

Vocabularies for an infrastructure serving a heterogeneous community

Earth and Env Vocabularies and ontologies today: how are they managed? how are they used by scientist?

SciDataCon, October 2021

Siddeswara Guru

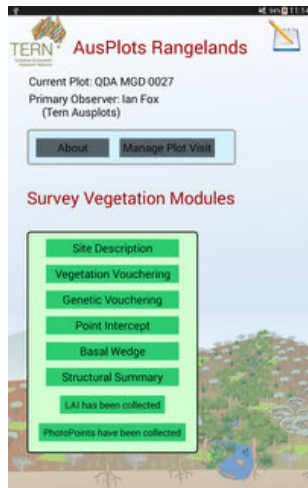
Program lead, TERN Data Services and Analytics Platform

s.guru@uq.edu.au

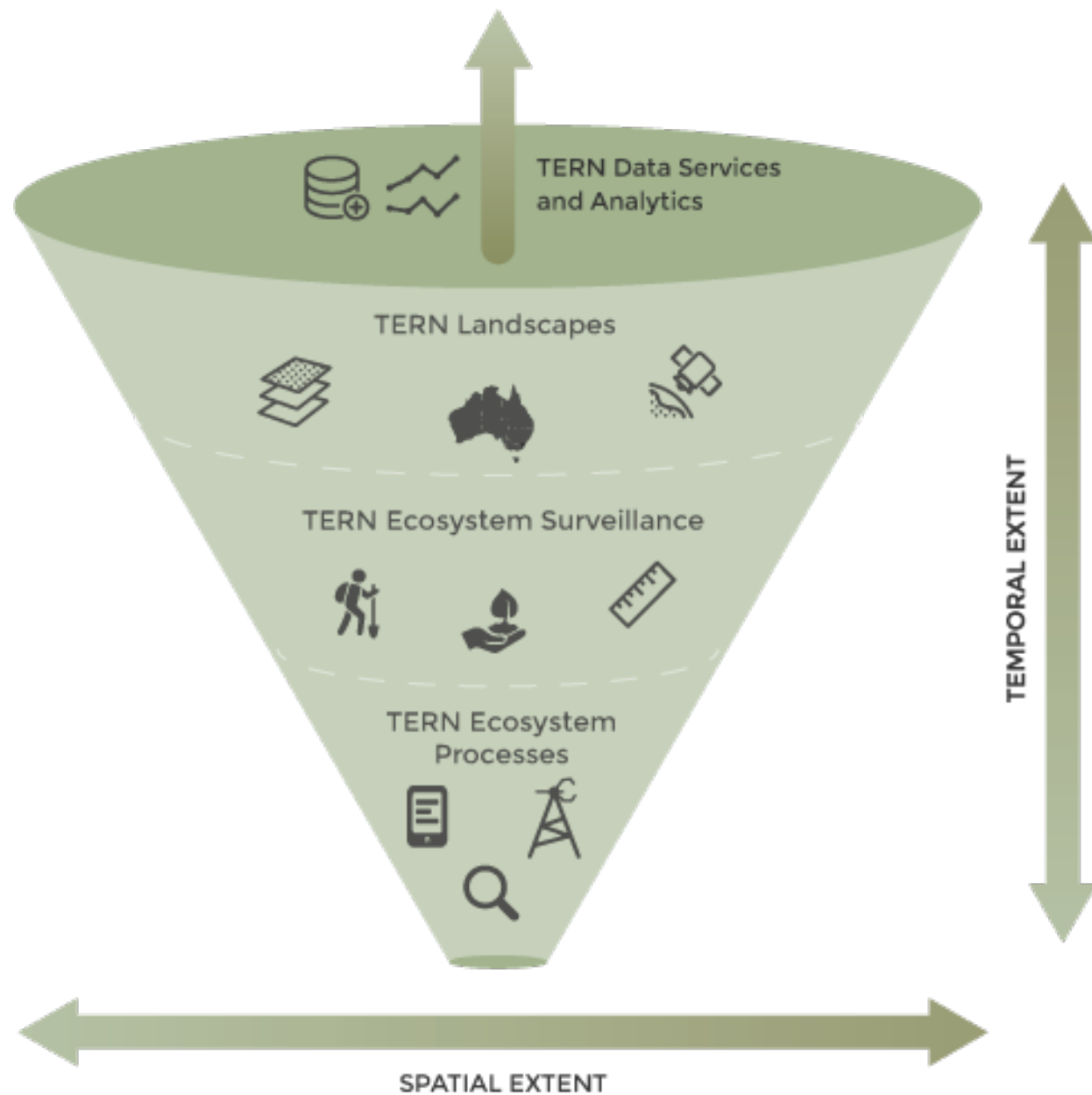


TERN Purpose¹

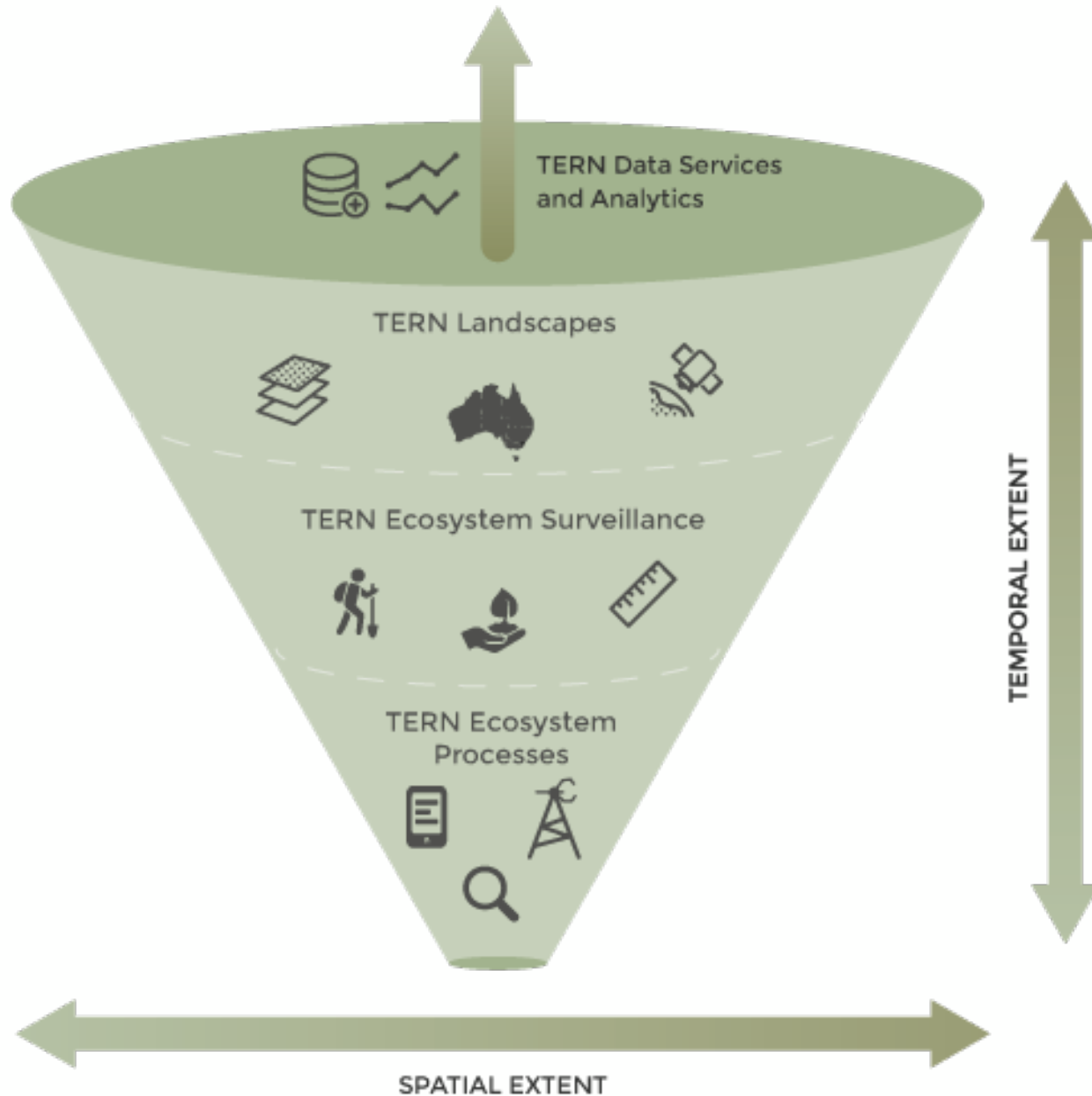
- National collaborative research infrastructure for collecting, collating, storing and sharing Australia's terrestrial ecosystem data sets and knowledge.

