESC considers
proposal,
**supporting
information,**
synthesis,
draft advice

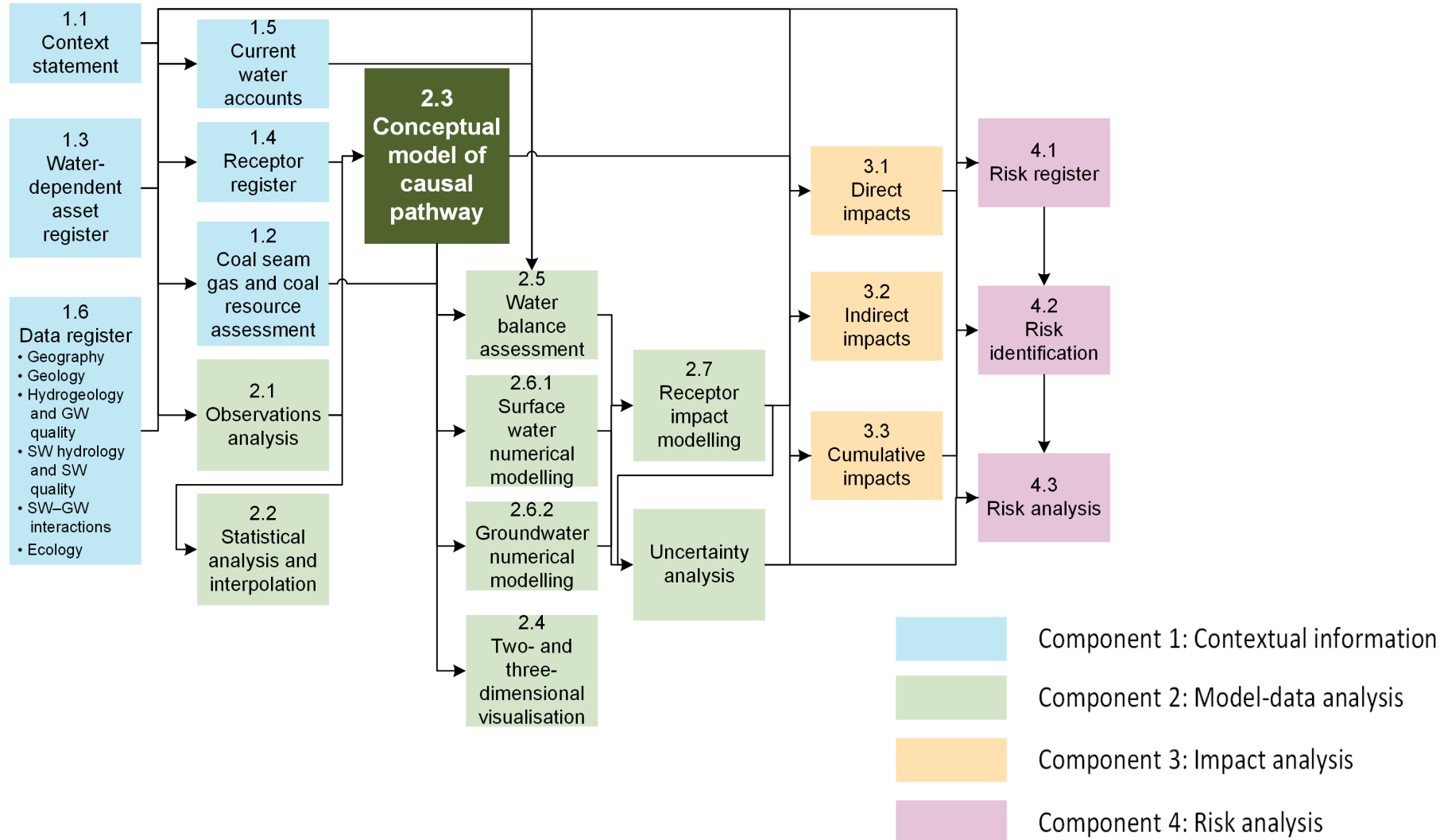


At IESC
meeting,
wordsmith
advice



Minister uses
advice to
make decision
on
development
approval

Products from bioregional assessments



2.7 Receptor impact modelling

D2.7e
What sort of
ecological receptor
impact modelling
should be
undertaken, if any?

