

Controlled Vocabularies and Data Integration

Presentation to ICSM Metadata Working Group Q1 2020

Jenny Mahuika, Data Librarian
TERN Data Services and Analytics
j.mahuika@uq.edu.au





TERN

Introduction to the Terrestrial Ecosystem Research Network

The challenge of harmonising diverse data

Process and Examples

Current Status





TERN

Introduction to the Terrestrial Ecosystem Research Network





Australia's Land Ecosystem Observatory



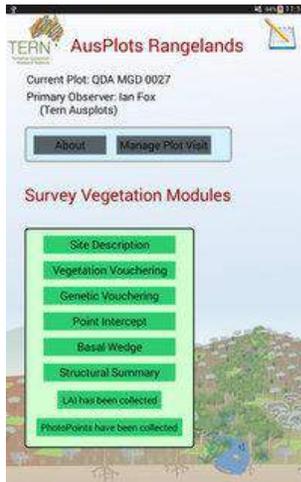
TERN's ecosystem observing sites

- ▲ ecosystem surveillance monitoring plots
- ecosystem processes monitoring SuperSites and eddy covariance towers
- landscape scale monitoring calibration and validation sites



TERN Purpose¹

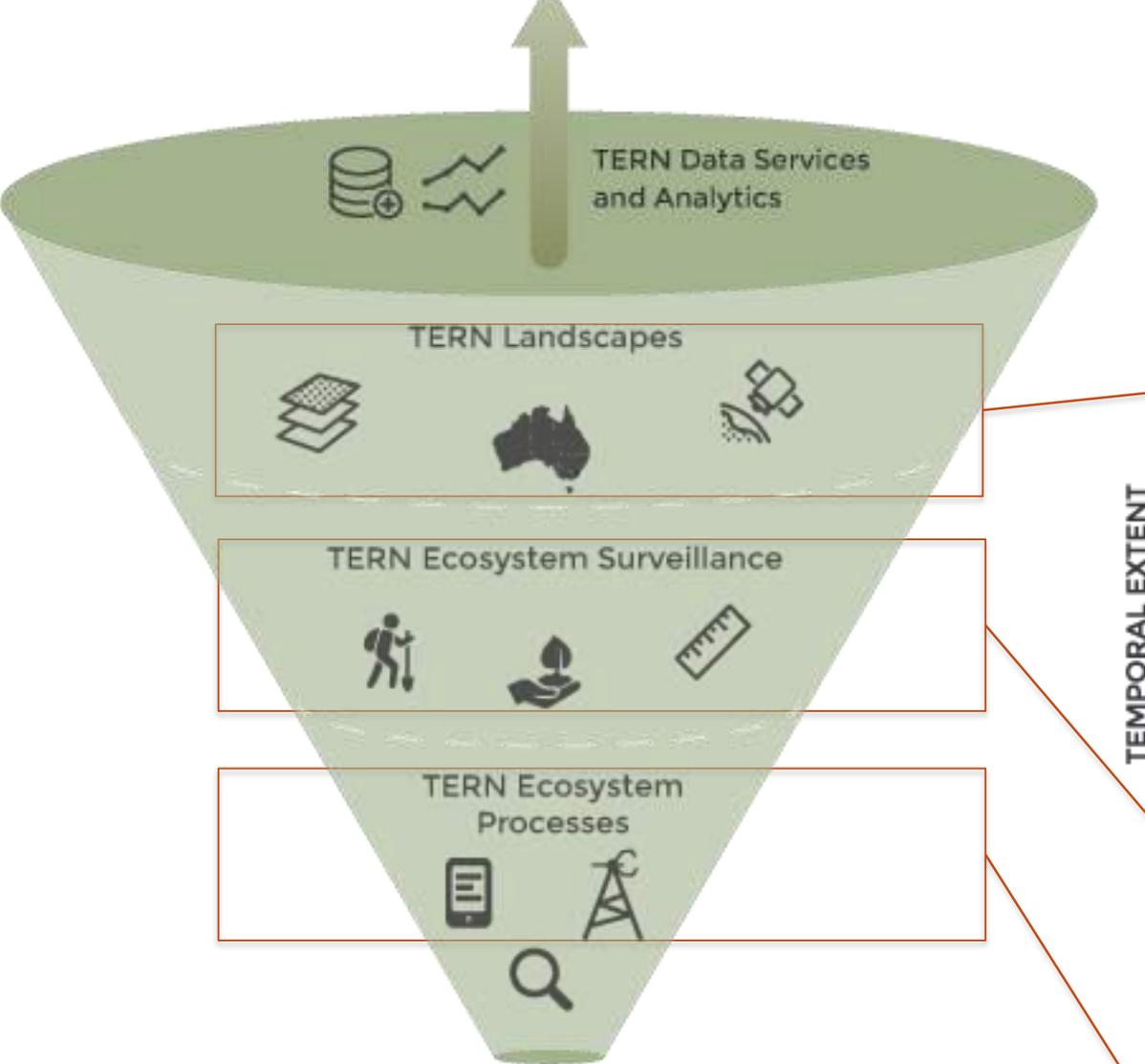
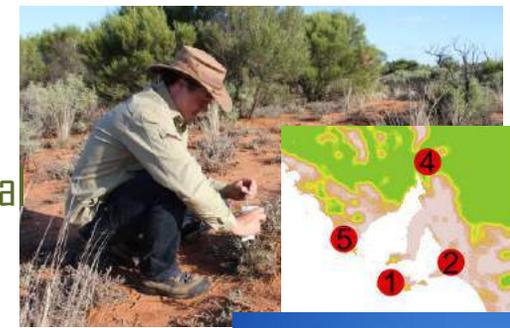
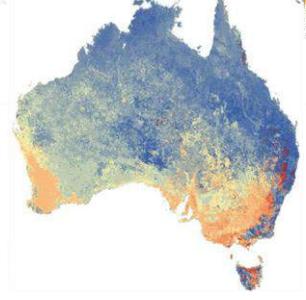
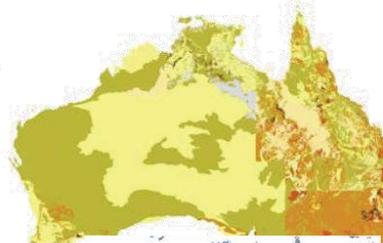
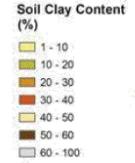
National infrastructure for collecting, collating, storing and sharing Australia's terrestrial ecosystem data sets and knowledge.



