March 3 2026 Agenda. Tues 11 am AEDT

meeting schedule & presentations, discussion forum, meeting topic ideas what's on agenda

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Junrong Yu (TERN) will present on *Vocabularies in TERN: current state and future directions*.	15 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentation	10 minutes 11.20 am
4	Kheeran Dharmawardena (Cytrax) will present on *A strategic community roadmap for an Australian FAIR vocabulary ecosystem*.	15 minutes 11.30 am
5	Group discussion: Chair ● Responses to the presentation and general discussion	10 minutes 11:45 am
6	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue June 2, 2026	5 minutes 11:55 am

December 9 2025 Agenda. Tues 11 am AEDT

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Peter Sefton (UQ), will present on RO-Crate, a lightweight approach to packaging research data with metadata. Peter will provide an overview and discuss a range of use cases. Including how vocabularies are used within RO-Crates for various purposes.	15 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentation	5-10 minutes 11.20 am
4	Nick Car (KurawongAI) will present on the VocPub SKOS profile and its community-governed future. Nick will also discuss VocSearch, a federated vocabulary system search application.	15 minutes 11.25 am
5	Group discussion: Chair ● Responses to the presentation and general discussion	10 minutes 11:40 am
6	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue March 3, 2026	10 minutes 11:50 am

for a WG to explore not just API but also how vocabs are referenced in other standards (i.e. profiling patterns to declare use of vocabularies). any W3C or OGC members please sign up to this!

June 3 2025 Agenda. Tues 11 am AEST

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Rob Atkinson, OCG, will present on *mechanisms for using controlled vocabularies in the context of 3D Cadastre profiles for ICSM & specific jurisdictions*, as well as *pathways toward federation of vocabularies*.	
3	Group discussion: Chair ● Responses to the presentation and general discussion	
4	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue September 2, 2025	

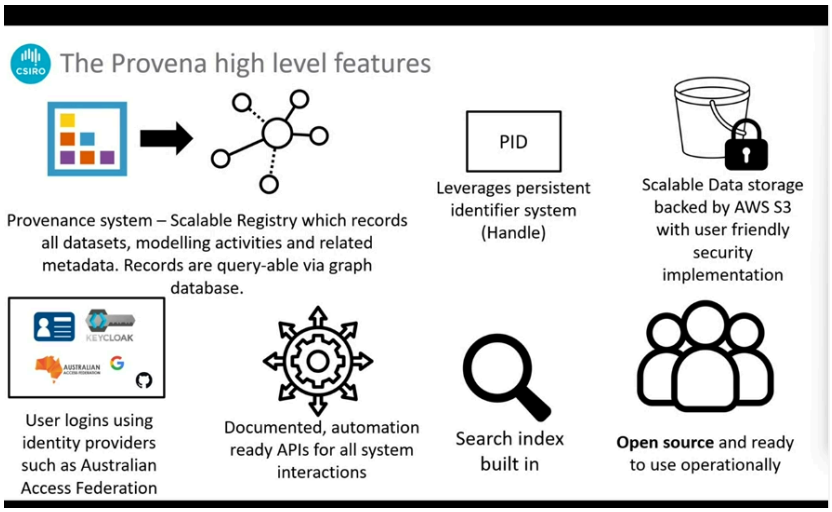
12 participants

March 4 2025 Agenda. Tues 11 am AEDT

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Jonathan Yu, CSIRO - *PROVENA (provena.io), a cloud-based system for registering, managing, and publishing scientific and modelling artefacts with built-in provenance tracking.*	
3	Group discussion: Chair ● Responses to the presentation and general discussion	
4	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue June 3, 2025	

20 participants



Comparison with RO-Crate
RO-Crate - is for metadata of objects as relates to datasets. It relies on external systems to bring in identifiers, and fit into data workflow. Provena is more holistic (within the larger workflow)
Applicability outside of modelling as target audience?
Modelling is just an activity ('model-run'), but could be extended to include sub-types eg 'analysis'.