Qualitative impact
modelling
Cost \$0.00
FTE 0

Semi-quantitative
impact modelling
Cost \$0.00
FTE 0

Quantitative impact
modelling
Cost \$0.00
FTE 0

Qualitative responses of ecological receptors to
forcing of state variables and fluxes (exposure,
threshold)

Qualitative statements of uncertainty

Semi-qualitative responses of ecological receptors
to forcing of state variables and fluxes (exposure,
threshold)

Qualitative statements of uncertainty

Quantitative responses of
ecological receptors to forcing of
state variables and fluxes
(exposure, threshold)

Variance/covariance of receptors

Cost \$0.00
FTE 0

Synthesise

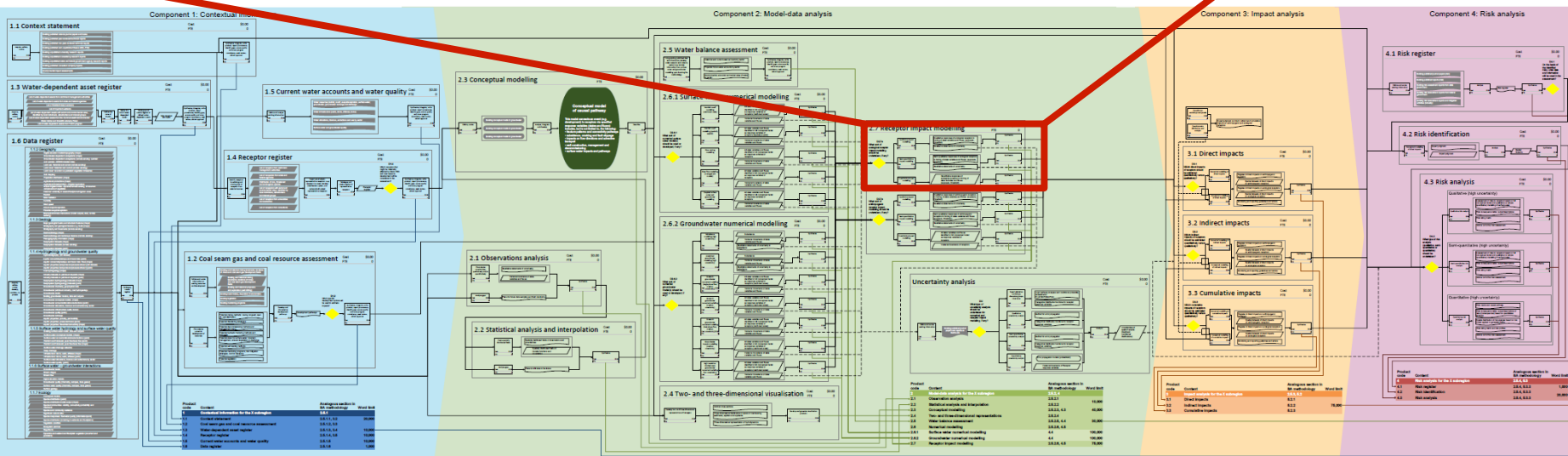
Cost \$0.00
FTE 0

Synthesise

Cost \$0.00
FTE 0

Synthesise

Cost \$0.00
FTE 0



The stats

What are we producing?

- >150 products
- 6,000,000 words
- >2000 elements (maps, charts, diagrams, tables)

Who is we?

- 200 collaborators – authors, editors
- multiple disciplines: ecology, hydrology, geology, hydrogeology
- 4 agencies
- >10 locations

To what standard?