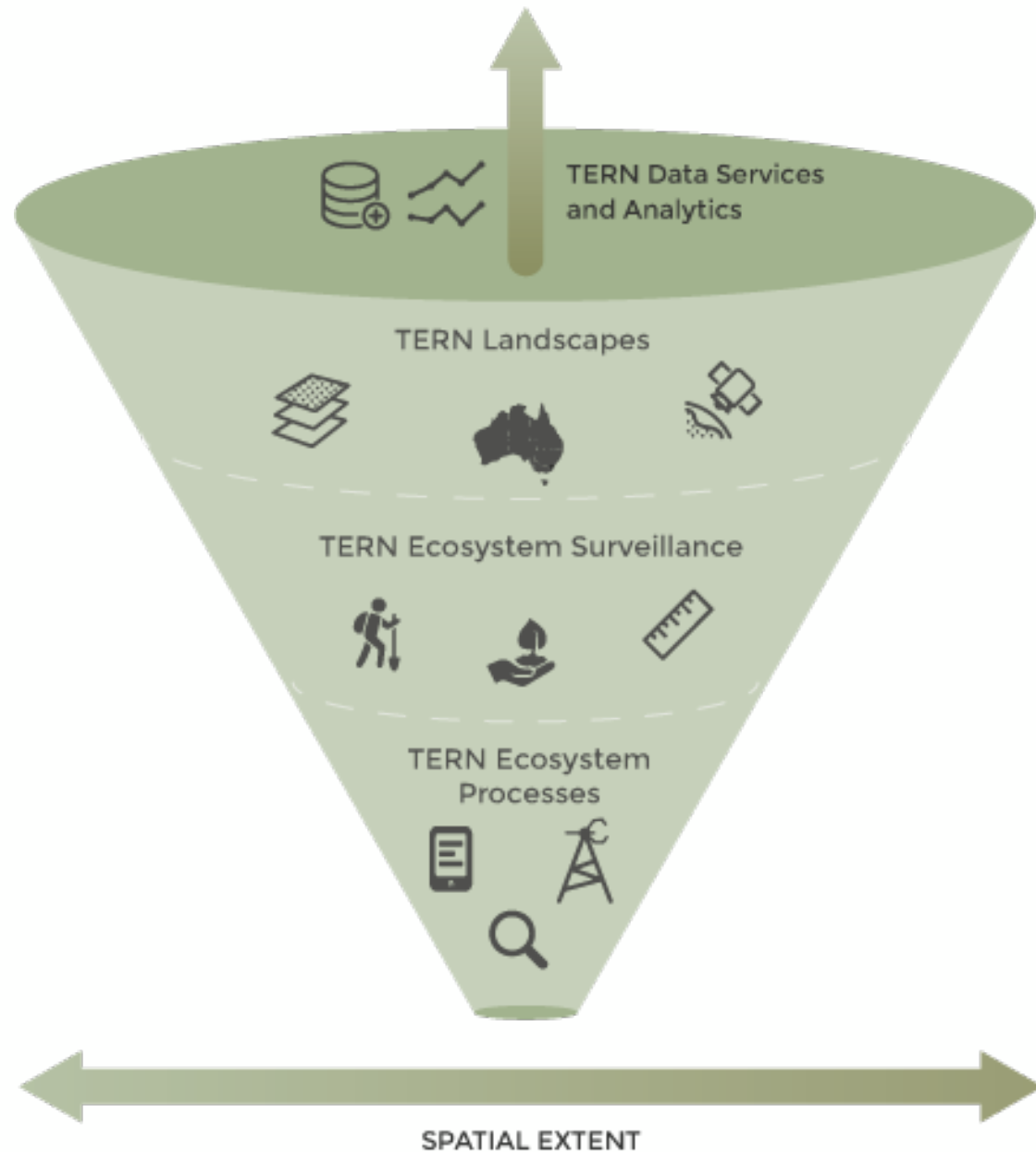


¹ TERN is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy from 2009.