Trialling integration of Python through workflows

Dec 3 Agenda. Tues 11 am AEDT

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Les Kneebone & Edmond Chuc (KurrawongAI) - *VocPub Editor as a free, user-friendly, browser-based tool designed to streamline the vocabulary management process. (eResearch presentation repeat - abstract here).*	15 minutes 11:05 am
3	Group discussion (Chair) ● Responses to presentation	15 minutes 11.20
4	Group discussion (Chair) ● What was your positive/win? ● What was interesting or a challenge? ● What do you look forward to in 2025? Opportunities? Foci? ● What's on? ● Ideas for presentations or discussion topics - welcome to add to. Here are some topics of interest - See Question set 4 -	20 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue March 4, 2025	5 minutes 11:50 am

Presentation, Les Kneebone (with Edmond Chuc) - Kurrawong AI

Les gave a history of his experiences in thesaurus editors - from Dept of Justice Subject Thesaurus - describing and organising content for a department-wide intranet (Term Tree, Pool Party, content management in MultiTes; Pool Party with content management - Drupal; Pool Party to VocExcel 0.7.0)
VocExcel - Excel template and processor
VocEdit - a browser based system, server based vocabulary management

hhtps://github.com/kurrawong/vocedit

- Server based create and edit
- For lowering the barrier to entry - increase adoption etc
- SKOS, with VocPub quality control with VocPub quality control - elements for better created/managed vocabs
- Save straight from and to your local file system
- Feedback welcome

- Extensive
- Future plans -
    - Hierarchical tree navigation
    - General UX improvements - GitHub - gather feedback and feature requests
    - VocEdit push vocabs update to GitHub
    - Versioning
    - Pushing out to services like RVA and others (for use cases)

Questions, answers -

Can vocabs search across, bring in vocabularies outside of the immediate imported ConceptScheme/vocabulary? - Not at present

Tree view and search coming

SC - Editing with GitHub - NVS, ENDO, DWC - are some other examples change requests, conversations, through GitHub issues

RA -

1. Is there a way to define and group catalogs for Prez?
2. thoughts about profiling an existing vocabulary (choosing a subset)?
    a. vocEdit is more of a general tool, SHACL-UI - if you wanted to create your own profile you could have that shape available through a form, and to manage other resources other than vocabs - eg declare a set of SHACL shapes to manage other resources beyond vocabs
3. github automation and governance processes - are these something that can be regularised?
    a. Eg think about the review cycle and how to express this - eg run validate every pull request. We'd like feedback loop to be available through US
    b. From Margie - GA - We have a two step "governance" of vocabs using GitHub - the vocab custodian looks after the content and the publish team validates on the commit and then once the custodian says it is ok, it is published.

SC - Why wasn't poolparty or vocbench good enough? What was the motivation for creating a new vocabulary editor??

Browser based, away from server administration

First we see files managed in GitHub

Server based tools require account manager

If you have GitHub account - you can use

Thoughts about capturing Provenance?

Plans to do that - in other products we deploy RDF Delta and Prov patterns - to capture changes made to dataset,who made changes and when - no reason why we could not integrate

Group discussion - general 2024 wrap up and into 2025

Highlights and what we are looking forward to in 2024 -

RAtkinson - 2024 achievement - the 3D Cadastre schema - best practices to modelling using standard schemas, standard models - unit testing level

MSmith - I'm looking forward to trialling images with a specific vocab at GA: https://demo.dev.kurrawong.ai/catalogs/exm:demo-vocabs/collections/demovocabs:image-test/items/imagetest:apron-image

RAtkinson - Planning to run a workshop in EU re ObservableProperties and cross-standard alignments of implementations - more modern. Raised the need to feature-type catalogue work across community (SC - there is INSPIRE - has limitations)

September 17 2024 Agenda. AEST

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 12 noon
2	Matthias Liffers (ARDC) and the AGLDWG co-chairs (Nick Car, John Machin, Terhi Nurmikko-Fuller, Andew Fitzgerald). *Introducing the AGLDWG / ARDC linked data identifier service*	10 minutes 12:05 pm
3	Chee Ng (DEMIRS). *Experience using the AGLDWG / ARDC linked data identifier service*.	10 minutes 12:15 pm
4	Simon Cox (Timely Logic) *Experiences of a user of a range of identifier services*	10 minutes 12:25 pm
5	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 12:35 pm
6	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, December 3, 2024	10 minutes 12:50 pm