- high editorial and scientific quality
- tight time frames
- provenance and uncertainty



16.10.2007

Existing QA/QC processes

BA-GIP-GIP-112-Geography-v00.docx

1.1.2 Geography

1.1.2 Geography

Summary

Summary of Section 1.1.2 for a public audience. Number of words about 10% of the total words in Section 1.1.2.

1.1.2.1 Physical geography

Physical context, general context and location, climate, landforms and land use

1.1.2.2 Human geography

Population, land use and water use

1.1.2.3 Climate

References

Component 1: Contextual information for the Gippsland bioregion

Existing QA/QC processes

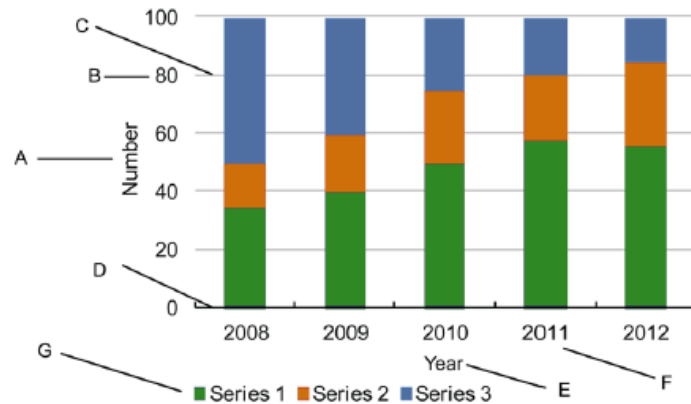


Figure 3 Common elements of a chart. Refer to Table 7 for descriptions and specifications.

Table 7 Specifications for common elements of charts

Element	Name	Specifications
A	y-axis label	Arial, black, 8 pt. Sentence case, rotated 270°
B	y-axis text	Arial, black, 8 pt. Sentence case, right aligned.
C	Gridlines	0.25 pt, grey (R194, G194, B194).
D	x-axis line	0.75 pt, black. Major tick marks inside.
E	x-axis label	Arial, black, 8 pt. Sentence case, centred. (Not applicable to this chart)
F	y-axis text	Arial, black, 8 pt. Sentence case, centred.

Numbers

- ^{18}O
- 4D (use 'one-dimensional' instead)
- ^{222}Rn
- 2D (use 'two-dimensional' instead)
- ^2H
- 3D (use 'three-dimensional' instead)
- ^4He
- ^{86}Sr
- ^{87}Sr

A

A Directory of Important Wetlands in Australia

activities (use 'development' instead in phrases such as 'impacts of coal seam gas and large coal mining development on water resources')

Acts (see 'legislation' in Table 3)

- actual evapotranspiration (AET)
- airborne electromagnetic (AEM)
- American Petroleum Institute units (API units)
- animals: common names (lowercase, do not italicise in text)
- animals: species names (italicise in text)
- anthropogenic receptor
- aquifer
- aquitard
- ArcGIS
- Arckaringa Basin
- Arckaringa subregion (in Lake Eyre Basin bioregion, do not shorten)
- artesian aquifer
- AS/NZS ISO 31000:2009 Risk management – principles and guidelines (on first mention, then subsequently 'the ISO 31000:2009 standard')

Products have elements

Context statement for the Galilee subregion

Product 1.1 from the Lake Eyre Basin Bioregional Assessment

28 May 2014



A scientific collaboration between the Department of the Environment,
 Bureau of Meteorology, CSIRO and Geoscience Australia