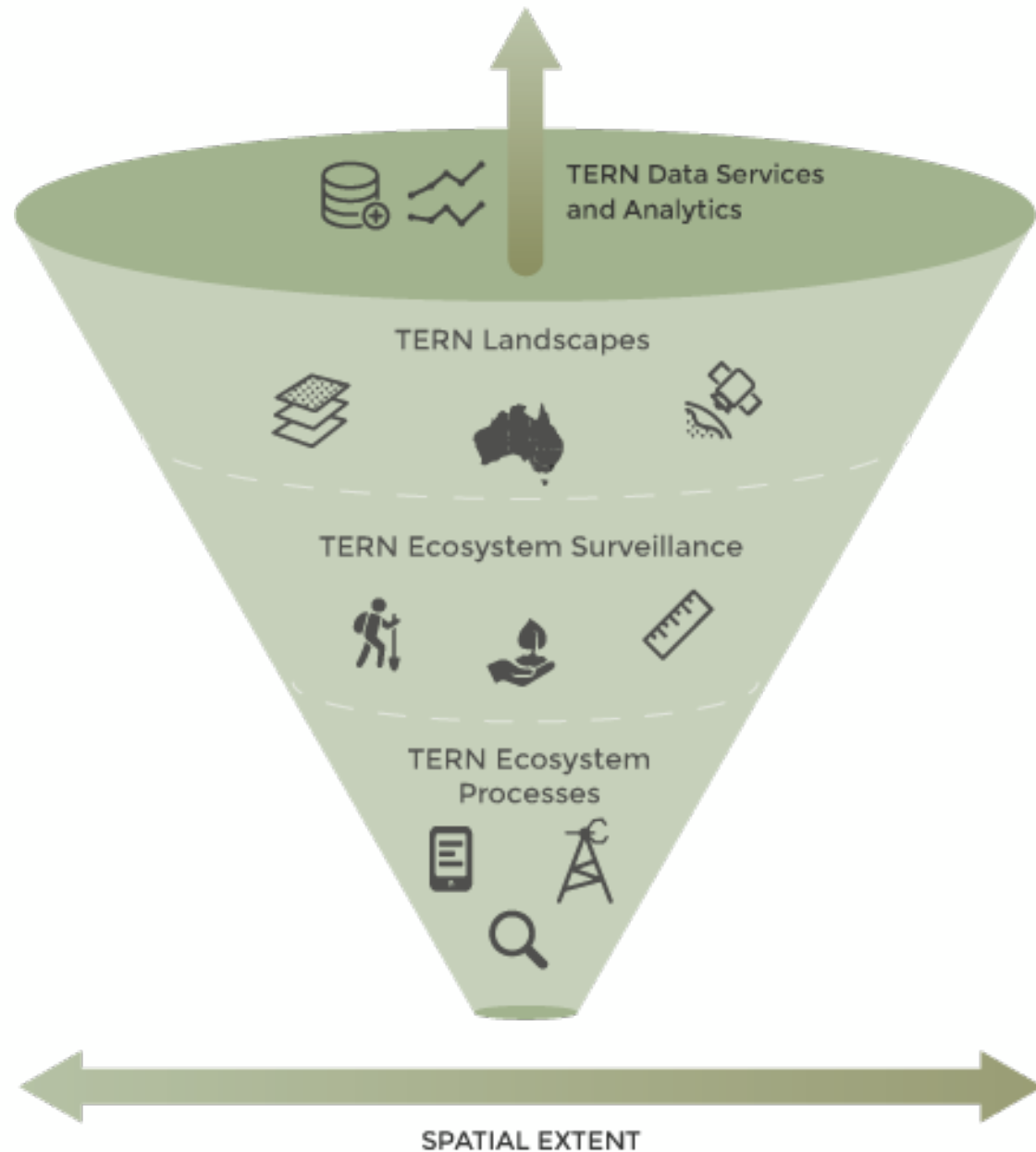
TERN in Operation





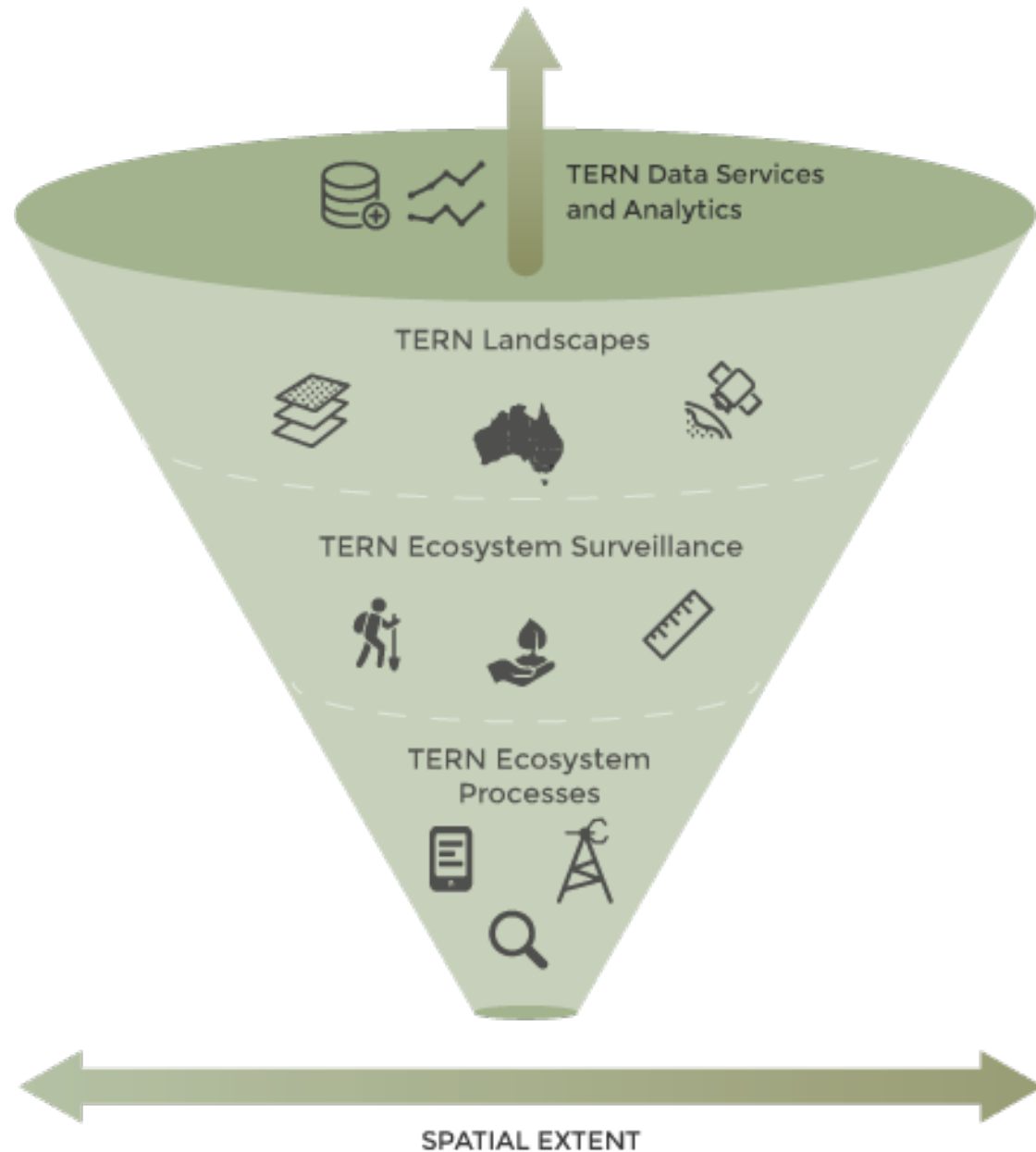
TERN in Operation

- Satellite remote sensing products
- Land cover dynamics and phenology
- Vegetation composition and structure
- Fire dynamics and impacts
- Continental Soil & Landscape data



