

Semantic Web data Integration for Australia

By Nicholas Car
For DAMA, 26 Nov 2025



KurrawongAI

<https://kurrawong.ai>

A bit about me and my goals...

- I'm a data modeller
- I'm keen on large-scale data integration initiatives
- I'm trying to build large data collections of enduring value for Australia
 - Because I see a lot of waste through data loss, unnecessary tool expenses
- I have a long but arms-length association with DAMA



Outline

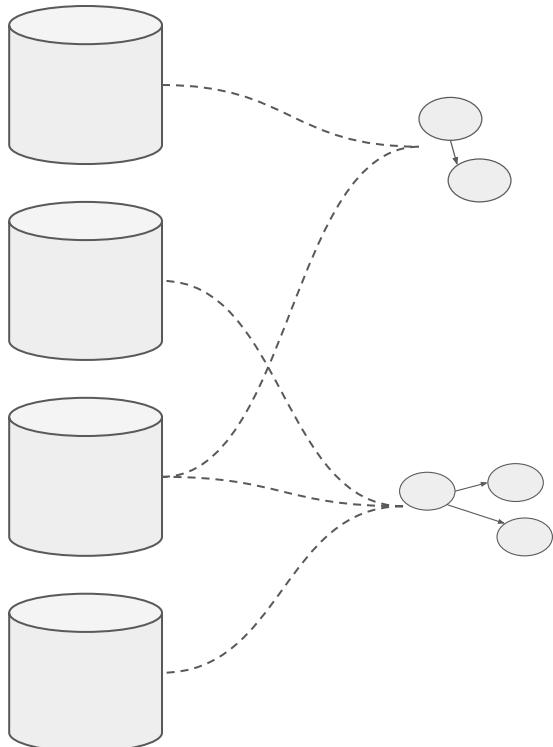
1. Recap: how to do data integration
2. What is the Semantic Web, really?
3. How SW does data integration
 - o And how it's different/better than other systems
4. Large-scale Australian Knowledge Graphs



KurrawongAI

Recap: how to do data integration

General data integration approach



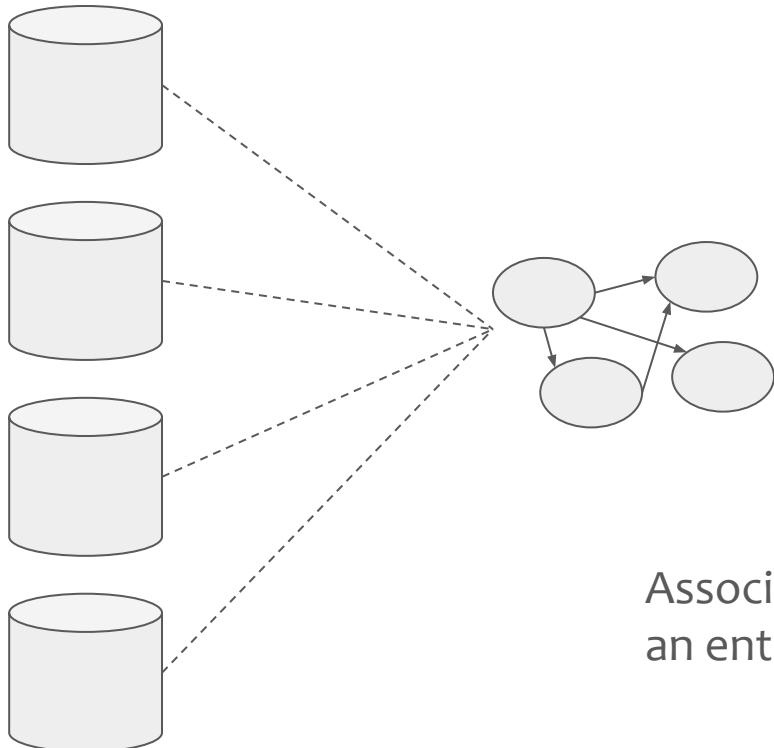
Reuse modelling patterns



KurrawongAI

Recap: how to do data integration

General data integration approach

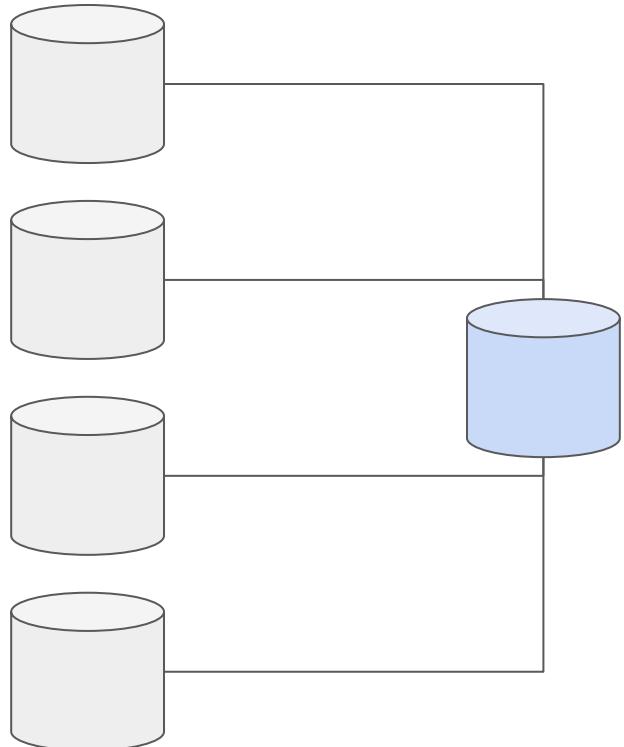


Associate dataset elements with an enterprise data model



Recap: how to do data integration

General data integration approach



Use Reference Data



KurrawongAI

Recap: how to do data integration

Assuming we have to make the enterprise data model:

- Where do we get the modelling patterns from?
- Are there any sources of general Reference Data?