Component 1: Contextual information for the Galilee subregion

1.1.4 Hydrogeology and groundwater quality

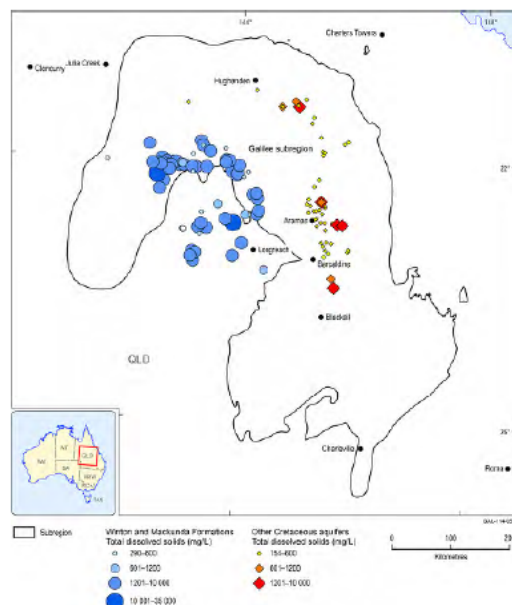


Figure 28 Groundwater quality of the Winton and Mackunda formation aquifers and other Cretaceous aquifers of the Eromanga Basin

Source data: RPS (2012) Appendix E

72 | Context statement for the Galilee subregion

1.1.4 Hydrogeology and groundwater quality

1.1.4.2.4 Cenozoic aquifers

The Cenozoic aquifers, which include the Quaternary alluvium and other Cenozoic sediments, are important groundwater resources in the subregion. In the RPS (2012) dataset, the groundwater sample depth ranged from less than 10 m to approximately 150 m (Figure 29), with most Quaternary alluvium sampled at depths of less than 30 m and the other Cenozoic aquifers sampled between 30 and 140 m. RPS (2012) suggested that there are probably at least twice as many bores as those shown in Figure 30 tapping into the Cenozoic aquifers.

The water quality of the Cenozoic alluvial aquifers ranges from fresh to saline (minimum and maximum of 48 to 13,618 mg/L TDS respectively). According to the Australian Drinking Water Guidelines classification (NHMRC and NRMCC, 2011), most groundwater in the Cenozoic aquifers is classed as fresh (<600 mg/L TDS) or fair to poor (600–1,200 mg/L TDS), with a median of 492 mg/L and a mean of 1057 mg/L TDS (Figure 29).

The water quality in the Quaternary alluvium aquifer shows an increase in salinity with depths (Figure 29), from very fresh (<100 mg/L TDS) to brackish (~3000 mg/L TDS). This may represent the chemical evolution of the groundwater as it flows from shallow recharge areas to deeper parts (~30 m depth) of the alluvium. In comparison, the water quality from other Cenozoic aquifers does not exhibit any distinct relation with depth or spatial pattern (Figure 30).

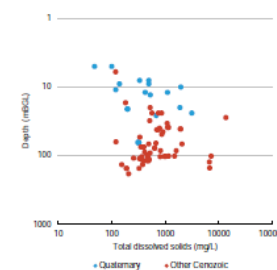


Figure 29 Graph of depth and groundwater quality for Cenozoic aquifers

Source data: RPS (2012) Appendix E

Context statement for the Galilee subregion | 73

Managing data and elements

BIOREGIONAL ASSESSMENTS

```
graph TD; MR[Modelling runs] --> D[Data]; D --> EW[Excel workbook]; EW --> PT[Plots Tables]; PT --> R[Report]; R --> EW; Note1[Base document links plots/tables in report to workbook]; Note2[If errors are found in plots in report, regenerate from workbook and then paste again into report to maintain audit trail.]
```

Modelling runs

Data

Excel workbook

Plots Tables

Report

Base document links plots/tables in report to workbook

If errors are found in plots in report, regenerate from workbook and then paste again into report to maintain audit trail.



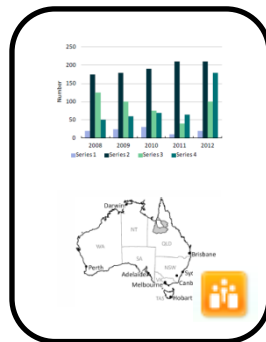
What are workbooks?