¹TERN is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy from 2009



TERN in Operation



- Satellite remote sensing products
- Land cover dynamics and phenology
- Vegetation composition and structure
- Fire dynamics and impacts
- Continental Soil & Landscape data
- Plot-based surveillance monitoring
- Soil sample, leaf tissue samples, LAI, Basal area
- Carbon, energy, water fluxes
- Phenocams
- Acoustic sensors
- Flora population

TERN



TERN

The challenge of harmonising diverse data



Ecosystem science data

- **Messy**

- Combination of human and sensor observation at different spatial and temporal extents

- **Diverse**

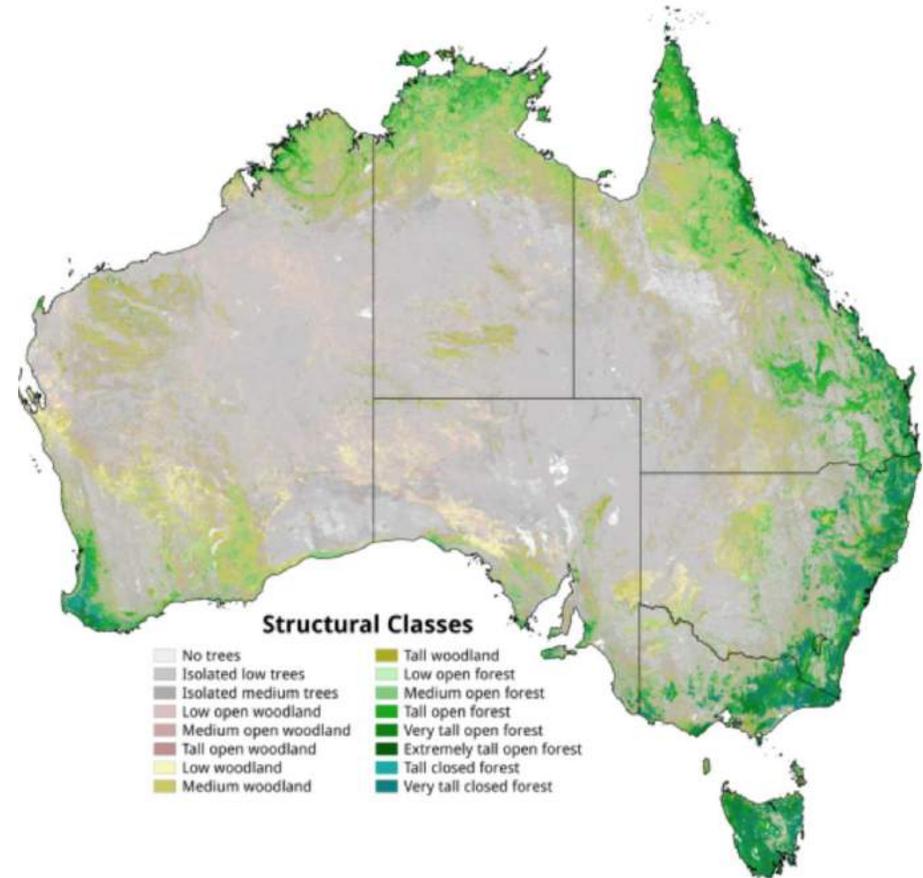
- As above but also different types and formats