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

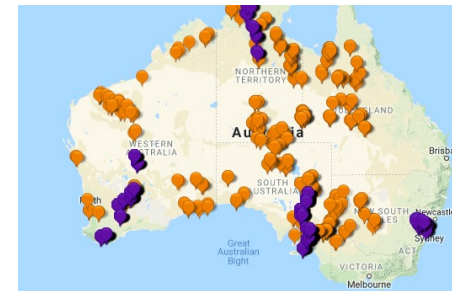
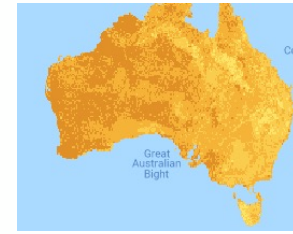


TERN in Operation

- 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

Data landscape of TERN

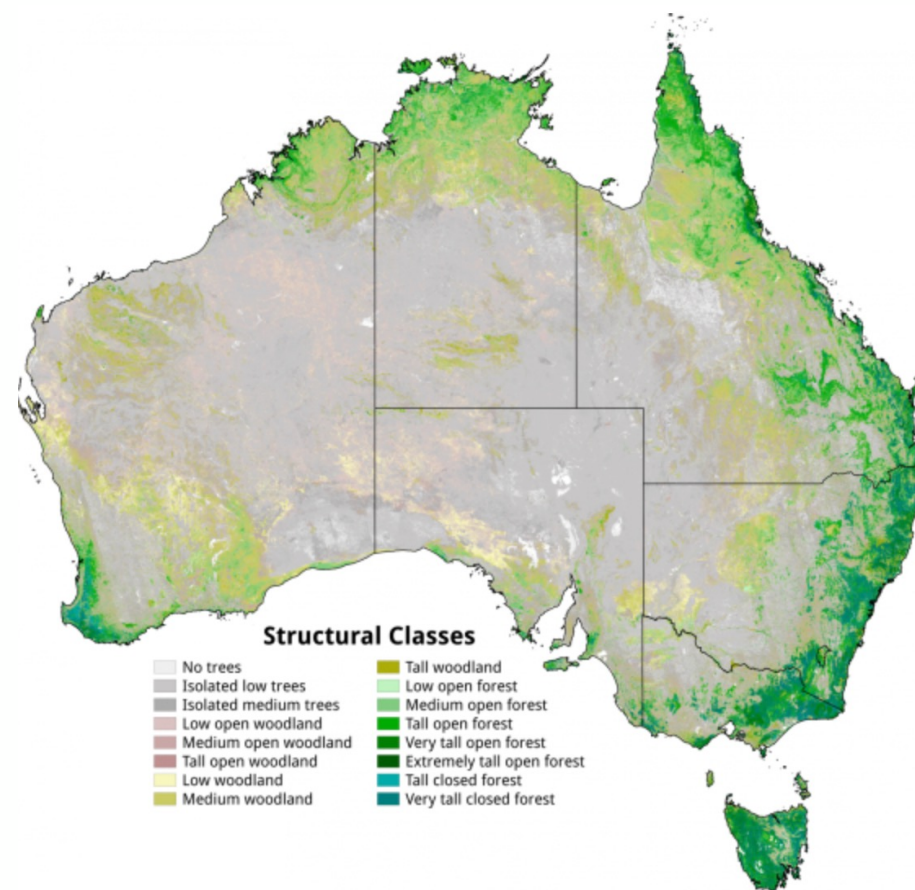
- Continental scale gridded data products : Remote sensing, Soil and landscape products
- Plot-based surveillance monitoring: Soils, vegetation - Human Observations
- Intensive monitoring
 - Flux tower - sensors
 - Phenocam - sensors
 - Acoustic monitoring - sensors
 - Plot-based vegetation monitoring – human observation
 - Calibration and validation data for remote sensing - sensor
- Institution survey data (state government agencies) – Human Observations



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 FORM ²	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 hermland CH	hermland H	open hermland OH	sparse hermland 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