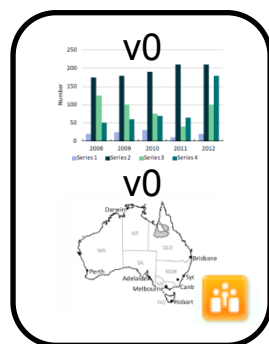
Element number	GIP-112-001
Data source	
Image source	
Other source	
Last updated	
Updated by	
Changes	
Element type	
Figure/Table number	
Caption	
Alt text	
Copyright approval needed?	
Copyright approval status	
Link to copyright documentation	
Notes	
Go to ControlSheet Go to WorkbookLog	

ELEMENT
GOES
HERE!

Instructions / WorkbookLog / ControlSheet / AAA-000-001 / **GIP-112-001** / GIP-112-002 / GIP-112-003 / GIP-112-004 / GIP-112-005

Elements





Ch1 v0



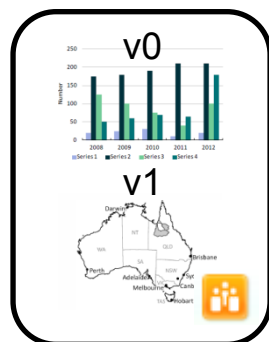
Ch2 v0



Ch3 v0



12 Feb



Ch1 v3



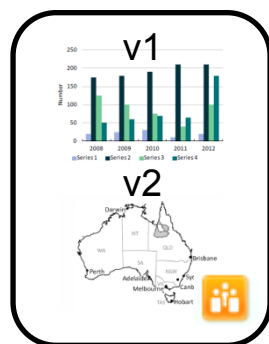
Ch2 v1



Ch3 v0



13 Feb



Ch1 v5



Ch2 v2



Ch3 v1



14 Feb

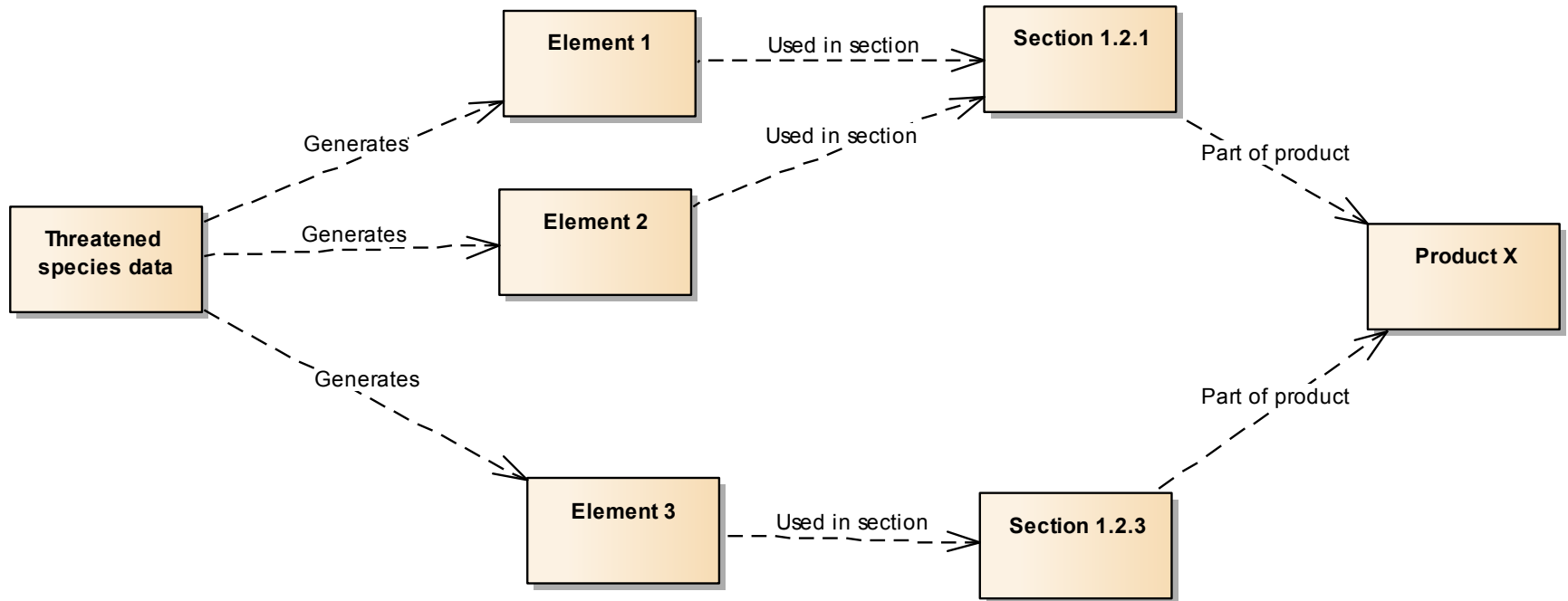
Why are we managing elements?