- Point, Grid, time-series, one-off, wide geographical extent

Structural Growth Form

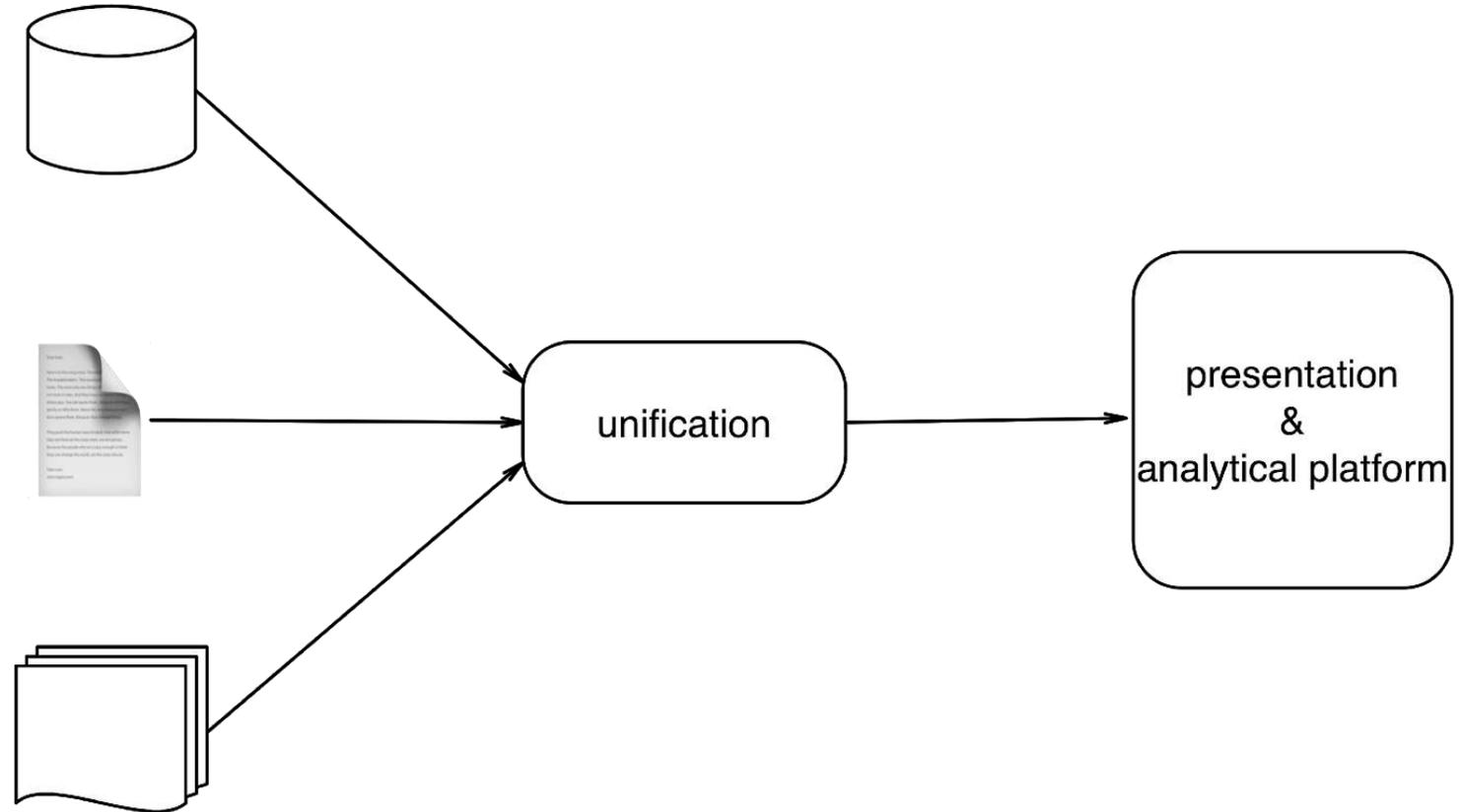
Table 28 Structural formation classes

Proj. foliage cover	>70%	>30–70%	10–30%	<10%
Crown class	Dense/closed	Mid-dense	Sparse	Very sparse
Crown cover % ¹	>80%	>50–80%	20–50%	<20%
GROWTH FORM²	Structural formation classes (qualified by height)			
Trees >30 m	tall closed forest TCF	tall open forest TOF	tall woodland TW	tall open woodland TOW
Trees 10–30 m	closed forest CF	open forest OF	woodland W	open woodland OW
Trees 2–10 m	low closed forest LCF	low open forest LOF	low woodland LW	low open woodland LOW
Shrubs 2– 8 m	closed scrub CSC	open scrub OSC	tall shrubland TS	tall open shrubland TOS
Shrubs 1–2 m	closed heath CHT or closed shrubland CS	open heath OHT or shrubland S	shrubland S	open shrubland OS
Shrubs <1 m	dwarf closed shrubland DCS	dwarf open heath DOHT	dwarf shrubland DS	dwarf open shrubland DOS
Succulent shrub	NA	succulent shrubland	succulent shrubland SS	open succulent shrubland OSS
Hummock grasses	NA	NA	hummock grassland HG	open hummock grassland
Tussock grasses	closed tussock grassland CTG	tussock grassland TG	open tussock grassland OTG	sparse tussock grassland STG
Herbs ³	closed herbland CH	herbland H	open herbland OH	sparse herbland SH
Forbs	closed forbland CFB	forbland FB	open forbland OFB	sparse forbland SFB
Rush	closed rushland CR	rushland R	open rushland OR	sparse rushland SR
Vines	closed vineland CVI	vineland VI	open vineland OVI	sparse vineland SVI
Ferns	closed fernland CFN	fernland FN	open fernland OFN	sparse fernland SFN
Sedges	closed sedgeland CV	sedgeland V	open sedgeland OV	sparse sedgeland SV



- Objective

combine data from different sources into usable and trusted information



Controlled vocabularies provide an opportunity to harmonise at different scales and across different domains