September 17 2024 Meeting Notes

21 attendees

Extracts from the video-conference chat

- ABN has non taxation compliance uses, it can also help businesses interact because it has something of the feel of a credential (and some fairly basic verification?) about it.
- Unfortunately not every organisation operating in Australia is eligible for an ABN. One of our newest DataCite Consortium members is an international treaty organisation that is headquartered in Tasmania. They are not eligible for an ABN and had to provide us a letter from (I think) DFAT to make our lawyers happy.
- LOCI was hosted by GA, then turned off

June 4 2024 Agenda. AEST

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Jo Croucher (NCI) - *Accessing w3id to identify vocabularies. Perspectives of a new user*	15 minutes 11:05 am
3	Valerie Dixon, (NASA) - *GCMD. The Global Change Master Directory Keywords. Past, Present & Future*	15 minutes 11.20
4	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, 2024	10 minutes 11:50 am

June 4 2024 Meeting Notes

Jo Croucher - Accessing w3id to identify vocabularies
Aus Scope, TERN, ARDC, NCI
MT Metadata documentation - paper, GitHub, zeondo, how to documentation
Metadata - wanted to use controlled vocabs, version controlled etc - a profile of the MT metadata, machine readable, PIDs
Jeff Rothenberg - "Digital information lasts forever - or for 5 years, whichever comes first"
W3id with Persistent links back to RVA
In vocabs aust - each concept has a w3ID URI - clicking redirects you to concept landing page
Steps - w3id -

- Check domain is available
- Fork repo
- Open github desktop - create readme and MT access file
- (redirect to RVA- could use this file as a template - Apache redirects)
- Merge back to master via a pull request
- Export - update with w3id and re-import. Test as new vocab before using the over right function

Choices

- What top level domain name you are using
- Planning for subsequent URL format

- URI generation pattern (sg use incremental numbers fo each concept)

Summary -

- W3id can be used to provide permanent identifiers for RVA vocabularies
- Setting up and maintaining w3id domains is not overly complex
- Carefully consider vocabulary maintenance & sustainability
- Thank you Rowan Browlee and RVA/ARDC for support

Valerie Dixon

Earth science DAta and Information System (ESDIS) Project

GSCD Keywords - hierarchical controlled vocabulary
Ever evolving to suit needs of NASA's EOSDIS community
GCMD Keyword Viewer
KMS - KEyword Management System - SKOS objects and static directory
GCMD Keyword Forum - a place to come to talk about keywords and talk about, discuss possible changes etc, updates re new releases etc

GCMD Keyword viewer -
Gets down to entities with descriptions
Easy to navigate (by humans)

Restful API -
Navigation to URL itself

Keyword viewer -
Download - as JSON, RDF, XML, CSV

KNM Use guide - best practices

GCMD - Keyword Forum
Come and discuss keywords
Can action every other week if changes are not impactful

Keyword use in EOSDIS
A consistent vocabulary is crucial for enabling programmatic search and discovery ad compatibility across tools and services

- ESDIS Standards Communication Office (ESCO) community review

GCMD Keywords has also contributed to Machine Learning (ML) tooling

- GCMD - Keyword Recommender (GKR)
- Earthdate search - uses science keywords - drill down, filter down hierarchy

- GCMD Keyword recommender (GKR) - earth data Metadata Management Tool (testing machine learning to make curators life easier)

Use for Inter-agency collaboration

Tomorrow …..
HIerarchical keywords - Semantic Ontology
How? When? With what resources?
ESIP - semantic web for EArth and Environment Terminology (SWEET).....? Future of/work on?

Knowledge Graph
- We use Graph QL and Graph DB for metadata association in CMD
- Currently lack the expertise and priority to look into how to use Graph for GCMD Keywords and or associations, but there is a path started
- Low hanging Fruit -
- Create unique IDs for each keyword
  - and cite those in metadata so keyword updates don't invalidate records

Links
https://2024januaryesipmeeting.sched.com/ - The meeting has a virtual option and there are 8 Australian Partners in Australia so you can get a discount rate.
Earth data forum
https://forum.earthdata.nasa.gov/viewforum.php?f=7&&ServicesUsage=101
Keyword searcher
https://gcmd.earthdata.nasa.gov/KeywordViewer/

Questions-
Rob A -NASA sponsors https://www.ogc.org/initiatives/open-science/...
OGC persistent demonstrator - how could bindings of GDMC to for eg GeoDCAT - be used as a demonstrator of best practice
Lesley - 2030 Geophys project - With the 2030 Geophysics project that Jo talked about, NCI were able to easily add a couple of keywords to GCMD - e.g. MTH5 (its a data format). Thanks GCMD and NASA