- Getting things done
- Quality assurance
 - Internal and cross-product consistency
- Change management
 - Which version are we using in the report, who changed it, why?
- Provenance/data management requirements
 - 3 months later, where did that number come from?

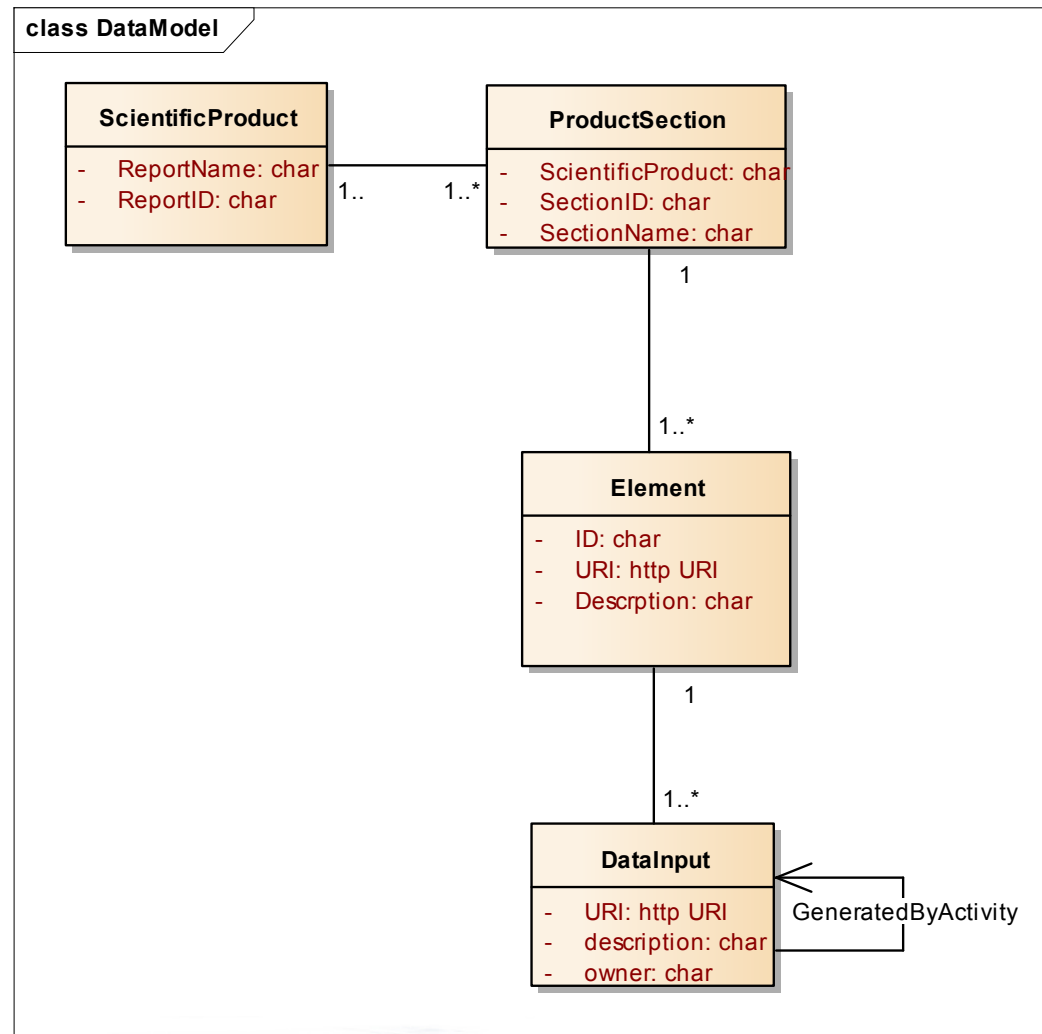


Workflow

act Example data workflow



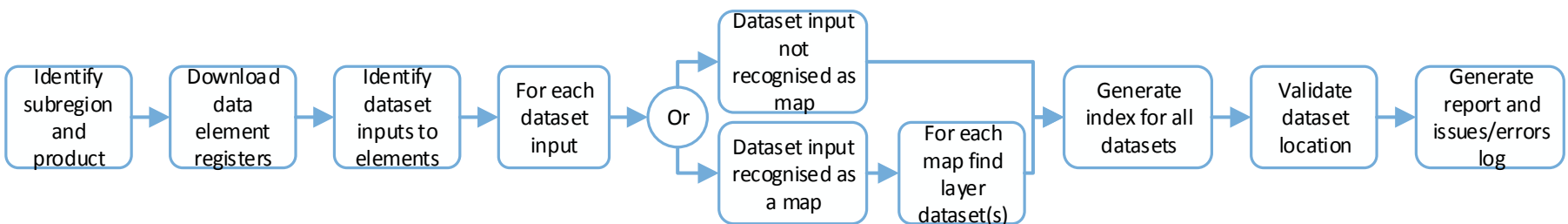
Data model



Provenance Reporting Tool

Reads element registers and delivers:

- Provenance information for each element
- Validation of dataset use
- Quality statistics
- Issue log