Harmonisation

General

>

Specific

GCMD Science Keywords

ANZ Fields of Research

– Platforms, Instruments

– Observed properties

– etc

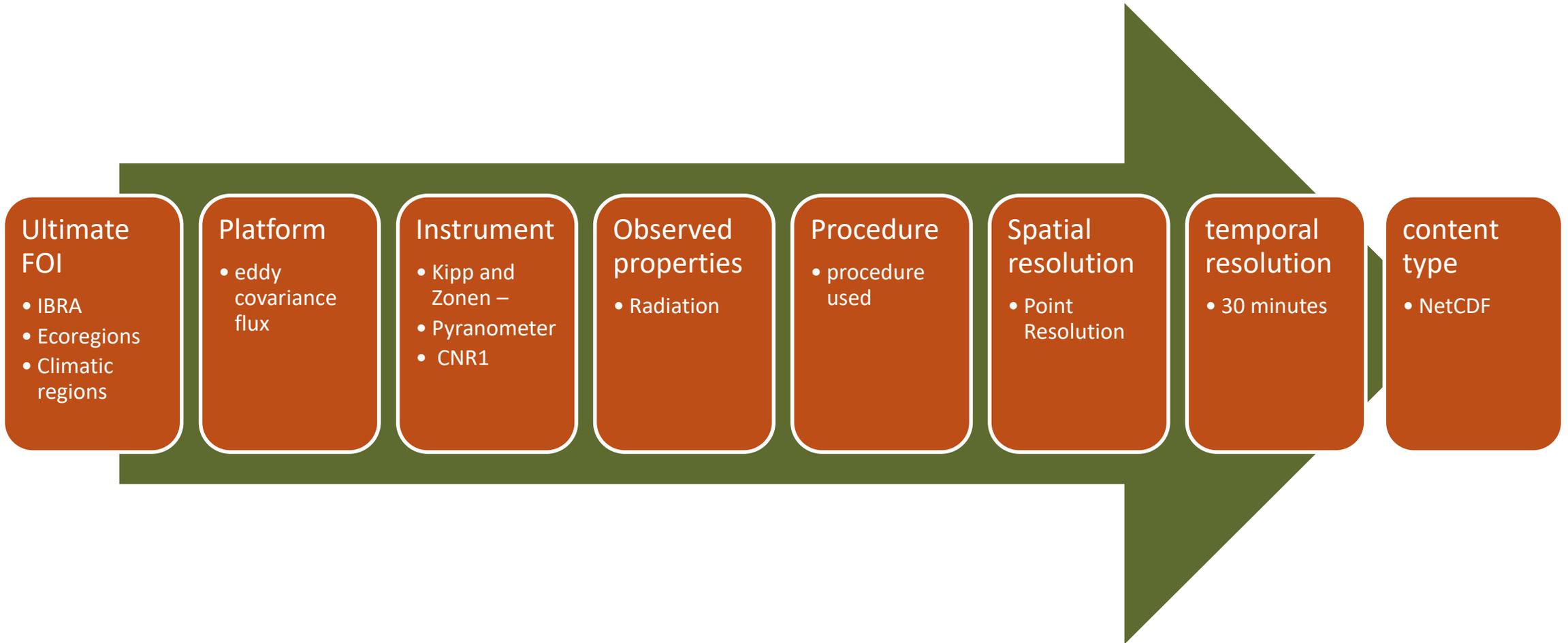
Vocabularies are key

- Platforms, Instruments - TERN vocabularies, based on SOSA ontology, aligned with GCMD
- Spatial regions – Australia's Bioregions (IBRA), Ecoregions, States and Territories
- Spatial resolution, Temporal Resolution, Content type - GCMD terms