March 5 2024 Agenda. AEDT

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Nick Car (KurrawongAI) - *Rescuing orphaned vocabularies*	15 minutes 11:05 am
3	Group discussion (Chair) ● Responses to Nick's presentation	15 minutes 11.20
4	Group discussion (Chair) ● Opportunities and challenges for 2024 Ran short on time - Add a key challenge or opportunity for 2024 here if you like - Megan Wong - a key challenge for soils community/authorities/funders (DAFF) is determining the governance/authoritative structure through which vocabs will be maintained and the PID domain. ● Ideas for presentations or discussion topics - welcome to add to. What would I like training on - See Question set 4 - from 2023 Vocabulary Symposium is a resource we may refer to ● What's on - feel free to add to, ALDWG will also cross-post/provide link	15 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, June 4, 2024	10 minutes 11:50 am

March 5 2024 Meeting Notes

Discussion of Nick's presentation

- Question about issues to be aware of when transferring ownership of a vocabulary. In the case of an initiative is to be wound-down. Good idea to have sufficient documentation within the vocabulary file, to enable it to be operated and understood at handover. For example, what is the scope of vocabulary holdings, are there relationships between vocabularies, is there an overarching logic to the set? Are these standalone vocabularies or are they used within a particular model? Where is the model? Are there derivatives of the vocabulary? Does the new owner know how to manage the identifiers (e.g. if moving from one server to another)?
- Question about what may have prompted NEII to create vocabularies rather than use existing examples. May not have known about existing examples. There may not be

existing examples. Often we need a national vocab that has some sort of context that must be preserved (even if the local is related to the international, or the local replicates or subsets the international). Wherever existing terms are reused, always a good idea to include attribution - indicate the source of the term, which assists the expression of relationships across local and international vocabs.

- Discussion about challenges involved in seeking to convince government agencies to publish their vocabularies in semantic format. Importance of the community in arguing for publication. Also important to have examples in government that demonstrate linked data as routine, to encourage take-up by other agencies ⇒ increased imperative to publish semantic expressions of vocabularies.
- Reminder to send updates to the mailing list. What are your opportunities and challenges for 2024? Ideas for presentations or discussion topics? Update the group on projects that you'll be working on this year. Let us know about events or activities of interest to the group

December 5 2023 Agenda. AEDT

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Presentation ● Steve McEachern (Australian Data Archive, ANU) : *A Roadmap for the Future of Vocabularies in Australia.*	15 minutes 11:05 am
3	Presentation ● Valerie Dixon (ESDIS): *GCMD: History, tooling, application, and future directions*	15 minutes 11.20
4	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, March 5, 2024	10 minutes 11:50 am

December 5 2023 Meeting Notes

September 5 2023 Agenda. AEST

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Presentation ● Margie Smith, (GA): *Developing a vocabulary publishing workflow at Geoscience Australia.*	15 minutes 11:05 am
3	Presentation ● Donald Hobern, (APPF)*: Relating vocabularies across data providers for the EcoAssets project*	15 minutes 11.20
4	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, Dec 5, 2023	15 minutes 11:50 am

September 5 2023 Meeting Notes

- Discussion of the OGC spreadsheet importer to convert spreadsheets to encoded vocabularies. OGC is working with ISO to formalise some of the vocabulary registration processes, workflows.
- Additional notes provided by Rob Atkinson on the OGC spreadsheet importer
    - *Here is an example configuration. We're setting up a more pedagogical example soon. https://github.com/ogcincubator/iso19157-3-sample*
    - *Another example I set up with a simpler config is*
    - *https://github.com/rob-metalinkage/uplift-ppi*
    - *Look in the github action of the second one for a little more commenting*
- Discussion of vocabulary reuse, profiling, federation, subsetting. Lots of initiatives. People do things in different ways. Can there be one or more common approaches developed, which then may be supported via tool development?
- Mention of https://www.checklistbank.org/ - developed by the Catalogue of Life and GBIF to marshall species lists, using standard data formats and tools that aim to make it easier to identify differences.
- Mention of a BoF being organised by Guru, at eResAust 23.