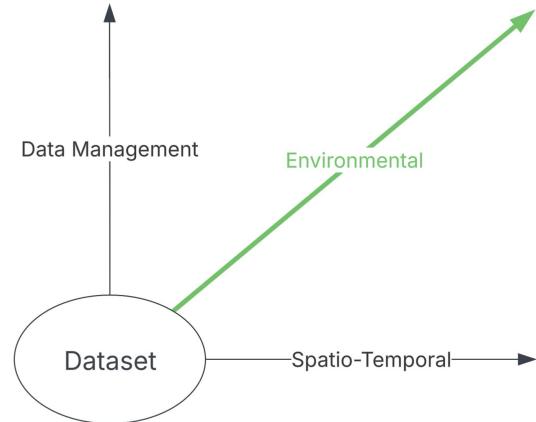


KurrawongAI

Recap: how to do data integration

Imagine trying to integrate *Environmental Information Australia's* datasets...

Integration dimensions may look like this:



Recap: how to do data integration

Assuming we have to make the enterprise data model:

- Where do we get the modelling patterns from?
- Are there any sources of general Reference Data?



KurrawongAI

What is the Semantic Web, *really*?

As originally specified, https://en.wikipedia.org/wiki/Semantic_Web:

“The goal of the Semantic Web is to make Internet data machine-readable.”



What is the Semantic Web, *really*?

As originally specified, https://en.wikipedia.org/wiki/Semantic_Web:

“The goal of the Semantic Web is to make Internet data machine-readable,
understandable



What is the Semantic Web, *really*?

As originally specified, https://en.wikipedia.org/wiki/Semantic_Web:

“The goal of the Semantic Web is to make Internet data machine-readable.”

understandable

- Use universal definitions for things - data, metadata, schema
- Allow universal lookups for more info



What is the Semantic Web, *really*?

As originally specified, https://en.wikipedia.org/wiki/Semantic_Web:

“The goal of the Semantic Web is to make Internet data machine-readable.”

understandable

- Use universal definitions for things - data, metadata, schema
- Allow universal lookups for more info

There are a bunch of technical things required for this but we are keeping to the conceptual today, assume they all work...!



KurrawongAI

What is the Semantic Web, *really*?

As originally specified, https://en.wikipedia.org/wiki/Semantic_Web:

“The goal of the Semantic Web is to make Internet data machine-readable.”

understandable

- Use universal definitions for things - data, metadata, schema
- Allow universal lookups for more info

Where do these definitions all come from?



KurrawongAI

What is the Semantic Web, *really*?

We now have a large body of SW models (ontologies)



ORG

SSN



DCA

PROV-OT

GeoSPARQL

Schema.org



OBO Foundry



What is the Semantic Web, really?

We now have a large body of SW models (ontologies)

How are these different from other bodies of models?

- It's not (just) about the tech
- It's about the intention
 - universal, comprehensive, free



.edu



KurrawongAI

What is the Semantic Web, really?

We now have a large body of SW models (ontologies)

How are these different from other bodies of models?

- It's not (just) about the tech
- It's about the intention
 - universal, comprehensive, free

Mention Linus Torvald's remark
about the Linux licensing here



.edu

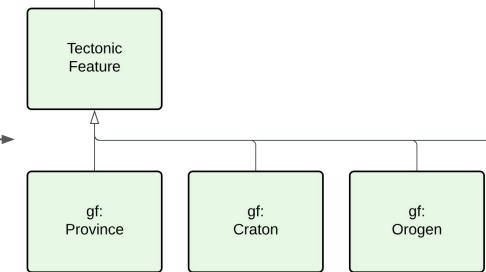


What is the Semantic Web, *really*?

The universality + comprehensive intention has lead to a fundamentals approach I don't think we see elsewhere:

“Upper” ontologies through to “domain” ones, to profiles

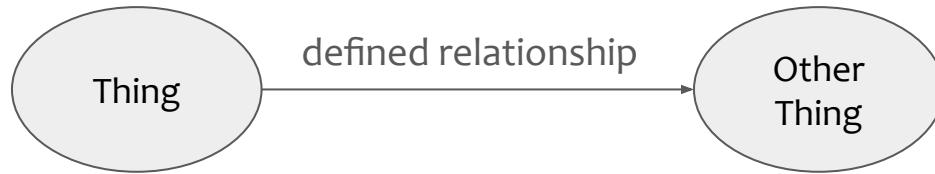
BFO → GeoSPARQL



What is the Semantic Web, *really*?

The universality + comprehensive intention has lead to a fundamentals approach I don't think we see elsewhere:

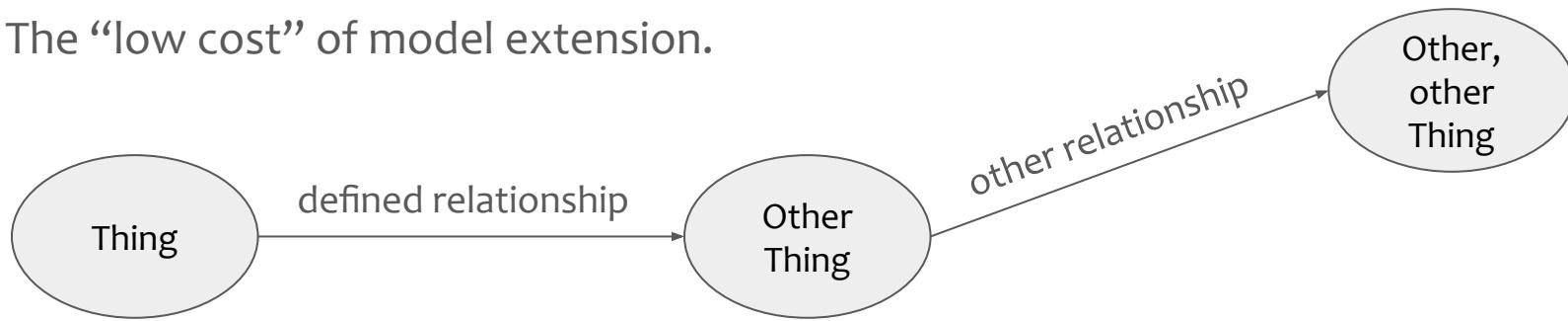
The “low cost” of model extension.