- UoM – QUDT ontology
- Observed properties – TERN vocabulary, RDF, aligned with EnvThes
- Methods/procedures – TERN vocabulary, RDF
- Organisations, Projects, People – TERN vocabularies, based on schema.org

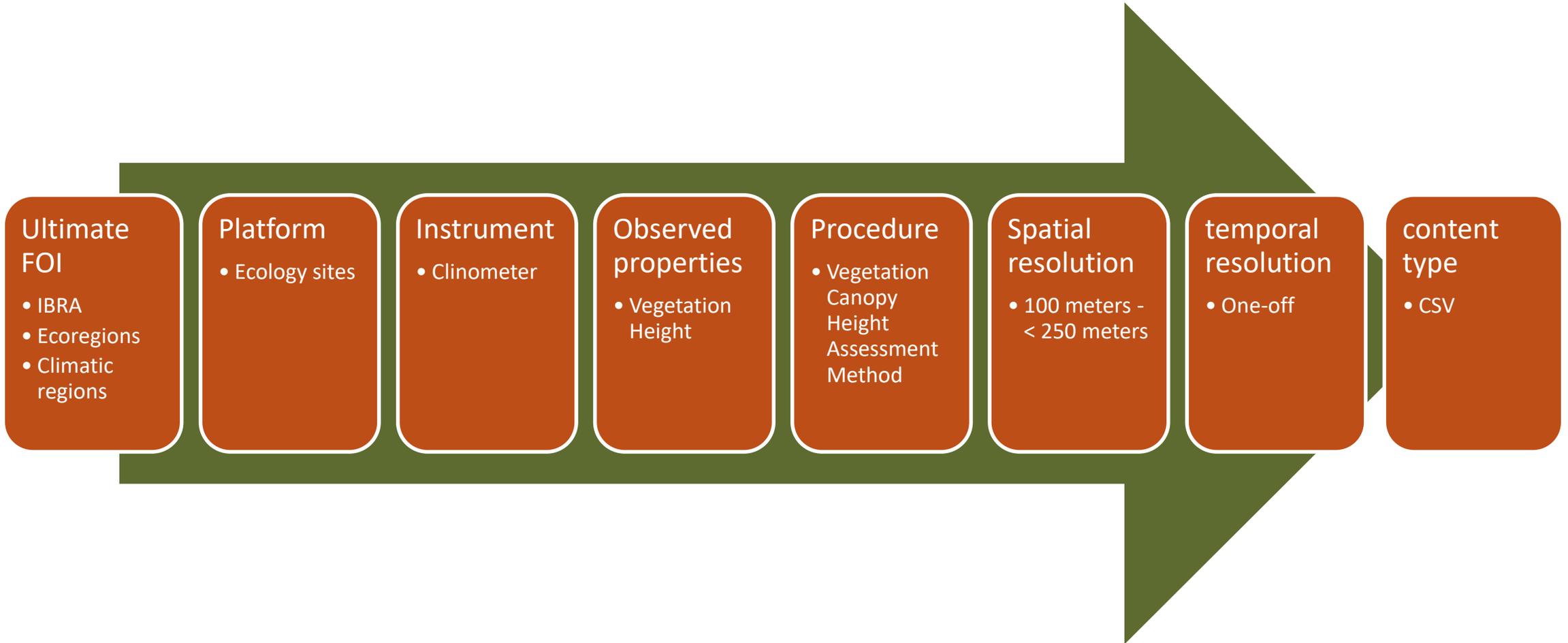
GCMD <https://gcmdservices.gsfc.nasa.gov/static/kms/> many also available through ANDS
EnvThes <http://vocabs.lter-europe.net/edg/tbl/EnvThes.editor>