June 6 2023 Agenda. AEST

meeting schedule & links to presentations, discussion forum, meeting topic ideas what's on

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Presentation ● Gabriele.Bammer (National Centre for Epidemiology and Population Health, ANU): Building a vocabulary for research integration and implementation useful for transdisciplinary and related research	15 minutes 11:05 am
3	Presentation ● Erin Kenna (CSIRO): eReef's NetCDF-LD implementation of the Parameter Usage Vocabulary (PUV) ontology through collaboration with BODC to use and extend vocabularies from the NERC Vocabulary server (NVS).	15 minutes 11.20
4	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 11:35 am
5	Brief updates, what's on & wrap-up: Chair ● Next meeting ○ Tue, 11 am, Sept 5, 2023	15 minutes 11:50 am

June 6 2023 Meeting Notes

Discussion of the question of whether Gabriele may choose to implement a hierarchy, or instead maintain separation between categories and tags. Gabriele is developing a vocabulary while also developing a discipline. Value of analysing domain literature to see level of matching between terms and occurrences in the literature.

Erin discussed an approach to relating a discipline layer to a vocabulary term.

Some discussion of the background to BODC. Running for 25+ years. A service that hosts large collections of parameters. Supports a PAV view of vocabularies and concepts, enabling a view of a concepts entire version history. Additional sophisticated tooling in planning for release at BODC. Nick Car may be able to provide an update at a future AVSIG meeting.

Mention of initiatives to standardise vocabularies for geochemistry. And a proposal for ICSM to publish their vocabularies related to their foundational spatial data. There is a strong requirement in the new wave of address models, to use vocabularies.

March 7 2023 Agenda. AEDT

meeting schedule & links to presentations, discussion forum, proposed meeting topics

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Group discussion: Chair ● What vocabulary-related activities are people undertaking in 2023? What are you planning and what issues are you facing? ● Public vocabulary patterns library (Nick Car) ● Report from OGC meeting (Rob Atkinson) (see screenshot below)	
3	Group discussion: Chair ● Are there events occurring in 2023 that you'd like others to know about?	
4	Brief updates & wrap-up: Chair ● Next meeting ○ Tue, 11 am, June 6, 2023	10 minutes 11:50 am

Agenda Item - Rob Atkinson, OGC

I'm in transit on Tuesday, however a few things to report from the OGC meeting that relate to best practices in this space

1. The Spatial Data on the Web Working Group held a discussion with the upshot that "observable properties" need to be handled in a way that supports both detailed models and simple vocabulary perspectives.

2. GeoDCAT profile of DCAT to be formalised in a new WG... profiles are how vocabularies are bound to structures, so a formalism for this is required and will be generally relevant.

3. We have initiated a vote to publish an Agriculture Domain model as a modular ontology (based on SOSA, PROV, RDF Datacube and other standard vocabularies

4. A range of profiling activities at the intersection of domain models and controlled vocabularies - such as buildingSmart IFC and CityGML commonalities.

In addition Nick Car and I will be collaborating on profiling a similar model, which largely involves mapping common concepts to a range of different vocabularies with jurisdictional versions.

The common thread here is of course the need for a canonical mechanism to bind typical schemas, such as metadata records and observations to controlled vocabularies to define profiles for a specific application domain.

Agenda Item – Nick Car, Kurrawong

- Public vocabulary patterns library
  - https://vocpatterns.kurrawong.net
  - KurrawongAI are hosting a free, open source, library of vocabulary patterns learned or discovered from many years of vocabulary implementation
  - Patterns are in response to vocabulary-related questions., for example, *How to I indicate the source of my vocabulary?*
  - Questions are answered with one or more Pattern with the above having the Vocab Derivation & Provenance patterns
  - It's early days for the library but we have 2 current projects that are contributing to the patterns now, so we hope to have perhaps 15 questions and perhaps 15 major patterns by July
  - We are happy for pattern contributions and for people to raise questions too - see the homepage for contribution / request details

March 7 2023 Meeting Notes

eRes - date is up. https://conference.eresearch.edu.au
Katherine - TADWIG - NAtional conference is in Hobart