What is the Semantic Web, *really*?

The universality + comprehensive intention has lead to a fundamentals approach I don't think we see elsewhere:

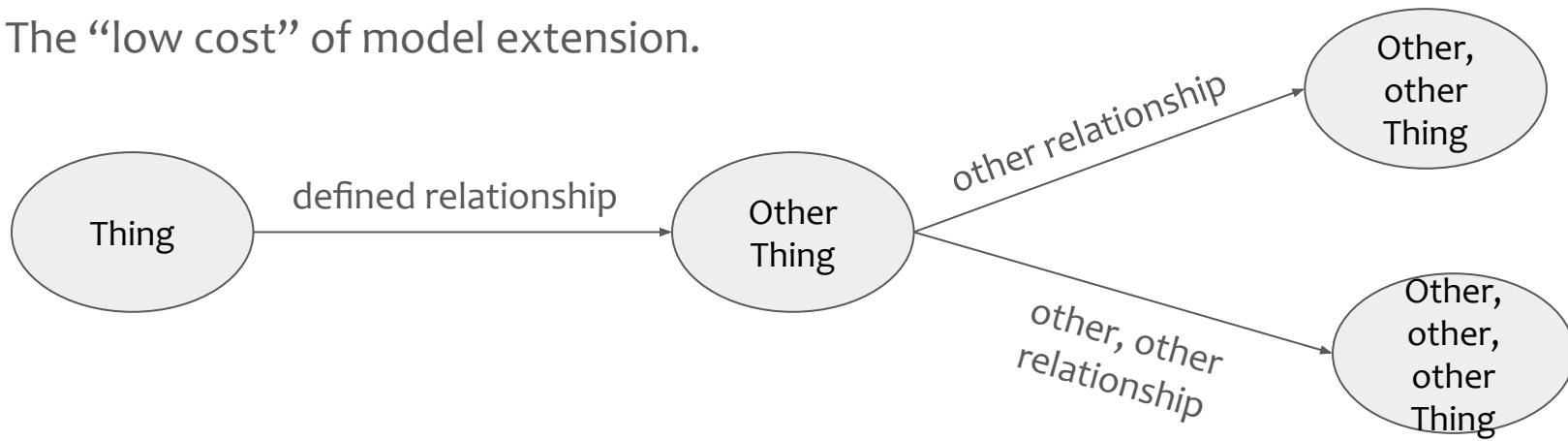
The “low cost” of model extension.



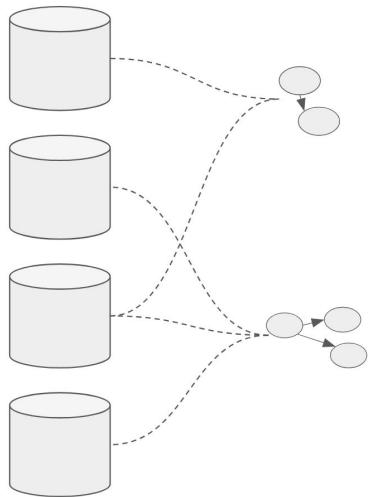
What is the Semantic Web, *really*?

The universality + comprehensive intention has lead to a fundamentals approach I don't think we see elsewhere:

The “low cost” of model extension.



How SW does data integration



from



SSN



ORG

DCA

PROV-OT

GeoSPARQL

Schema.org

BFO



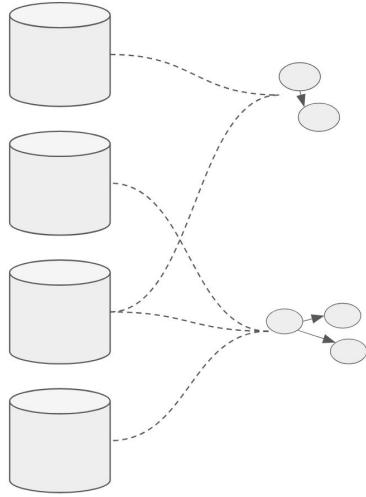
OBO Foundry

Fundamental & integrated models



KurrawongAI

How SW does data integration

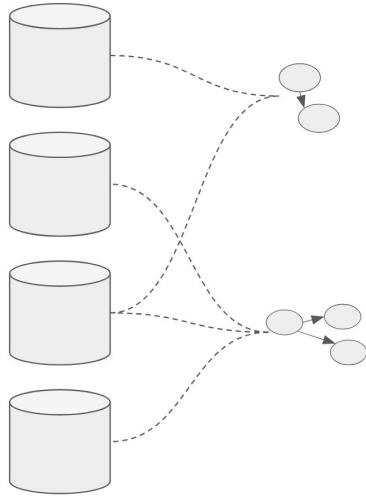


How are these models different from other collections of integrated models?

- e.g. ISO 19*, OMG standards...

Modelling patterns

How SW does data integration



How are these models different from other collections of integrated models?

- e.g. ISO 19*, OMG standards...