Data from Flux tower



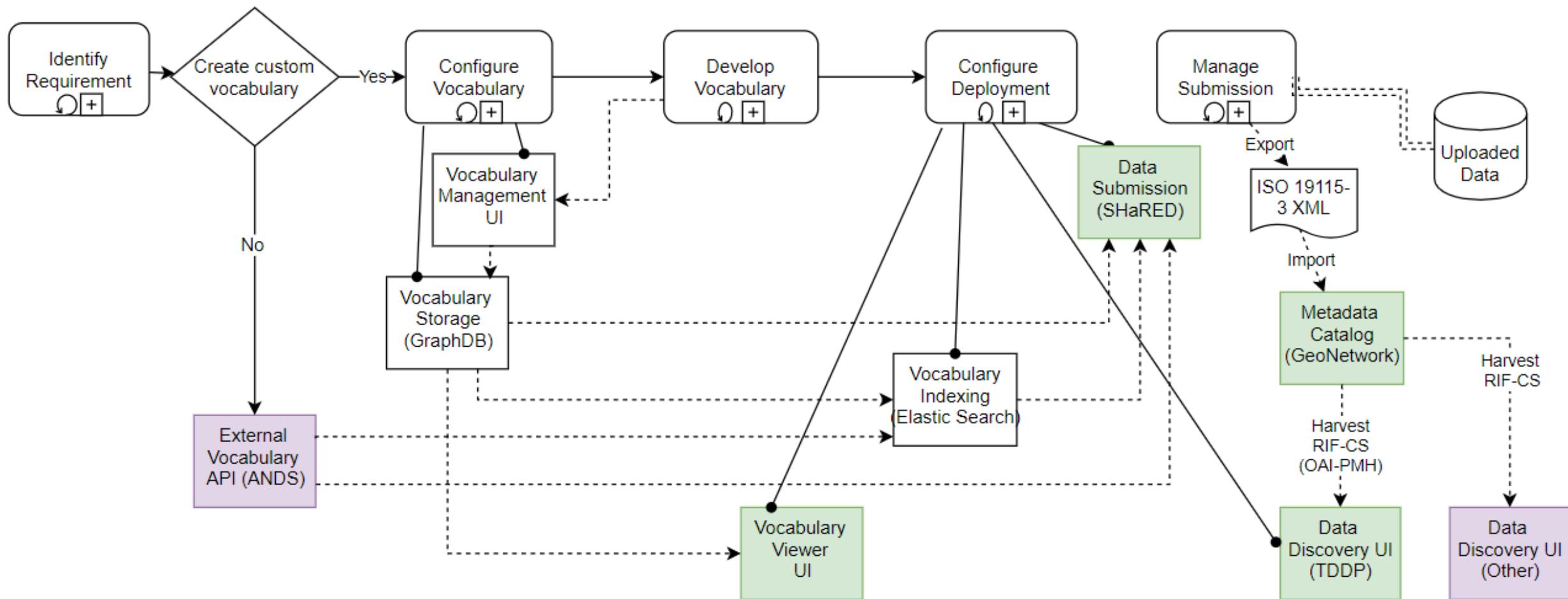
Data from Field Ecology



Process and Examples



To-Be Process



Viewer



[Home](#) [Vocabularies](#) [Concepts](#) [Observable properties](#) [Observation groups](#) [Methods](#) [Categorical variables](#)

AusPlots Rangelands Vocabularies

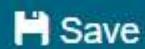
A Linked Data API for vocabularies encoded in SKOS.



Data Submission



SHaRED TERN SHaRED Data Submission Tool



j.mahuika@uq.edu.au / Demonstration for MDWG Draft

Last edited a few seconds ago

Data identification

What *

When *

Where *

How

Who *

About

Data sources

Lodge

There are multiple fields on this page that require your attention:

- Research theme keywords, GCMD Science Keywords: This field is required
- Abstract: This field is required
- Research theme keywords, ANZSRC Fields of Research: This field is required



Acknowledgement:
TERN
acknowledges
initial development
of the tool and
documentation by
the Australian
Ocean Data
Network (AODN)
and the Institute for
Marine and
Antarctic Studies
(IMAS).

SHaRED TERN SHaRED Data Submission Tool

j.mahuika@uq.edu.au /

 Archive

 Save

TERN AusPlots Forest

Monitoring Network - Large Tree Survey - 2012-2015

Submitted

Last edited 10 days ago

Data identification

What

When

Where

How

Who

About

Data sources

Lodge

1. Data Identification

Title *

TERN AusPlots Forest Monitoring Network - Large Tree Survey - 2012-2015

Clear and concise description of the content of the resource

Date the resource was created *

 23-03-2015

Topic Categories *



HUMAN DIMENSIONS > ENVIRONMENTAL GOVERNANCE/MANAGEMENT

ENVIRONMENTAL ASSESSMENTS



HUMAN DIMENSIONS

ENVIRONMENTAL GOVERNANCE/MANAGEMENT



Start typing to filter list...



 Browse

Research theme keywords, ANZSRC Fields of Research *

Select up to 12 research theme keywords describing your data

ECOLOGY



PLANT BIOLOGY



Start typing to filter list...



 Browse

Additional theme keywords

Enter your own additional theme keywords as required and click + to add

Tall Open Forests