Output of tool in pseudo PROV-O format

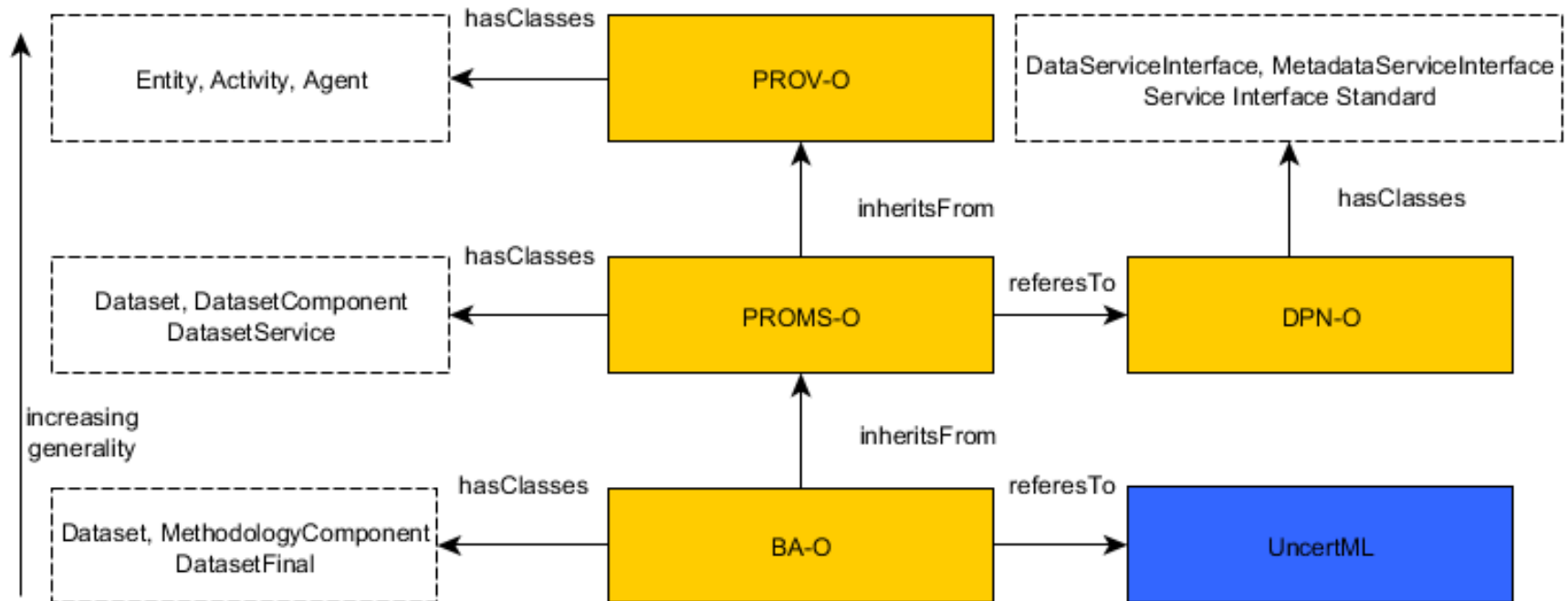
```
:Dataset_5
  a      Entity;
  value"
  \\wron\project\BA\NIC\AI\Maps\Templates\Shapefiles\State\_Borders\_10million.shp";
  title  State_Borders_10million;
.
:Activity_1
  a      Activity;
  used   Dataset_8;
  used   Dataset_2;
  used   Dataset_5;
  used   Dataset_10;
  used   Dataset_11;
  used   Dataset_14;
  used   Dataset_26;
  used   Dataset_16;
  used   Dataset_17;
  generated Element_CEN-112-009;
.
:Activity_11
  a      Activity;
  used   Element_CEN-111-001;
  ...
  used   Element_CEN-112-009;
  generated CEN_Product_1-1;
.
```



Leveraging new data management systems

- To store and query provenance information
- DIDS – Data Identity System
- PROMS – Provenance Management System
- Requirements – during project
 - Tracking of product evolution
 - What are all the downstream impacts when science modelling is rerun (new data, corrections, etc.)?
- Requirements – post-project
 - Audit trail.
- Goal: centralise element information and query for product specific information for logs/workbooks.

Provenance



Using the existing QA/QC

Element number	GIP-112-001
Data source	
Image source	
Other source	
Last updated	
Updated by	
Changes	
Element type	
Figure/Table number	
Caption	