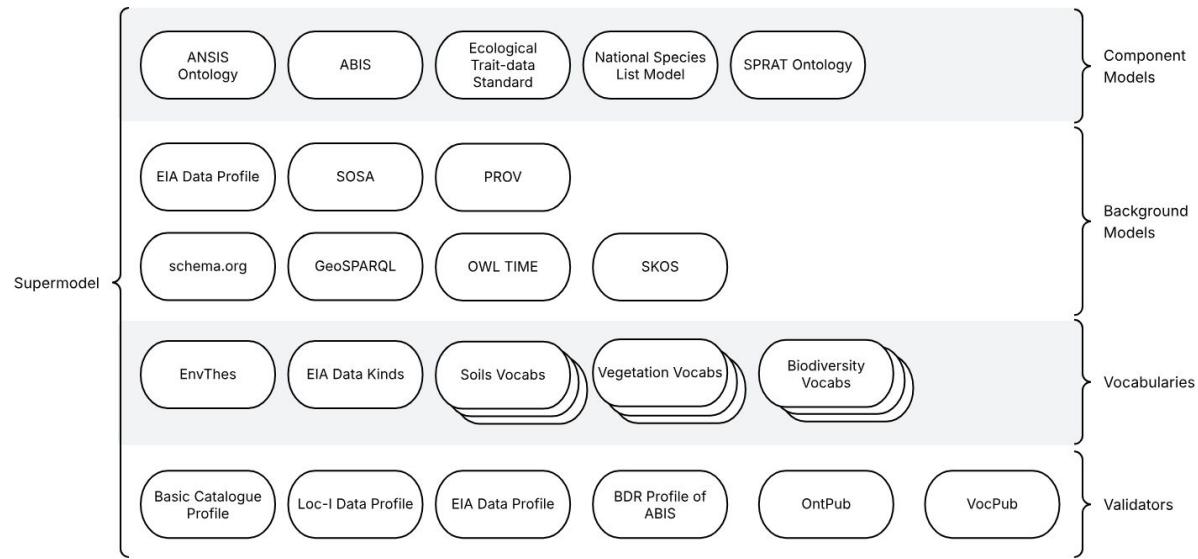
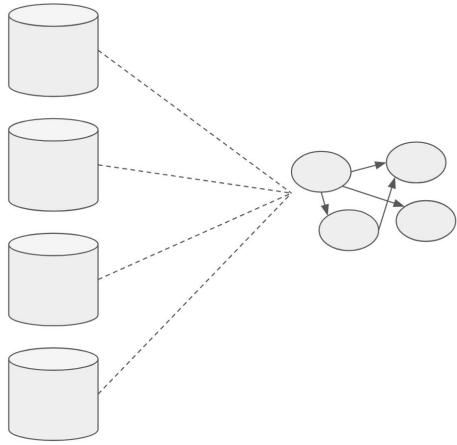
Intention: ... universal, comprehensive, free...

Tech:

- Tighter conceptual modelling → orderly implementation
- Validation

Modelling patterns

How SW does data integration



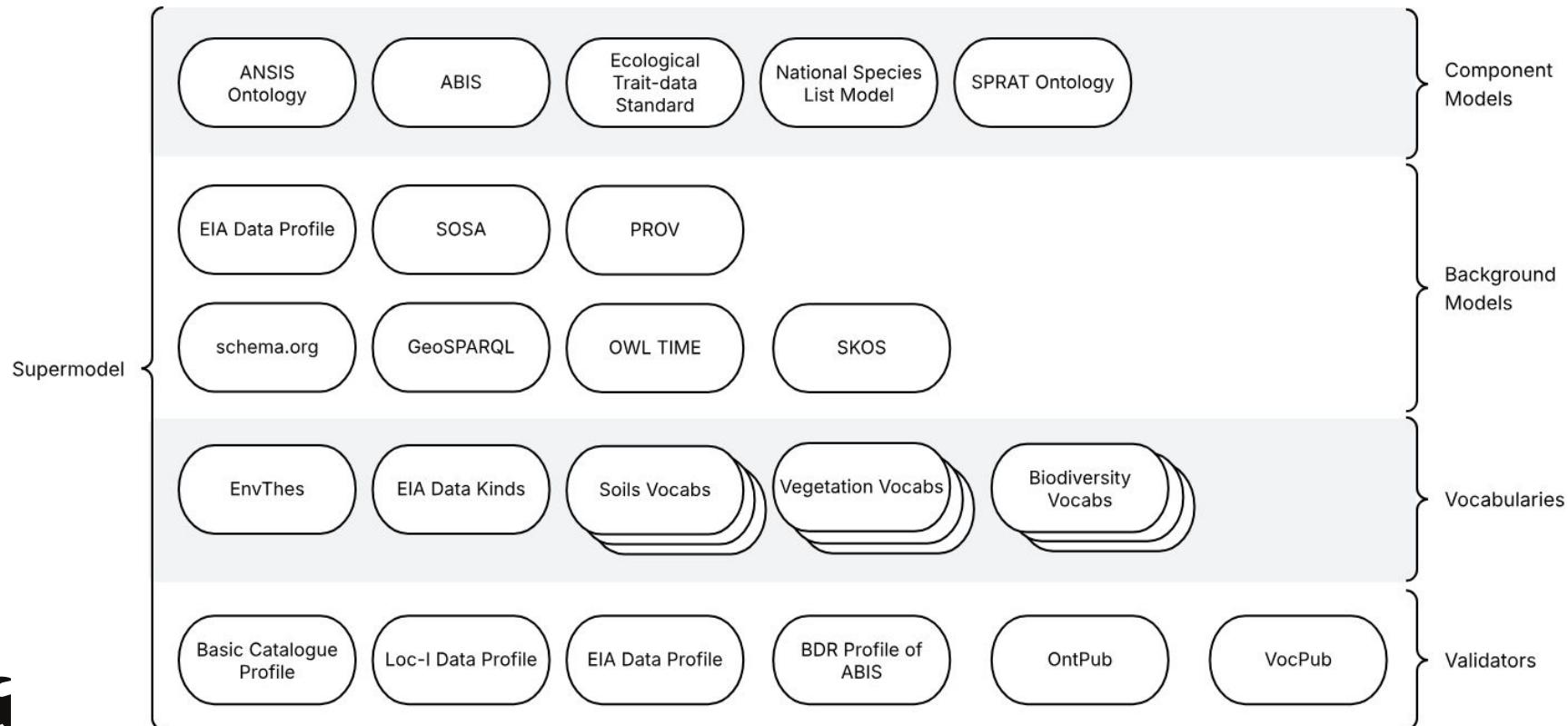
Enterprise data model

Independently-modelled data magically lines up,
if using the same background models - EIA
experience