Biogeography | Long-Term Community Monitoring | Long-Term Species Monitoring | Macroecology



5: How

Name	Description
Species identification	Field identification of all trees >10 cm diameter at breast height to genus and species level. Voucher specimens from unknown species (prefix UNN) were collected and submitted to local herbarium. [see Manual]
Tree Description	All trees >10 cm diameter at breast height were described by field observation. The following information was collected for each tree: [1] Tree Condition (i.e. Live or Dead); [2] Tree Status (Multi-stem, Buttressed, Burnt, Hollow, etc.); [3] Growth Stage (Regeneration, Regrowth, Mature, Senescent, etc.); [4] Crown Class (Suppressed, Intermediate, Co-dominant, Dominant, etc.); [5] Mode of Death (Standing, Broken, Anthropogenic, Burnt, Lightning, etc.). [see Manual]
Diameter measurement	For all trees >10 cm diameter at breast height, the diameter was measured (in centimeters) at an ascribed point of measurement with a diameter tape. The standard point of measurement was a height of 1.3 m, except in the case of buttressed or 'problem trees' whereby strict rules governed an alternative point of measurement. [see Manual]

7: About Dataset

Data parameters

Name	Units	Instrument	Serial No.	Platform
Tree Diameter	centimeter	--	--	AusPlots
Tree Condition	unitless	--	--	Ecology Plot
Tree Status	unitless	--	--	--
Growth Stage	unitless	--	--	--
CROWN	unitless	--	--	--

GeoNetwork

TERN AusPlots Forest Monitoring Network - Large Tree Survey - 2012-2015

The dataset comprises data from the first survey of ~24,000 large trees (>10 cm diameter at breast height; DBH) within 48 1 ha forest monitoring plots established across Australia between 2011 and 2015. Data includes: [1] Site identifiers (ID and Site Name); [2] Plot Establishment Dates; [3] Tree identifiers and descriptors (ID, Species, Status, Growth Stage, Crown Class); [4] Tree measurements (Diameter, Point of Measurement, Height, Location); [5] Comments and ancillary information; and [6] List of Metagenomic Sample Identifiers.