Alt text	
Copyright approval needed?	
Copyright approval status	
Link to copyright documentation	
Notes	
Go to ControlSheet Go to WorkbookLog	

ELEMENT
GOES
HERE!

Press this button to output to DIDS
and PROMS and get URI

< ▶ ▶ Instructions / WorkbookLog / ControlSheet / AAA-000-001 / **GIP-112-001** / GIP-112-002 / GIP-112-003 / GIP-112-004 / GIP-112-005

Benefits of provenance reporting

- Increased transparency
- Status reported more easily
 - Benchmarks and checkpoints for product approvals
- Data management improved
- By using existing process, demystifies data models and workflows and provenance for researchers – so that they are more open to more sophisticated solutions on horizon

Feedback loop: QA/QC processes improved

- Automation increases quality
- Measure quality improvements
- URIs and vocabulary services improve quality of products
- Tagging prepares for delivering via information platform (not just reports)



CSIRO

Becky Schmidt

Research Team Leader, Knowledge Delivery Team

t +61 2 6246 5923

e Becky.Schmidt@csiro.au

w <http://www.csiro.au/Organisation-Structure/Divisions/Land-and-Water/BeckySchmidt.aspx>



Australian Government
Department of the Environment
Bureau of Meteorology
Geoscience Australia



www.bioregionalassessments.gov.au