AVSIG Meeting Agenda

https://www.tdwg.org/

RDA - later this month - provides recoding
SciDataCon

If we could put the list into the google groups space it might be possible for anyone to add events, rather than relying on a single admin
https://calendar.google.com/calendar/u/0/embed?src=sb9cnmaknqpsifs7ttn9ksobu0@group.calendar.google.com

https://www.tern.org.au/integrated-earth/

Les Kneebone - Biosecurity - will be interested in Ag Domain model developments. Biosecurity Thesaurus
Mieke Strong - is interested in learning more about the profiling
Megan - Working on Observable Properties of soil physical methods

https://documentation.ardc.edu.au/display/DOC/Subscribing+to+vocabularies

Natalia Atkins - Maintaining vocabs under IMOS, keeping eye on developments in vocab space. There is an architecture update of system planned - should not result in to many disturbances - business as usual

Les Kneebone - Centre of Excellence for Biosecutiry and risk analysis - digital library - controlled vocabularies required. Two that have been released in research vocabularies Australia. Priority pest lists - compiling -

Mieke Strong- BDR - trying to bring all kinds of vocabs together. What we run into - knowing about the vcoabs that are out there, the providers knowing we would like them to use the vocabs, in the system - what are the linked terms in the BDR content

Erin Kenna - (Environment, Dutton Park) BODC Parameter Usage vocabularies PUV and its semantic model - https://github.com/nvs-vocabs/S06 , continuing to work on. Including what quality codes to use in data - started making a list in

Katherine Tattersall - NCMI - map to IMOS standards - letting you know there is metadata work going on in marine data exchange (did not grab the name)

Gabi Ceregra - working with NFF on Farm Data Code. Where do we source agricultural vocabularies, and encourage their re-use?

December 6 2022 Agenda. AEDT

meeting schedule & links to presentations, discussion forum, proposed meeting topics

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	● Report back on the recently held Vocabulary Symposium and Workshop ● Speakers will include ○ Rowan Brownlee (ARDC): *Symposium background & purpose, attendees, keynotes, presentations & outputs*. ○ Megan Wong (Federation University): *Workshop background & purpose, attendees, activities, outputs and next steps.*	30 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentations and general discussion	15 minutes 11:35 am
4	Brief updates & wrap-up: Chair ● Next meeting ○ Tue, 11 am, March 7, 2023	10 minutes 11:50 am

December 6 2022 Meeting Notes

The group expressed interest in

- understanding how we may progress development of the proposed roadmap, so as not to lose momentum
- advocating and communicating across domains, the aims of the roadmap, and the value of participating in this activity.

The organising committee responsible for the Symposium and Workshop are planning to

- contact roadmap document section coordinators to discuss next steps. Section coordinators should expect contact within the next week.

There was discussion about the need for more coordination at the basic technical level.

- The I in FAIR (Interoperability) requires technical coordination. For example, we need to reach agreement on the mechanisms for exchanging data. When exchanging data, we may agree on how that data will be structured and encoded. But we haven't reached

agreement on how the technology will operate to enable that data exchange, and we haven't reached agreement on the governance frameworks for that technology. And we need to ensure that an AU vocabulary roadmap connects internationally.
- Different domains use different approaches to managing and enabling access to vocabularies. We need to advance technologies and implementations toward convergence of features. Hard to see 5-10 years in the future, so start with a year. For example in the health & medical domain, an approach that seeks to standardise the use of OntoServer as a platform for terminology access. We need to have those international dialogues and longer term thinking, while seeing what we may achieve in the short to medium term.

The next OGC member meeting is at Frascati, southeast of Rome in February 2023
- to continue work to revamp the OGC reference architecture and to explicitly call out the role of vocabularies. OGC would welcome ideas, comments and responses from AVSIG.
- Interoperability architectures have been component-centric rather than information-centric. There are lots of technical ways of moving components, but the thing that is constant is the exchange of information and the semantics of that. And yet those things are not easily done. So we are going to be moving towards looking at a design pattern catalogue for implementation architectures. And in particular trying to work out how we would call out the role of vocabularies in those architectural design patterns.

September 6 2022 Agenda

meeting schedule & links to presentations, discussion forum, proposed meeting topics

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Presentation ● Andrew Hancock (Stats NZ). *Modernising management of statistical classifications at Stats NZ*	30 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentation and general discussion	15 minutes 11:35 am
4	Brief updates & wrap-up: Chair ● Next meeting ○ Tue, 11 am, December 6, 2022	10 minutes 11:50 am

September 6 2022 Meeting Notes

Questions to Andrew

Although we need to move away from a code-based approach to organising classifications, to what degree should we seek to maintain legacy systems that depend on these codes? By maintaining these legacy systems, are we prolonging a problem?