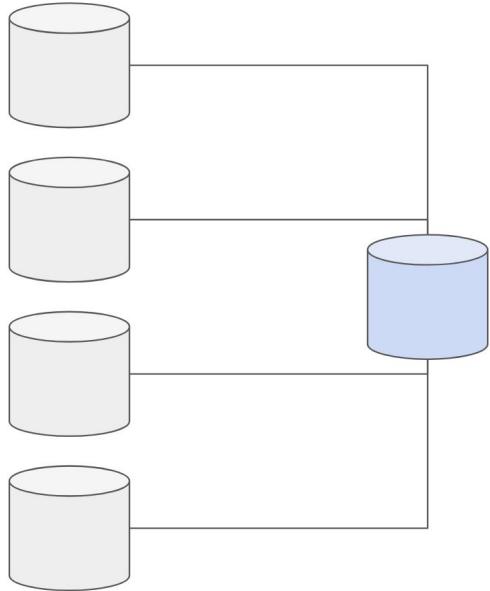


KurrawongAI

How SW does data integration



How SW does data integration



Use Reference Data



KurrawongAI

Mechanisms: Linked Data. You can go to it directly.

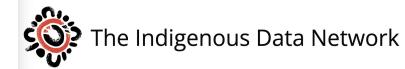
Instances:

The screenshot shows a search interface for 'fossil'. The search bar contains 'fossil'. Below it, there are tabs for 'All (89)', 'Vocabularies (2)', 'Collections (0)', and 'Concepts (87)'. The 'Concepts' tab is selected. A sidebar on the left lists 'Endpoints' with checkboxes for 'DEMO', 'GSWA', 'GSSA', 'OGC', 'IDN', 'GA', 'CGI', and 'GGIC'. The main content area shows search results for 'fossil':

- fossil**
https://pid.geoscience.gov.au/def/voc/ga/GeologySample
A specimen of remains or traces of a plant or animal preserved
Score: 100 Has Children
- Fossil Downs Station Illa**
https://data.id.au.org/pid/nntt/WI2016-814
Score: 100
- FOSSIL FUEL BURNING**
http://gdservices.gsfc.nasa.gov/kms/concept/edfbff54-e1b2-3d4d-74f9

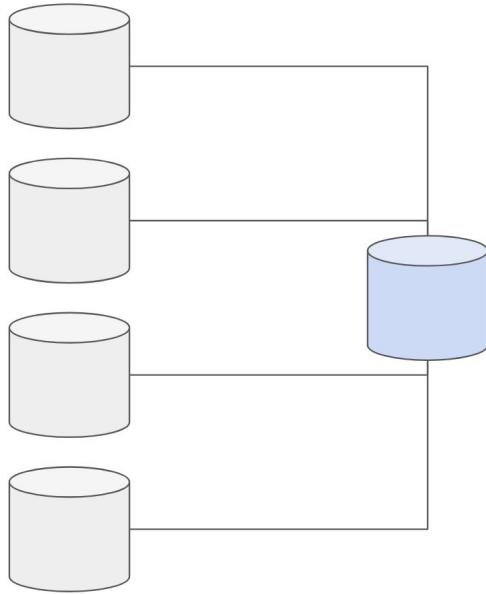


The screenshot shows the 'Collections' page of the Australian Government website. The header includes the Australian coat of arms and the text 'Australian Government' and 'Department of Climate Change, Energy, the Environment and Water'. Below the header, there are links for 'Home', 'Catalogs', 'EIA Test Dataset Catalogue', 'Collections', and 'Search'. The main content area is titled 'Collections'.



OGC RAINBOW

How SW does data integration

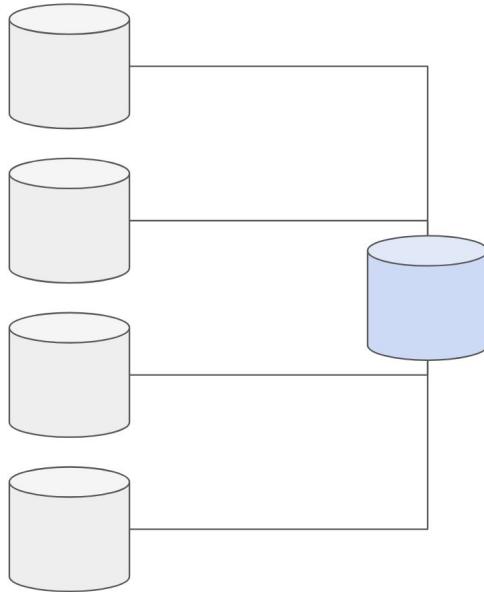


Use Reference Data

The screenshot shows the KurrawongAI search interface. The search bar contains the query "fossil". Below the search bar, there are tabs for "All (89)", "Vocabularies (2)", "Collections (0)", and "Concepts (87)", with "Concepts" being the active tab. A section titled "Endpoints" shows "87 Concepts from 4 endpoints" with a list of sources: DEMO, GSWA, GSSA, OGC, IDN, GA, CGI, and GGIC, each with a checkmark. The main search results for "fossil" include:

- fossil**
<https://pid.geoscience.gov.au/def/voc/ga/GeologySample>
A specimen of remains or traces of a plant or animal preserved.
Score: 100 Has Children
- Fossil Downs Station Ilua**
<https://data.idnau.org/pid/nntt/WI2016-014>
Score: 100
- FOSSIL FUEL BURNING**
<http://gcmdservices.gsfc.nasa.gov/kms/concept/edfbff154-e1823b4d7f49>

How SW does data integration



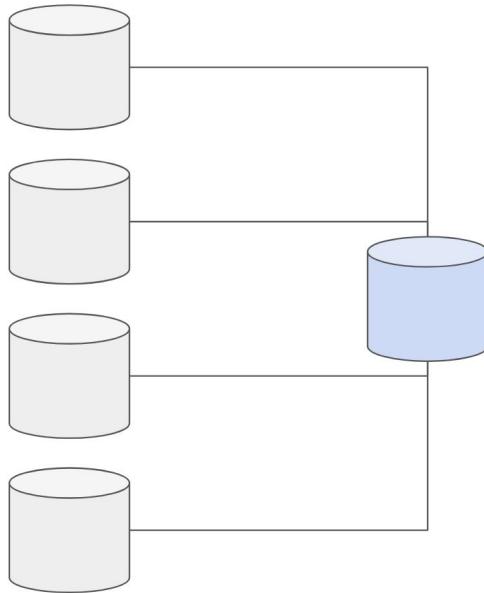
Use Reference Data

The screenshot shows a dark blue header with the Australian Government logo and the text "Department of Climate Change, Energy, the Environment and Water". Below the header is a navigation bar with links for "Home", "Catalogs", "Search", "SPARQL", and "Pro". The main content area has a light gray background. The breadcrumb navigation shows "Home / Catalogs / EIA Test Dataset Catalogue / Collections". The title "Collections" is displayed in a large, bold, dark font. Below the title is a section titled "Item" with a list of dataset links:

- [Australian National Soil Information System \(ANSIS\) Dataset ↗](#)
- [AusTraits to NSL Mapping Linkset ↗](#)
- [Biodiversity Data Repository Dataset for the ACT ↗](#)
- [Habitat Condition and Assessment System \(HCAS\) Dataset ↗](#)



How SW does data integration



Use Reference Data

The Indigenous Data Network

Home Catalogues Search SPARQL About API Documentation

Home > Catalogues > IDN Spatial Reference Data > Collections

Collections

Item

Indigenous Languages Multi-map ↗
Multiple indigenous language maps aggregated into one dataset by the IDN

AIATSIS Place Thesaurus ↗
The AIATSIS Place Thesaurus contains headings for place names, using the Indigenous place names and place names in English.

Collaborative Australian Protected Areas Database (CAPAD) 2020 ↗
The Collaborative Australian Protected Areas Database (CAPAD) 2020 provides both spatial and non-spatial information on Australian protected areas.

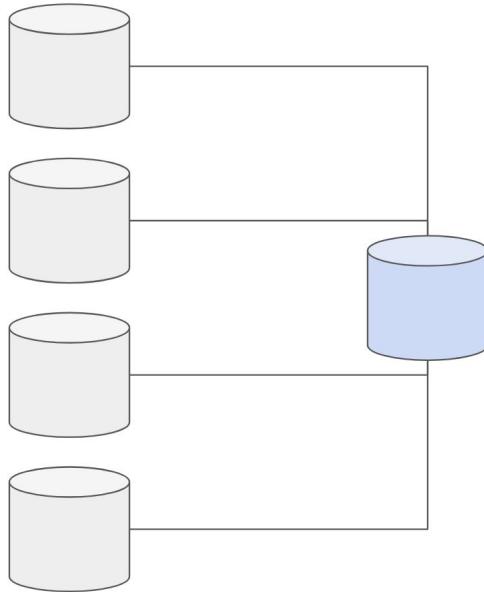
Australian Government Indigenous Programs & Policy Locations (AGIL) dataset ↗
This dataset has been developed by the Australian Government as an authoritative source of information on Indigenous programs and policy locations.

Indigenous Structures within the ASGS Ed. 3 ↗



KurrawongAI

How SW does data integration



Use Reference Data

The Indigenous Data Network (IDN) Catalogues page. The URL is <https://data.idnau.org/pid/atns/feature/0321>. The item is a Feature. The map shows the location of the feature, with labels for Rungula National Park and Giringun Indigenous Protected Area.

The Indigenous Data Network

Home Catalogues Search SPARQL About API Documentation

Home > Catalogues > IDN Spatial Reference Data > Collections > ATNS Spatial Features > Items > ATNS Entity Areas > Features > Union Mining, the Ewamian People and the State of Queensland Agreement (November (1997))

Union Mining, the Ewamian People and the State of Queensland Agreement (November (1997))

IRI <https://data.idnau.org/pid/atns/feature/0321> [Open in new tab](#)

Type Feature

Rungula National Park

Giringun Indigenous Protected Area

Large-scale Australian Knowledge Graphs

GSWA Supermodel

Search

The Supermodel

Overview

This is a documentation website for a Supermodel: a multi-part enterprise data model.

This particular Supermodel is for a public expression of the data holdings of the [Geological Survey of Western Australia](#).