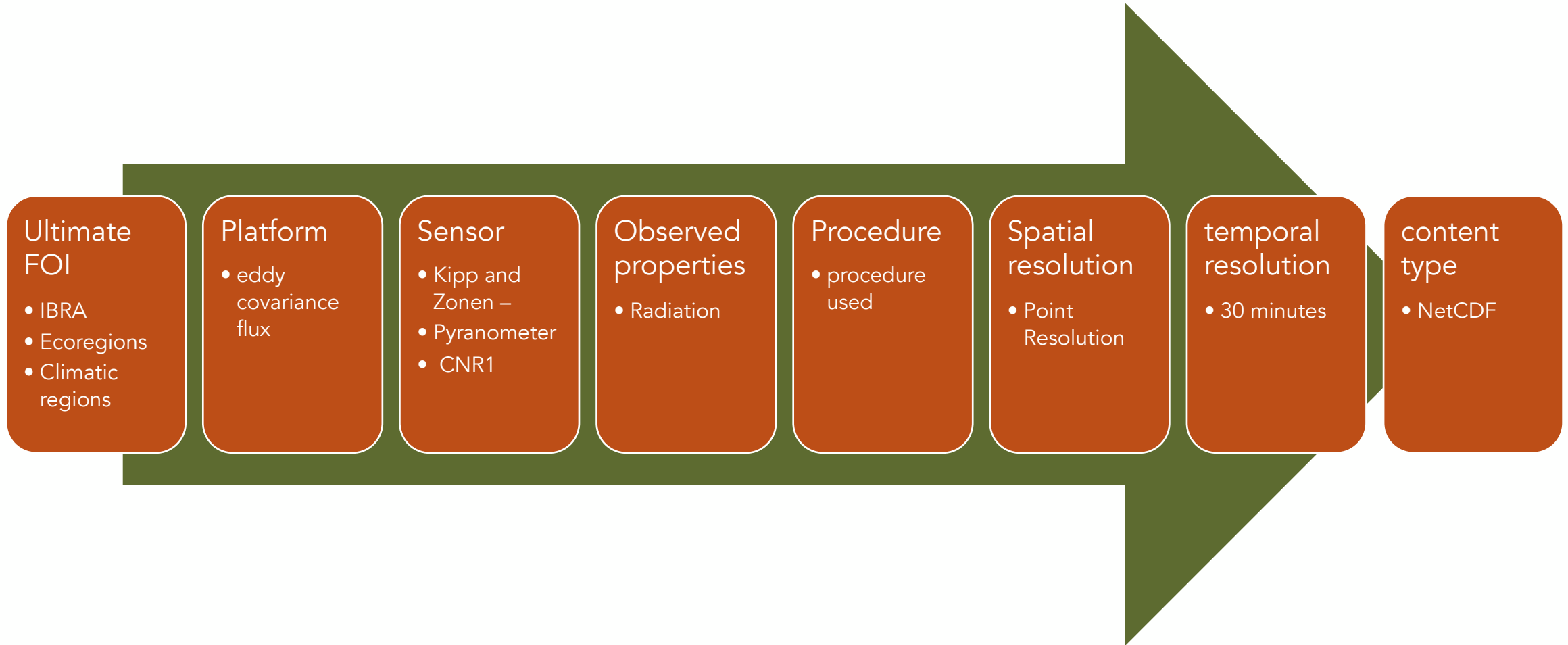


controlled vocabularies provide an opportunity to
harmonise at different scale.