Completed

Download and links



Wood, S. W., Prior, L. D., Stephens, H. C., & Bowman, D. M. (2015). Macroecology of Australian Tall Eucalypt Forests: Baseline Data from a Continental-Scale Permanent Plot Network. *PLoS one*, 10(9), e0137811.

Open link





data_site_description.csv

Download

Data file

data_site_description.csv

About this resource

Categories

TERN Surveillance

 Environment

GCMD Earth Science and Earth Science Services Keywords

- [BIODIVERSITY FUNCTIONS](#) 
- [ENVIRONMENTAL ASSESSMENTS](#) 
- [ENVIRONMENTAL GOVERNANCE/MANAGEMENT](#) 

ANZSRC Fields of Research

- [ECOLOGY](#) 
- [PLANT BIOLOGY](#) 

otherKeywords

- [Tree Diameter](#) 
- [centimeter](#) 
- [AusPlots](#) 
- [Tree Condition](#) 
- [unitless](#) 
- [Ecology Plot](#) 

Data Discovery Portal

TERN AusPlots Forest Monitoring Network - Large Tree Survey - 2012-2015



Description

The dataset comprises data from the first survey of ~24,000 large trees (>10 cm diameter at breast height; DBH) within 48 1 ha forest monitoring plots established across Australia between 2011 and 2015. Data includes: [1] Site identifiers (ID and Site Name); [2] Plot Establishment Dates; [3] Tree identifiers and descriptors (ID, Species, Status, Growth Stage, Crown Class); [4] Tree measurements (Diameter, Point of Measurement, Height, Location); [5] Comments and ancillary information; and [6] List of Metagenomic Sample Identifiers.

Credit

Funding was provided by Education Investment Fund (EIF). LD Prior (UTAS) assisted in the study design. J Foulkes (UA), B Sparrow (UA) and I Fox (UA) provided administrative support.



Citation and Identifier

How to cite this collection:

Wood, S., Stephens, H., Bowman, D. (2020): TERN AusPlots Forest Monitoring Network - Large Tree Survey - 2012-2015. Terrestrial Ecosystem Research Network (TERN). doi: <https://doi.org/10.4227/05/552486484985D>

Rights and Licensing



Creative Commons Attribution 4.0 International License

ANZSRC - FOR

- ECOLOGY
- PLANT BIOLOGY

GCMD Science

- BIODIVERSITY FUNCTIONS
- ENVIRONMENTAL ASSESSMENTS
- ENVIRONMENTAL GOVERNANCE/MANAGEMENT

Parameters

- Tree Diameter (centimeter) AusPlots
- Tree Condition (unitless) Ecology Plot
- Tree Status (unitless)
- Growth Stage (unitless)
- CROWN (unitless)
- Bole Height (meter)
- Tree Height (meter)
- Soil Metagenomic Sample Identifier (unitless) AusPlots

Taxonomic Group

- *Acacia melanoxylon* | *Allocasuarina decussata* | *Allocasuarina torulosa* | *Corymbia intermedia* |



Current Status

Goals	Status
Improve data submission capabilities	SHaRED v3.0 Pilot Testing metadata migration and refining process
Adopt or develop controlled vocabulary to describe platform, instruments, Observable properties, UoM, Spatial and temporal resolution, organisations and people.	Adopted GCMD terms for spatial and temporal resolution Adopted QUDT terms for UOM* Developed organisations and people Work in Progress: platforms, instruments, Observable properties



tern.org.au

TERN Vocabs: <https://linkeddata.tern.org.au>

Data Access: <https://portal.tern.org.au>

Data Visualisation: <https://maps.tern.org.au>

Cloud and Virtual desktop platform: <https://coesra.tern.org.au>
<https://ecocloud.org.au>



Acknowledgements:

Supported by



Hosted by



In partnership with