Question	Answer
What is a Supermodel, in detail?	See the next few sections See the generic description of a Supermodel
What is the status of this Supermodel?	See Status below
What other Supermodels?	See the Related Supermodels section of this Supermodel's description

Table of contents

Overview

Introduction

Related Supermodels

Supermodel Structure

Technical Assets

Supermodel Definition

Modelling Documentation

Diagram Conventions

Namespaces

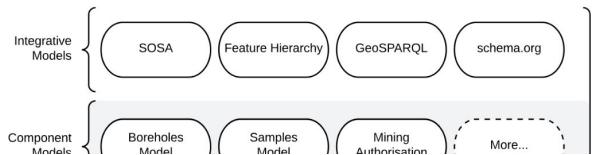
RDF code

SPARQL queries

Status

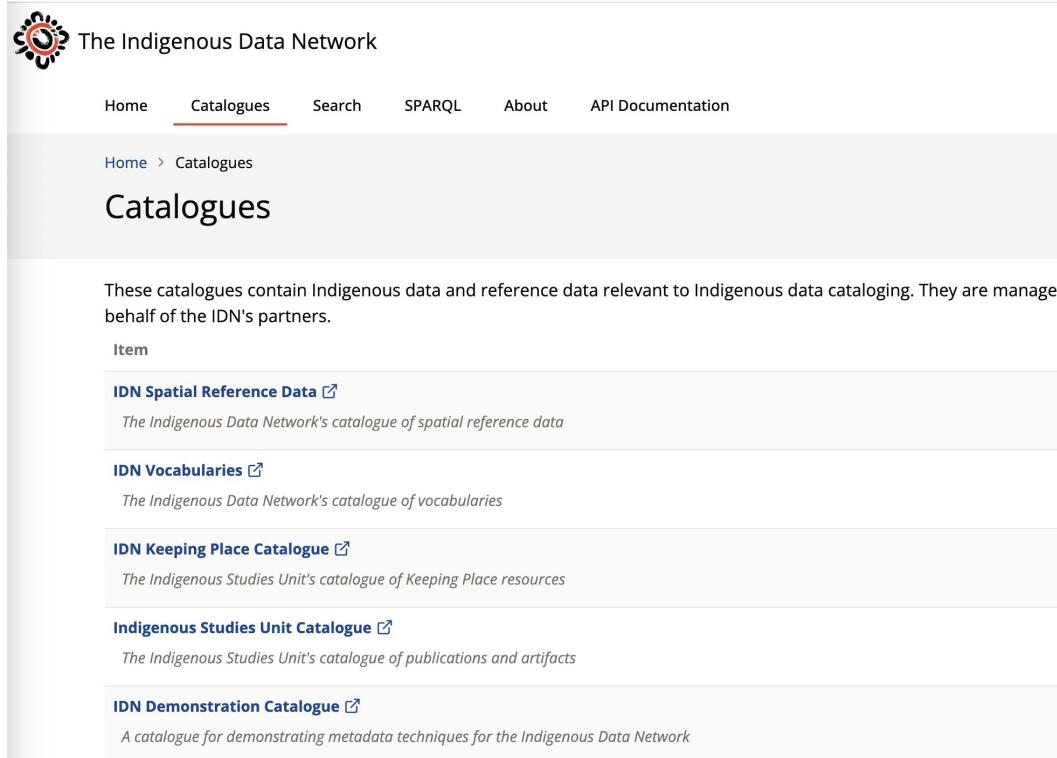
License & Rights

Contacts



Geological Survey of WA

Large-scale Australian Knowledge Graphs



The Indigenous Data Network

Home Catalogues Search SPARQL About API Documentation

Home > Catalogues

Catalogues

These catalogues contain Indigenous data and reference data relevant to Indigenous data cataloging. They are managed | behalf of the IDN's partners.