- Back in the 1980s when we were using mainframes and there were considerable memory constraints, we had to use codes. There was no other option. And so we used codes to arrange classifications. But this means that there are only a certain number of numeric slots available at each level. This is quite a constraint. Now we can do more with words and labels, and computer storage is not the issue that it was. Until the legacy systems can be switched off, we can maintain the codes, but over time, transition away from them

Are you exploring the use of XKOS?

- We are planning to explore XKOS in future. At the moment this is beyond our immediate requirements. Other agencies such as Stats Canada & the French National Statistical Agency may have progressed further with XKOS

Where are you at with the publication of NZ classifications as RDF?

- Although this is in the pipeline, there are other priorities at the moment, such as the 2023 Census. Then we will return to system review and development.

What is the status of the Aria software as used by Stats NZ. Is it commercial? Is it accessible for use by non-commercial users?

- The long term goal of the vendor is to make the software available as open source. And then if users wish to be assisted with customisation, services may be contracted with the vendor. Stats NZ has a perpetual licence that allows them to make the system available to Pacific Island countries free of charge.

Can you comment on the use of Aria by Maor and Iwii?

- Aria is multilingual. As we approach revising the NZ Maori and Iwi classifications, Stats NZ will have discussions with representative groups including the Iwi Leaders Group and the Iwi Data Forum. Aria has been designed from a Maori perspective, incorporating familial relationships and contexts.

Additional questions to Andrew

The meeting didn't have time to cover the following two questions. Andrew provided answers via email. The questions and his answers are copied below.

Automation and metadata modelling are of interest to me. How will research/researchers be impacted?

- At the moment a researcher can get counts from a survey about a particular category or variable, and then they have to do a lot of work to expand the narrative around the data by cross-tabulation with other variables, all of which is very much based on stand-alone classification systems. We have started some work on mandating standards for key linking variables such as name, address, sex/gender to support the Integrated Data Infrastructure (IDI) which is our main tool for researchers. See https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/ But all that research can only be done in a secure datalab, and obviously has to be highly confidentialised, and the resulting research has to be quite generalised.

- Moving forward to a concept-based approach will negate a lot of the problems of trying to link classifications – which is a very manual process and will also allow us to basically use a classification as a portal to more information. An example being occupation where you have a code, a label, a one sentence task definition, and a statement about the formal qualification needed – it doesn't tell you much. So, using apis we can link systems

together either through a classification or concept or an entity. That allows us to bring in competencies, roles, income ranges, employers etc as information that a classification can't provide. It then allows us to be more dynamic in keeping content relevant because the classifications are currently reviewed in a cyclical nature and are always behind the contemporary reality, so researchers aren't able to do real-time analysis. Further we can introduce AI and machine-learning to update and maintain content, whether it is what we currently see as classification categories or whether it is things like the words used in coding indexes. We can also then take the research on a dataset and bring that back into the system as a means of updating and informing. Hence the creation of reusable views of content, and then we use automation or a stock-market analytics to articulate the most common views or 'standard/most popular' giving researchers more options.

Does the new Aria system extend to alternative or traditional Māori knowledge views or contexts?

- Yes. The vision, philosophy and tool itself has been thought of from a te ao Māori world view rather than the Eurocentric model that we have traditionally applied. It is about the ability to introduce contextual fluidity into the views/classifications to tell better stories and integrate content away from static, mutually exclusive boxes. For example, a person may affiliate with an Iwi for Treaty settlement purposes, or with a hapū for fishing rights, or a marae for social gathering or community, or a whanau for familial support or well-being, and those needs, and contexts can happen continuously /concurrently in a day. From a traditional classification perspective, we have to choose one of those ie you can only belong to an iwi, or a hapū – you can't be both at the same time and if you are we assign one. The same applies with geographies where we seek to put geographic context round iwi and hapū for analytical and statistical purposes but use local government boundaries as opposed to the mountains, rivers, foreshores etc that Māori use when telling their narrative.
- The second part is about bilingual content – we traditionally work in English and then translate to Māori and Ariā enables both languages. However, the underlying model means we can actually store content as developed and described in te reo and then translate back to English – giving two perspectives. The fluidity and neural networking approaches utilised in how we integrate the underlying model and processes is heavily influenced by tikanga Māori and mātauranga Māori as the content becomes taonga.
- The naming for the vision, philosophy and tool itself also was important. We could have had the Classification Management System or Classifications+ or something like that but we proposed using a te reo term and Ariā was what the vendor chose (on our suggestion) as the word represents - likeness, resemblance, notion, idea, concept, theory, feeling, theme – as opposed to just translating the word classification.