Item

IDN Spatial Reference Data 

The Indigenous Data Network's catalogue of spatial reference data

IDN Vocabularies 

The Indigenous Data Network's catalogue of vocabularies

IDN Keeping Place Catalogue 

The Indigenous Studies Unit's catalogue of Keeping Place resources

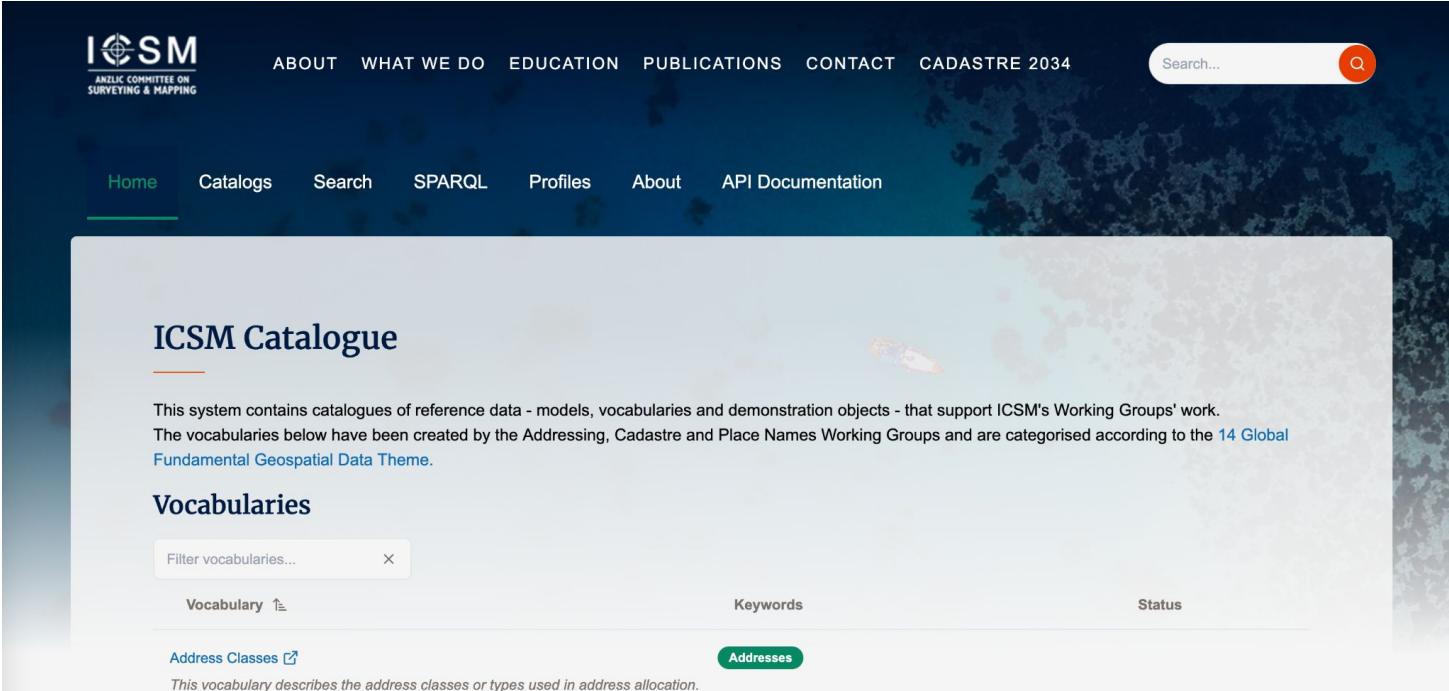
Indigenous Studies Unit Catalogue 

The Indigenous Studies Unit's catalogue of publications and artifacts

IDN Demonstration Catalogue 

A catalogue for demonstrating metadata techniques for the Indigenous Data Network

Large-scale Australian Knowledge Graphs



The screenshot shows the ICSM Catalogue interface. At the top, there is a navigation bar with links for Home, Catalogs, Search, SPARQL, Profiles, About, and API Documentation. The 'Home' link is underlined, indicating it is the active page. Below the navigation bar, the title 'ICSM Catalogue' is displayed. A sub-section titled 'Vocabularies' is shown, featuring a search bar labeled 'Filter vocabularies...'. The vocabulary list includes 'Address Classes' (with a description: 'This vocabulary describes the address classes or types used in address allocation.') and 'Addresses'. The 'Addresses' entry is highlighted with a green button-like background.

ICSM Catalogue

This system contains catalogues of reference data - models, vocabularies and demonstration objects - that support ICSM's Working Groups' work. The vocabularies below have been created by the Addressing, Cadastre and Place Names Working Groups and are categorised according to the 14 Global Fundamental Geospatial Data Theme.

Vocabularies

Filter vocabularies...

Vocabulary	Keywords	Status
Address Classes 	Addresses	

This vocabulary describes the address classes or types used in address allocation.

Intergovernmental Committee on Surveying & Mapping - distributed

Large-scale Australian Knowledge Graphs



Intergovernmental Committee on Surveying & Mapping - distributed

Large-scale Australian Knowledge Graphs



Australian Government

Department of Climate Change, Energy,
the Environment and Water

Biodiversity Data Repository



Biodiversity
Data
Repository

Environment Information Australia

[Home](#) [Resources](#) [Submit](#)

Home

The Biodiversity Data Repository (BDR) is an initiative to centralise and improve the management of Australia's biodiversity information.

Using Knowledge Graph technology, the BDR aims to consolidate biodiversity data currently scattered across government, industry, and research institutions. It supports Australia's [Nature Positive](#) agenda by making biodiversity data more accessible, accurate, and interoperable with existing environmental information systems.

Biodiversity Data Repository



KurrawongAI

Large-scale Australian Knowledge Graphs

EIA Test Catalogue

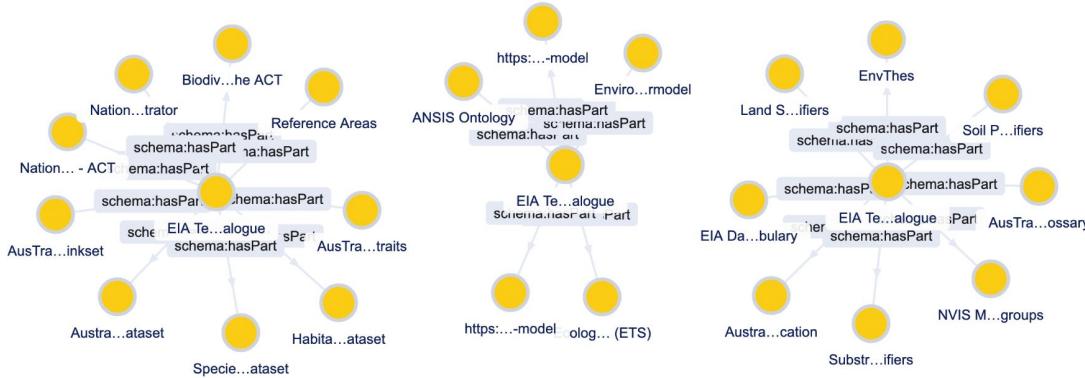
This catalogue was created in mid-2025 to showcase the integration of [Environment Information Australia \(EIA\)](#) datasets.

The catalogue lists test versions of major Australian environmental datasets within EIA's scope. It also lists models and vocabularies needed to support the integration of the datasets.

The catalogue tool also provides the [EIA Scenario Demonstrator](#) which describes a series of data discovery scenarios that show off different aspects of the dataset's integration using [Semantic Web](#) and [Knowledge Graph](#) methods.

The enterprise data model - a "Supermodel" - within which all elements of this catalogue are positioned is online at:

<https://linked.data.gov.au/def/eia-supermodel>



Large-scale Australian Knowledge Graphs

GSWA + IDN + ICSM + BDR + EIA ?



KurrawongAI

Large-scale Australian Knowledge Graphs

GSWA + IDN + ICSM + BDR + EIA ?



Which of these would
you need to underpin
your Digital Twin?



What next?

- True distributed publication
 - Indexation & efficient access are un-solved challenges



What next?

- True distributed publication
 - Indexation & efficient access are un-solved challenges
- Keeping government focus on “the plan” (any plan)



Thanks!

Nicholas Car

nick@kurrawong.ai

KurrawongAI

Brisbane, Australia

<https://kurrawong.ai>

info@kurrawong.ai



KurrawongAI