June 7 2022 Agenda

meeting schedule, discussion forum, proposed meeting topics

Item No	Topic	Time (Start)
1	Introduction - Chair, Megan Wong (Federation University)	5 minutes 11 am
2	Presentation ● Nick Car, (Kurrawong AI). *Recent progress within the ICSM and GGIC communities to govern and technically manage vocabularies*	15 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentation and general discussion	30 minutes 11:20 am
4	Brief updates & wrap-up: Chair ● Next meeting ○ Tue, 11 am, September 6, 2022	10 minutes 11:50 am

June 7 2022 Meeting Notes

Update from John Timermanis on where things are at with the decommissioning of NEII. More information available here.

Questions to Nick

Are you encouraged that the initiative involving GGIC and ICSM is likely to succeed?

- They are taking a gradual approach to testing out the interest of GGIC and ICSM. This is a reasonably straightforward proposal to publish some vocabularies and run some checks & management procedures around them. Not a complex activity.

Interested in the approach involving use of version control tools as applied to vocabularies

- If we have tools that work in the management of textual information assets, we may as well reuse them.

Comment on the ICSM part of Nick's presentation. Every published vocabulary will be

registered with the ARDC vocabulary service, to be accessible to the wider community. The published vocabularies will belong to the Digital Atlas of Australia.

Is the tooling that checks the validity of vocabularies easy enough for non-technical people to use?

- To keep the management demands low, the tooling expects a particular mode of vocabulary expression. The tooling validates against a data file (a profile), that expresses the rules that must be followed. There are different ways to provide the required validation file. The challenges are not so much how to generate a validation file, but to conceptualise how you want a vocabulary to be structured. The profiles can be published, and provide starting points for creation of other profiles. For example, the vocpub profile is published and can be commented upon

Good to see activity around the development of shared profiles. How does a registry work out which metadata to implement. Decision-making on this is assisted if registry owners can see that particular profiles are being reused, advocated, within a domain community.

- Very important to understand whether a concept is being defined within this vocabulary, or is being reused from an existing vocabulary. Important to include a link to the source of a concept. Helps in understanding where a concept is used across semantic resources.

March 2022 Agenda

meeting schedule, discussion forum, proposed meeting topics

Item No	Topic	Time (Start)
1	Introduction - Chair, Natalia Atkins (IMOS)	5 minutes 11 am
2	Presentations ● Kheeran Dharmawardena. *Introducing the Dagstuhl Downunder vocabulary workshop* ● Simon Cox. *Vocabulary governance* ● Nick Car. *A catalogue of vocabulary technology tools*	30 minutes 11:05 am
3	Group discussion: Chair ● Responses to the presentations and general discussion	20 minutes 11:35 am
4	Brief updates & wrap-up: Chair ● Next meeting ○ Tue, 11 am, June 7, 2022	5 minutes 11:55 am

March 2022 Meeting Notes

Question to Nick.

RDA VSSIG had expressed interest in taking on the catalogue, to be maintained and developed within one of its task groups. Any updates?

- Nick is unaware of further steps taken by RDA VSSIG. Rowan contacts the RDA VSSG chairs on March 1. No update available at the time of drafting these notes (March 9)

Question to Simon.

Will there be another vocabulary workshop in 2022.

- Yes, currently in development. Information to be circulated in future. To be developed in partnership between ARDC, Australian Data Archive and CODATA. Will be seeking suggestions from the community. If you have ideas, circulate to the AVSIG group, or email rowan.brownlee@ardc.edu.au

Interest expressed in approaches to building a vocabulary from scratch. There may be discussion starting on the AVSIG mailing list.

Comment from Shawn Ross that work is gradually progressing on the Munsell Colour Chart as a vocabulary. There had been discussion at the September 2021 Vocabulary Workshop.

Comment from Nick that his company is rolling out more vocabulary systems. Most recently, one for the Geological Survey of South Australia.