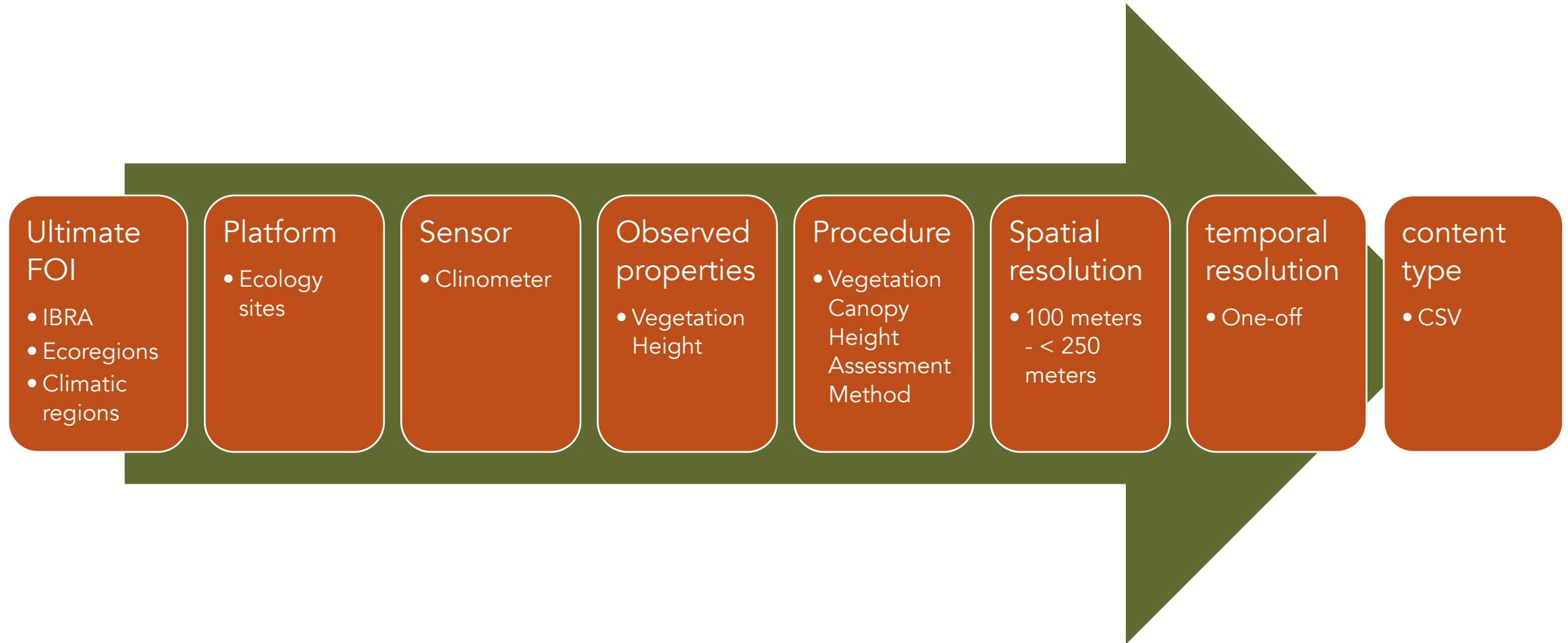
Harmonisation

- Vocabularies to describe artefacts
 - Platform -- based on SOSA
 - sensors – based on SOSA (some are SKOS-based)
 - Organisations – based on schema.org/ W3C org ontology
 - People – based on schema.org
 - Observed properties – parameter terms (SKOS-based)
 - Methods/procedures – Building Vocabularies for methods
- Spatial regions – Australia's Bioregions (IBRA), Ecoregions, States and Territories
- Spatial resolution -- RDF (align with GCMD terms)
- Temporal Resolution – RDF (align with GCMD terms)
- Content type – RDF (align with GCMD terms)
- UoM – QUDT ontology

Data from TERN Flux towers



Data from Field Ecology



Highlights

- Controlled Vocabularies
 - mixture of object of ontology classes and SKOS concepts
- Reuse of vocabularies
 - EnvThes, NERC, QUDT, CF, GCMD
- Fulfill business use cases
 - Asset registry
 - use the instances of sensors as an asset registry of all sensors deployed in TERN
 - same for platform
- Ability to track exact instance of a platform and sensors used in the data collection.

Conclusion

- Vocabularies are key for consistent representation of data and contextual information
- Improve data discoverability, access and reuse
- Ability to index data at several agreed terms
- Support interoperability between repositories
- Enable FAIR data.



tern

Ecosystem Research Infrastructure



We at TERN acknowledge the traditional owners and their custodianship of the lands on which TERN operates. We pay our respects to their ancestors and their descendants, who continue cultural and spiritual connections to country.

TERN is enabled by NCRIS.

Our work is a result of collaborative partnerships with many universities and institutions.

To find out more please go to **tern.org.au**.

Acknowledge to TERN Data Services Team, Simon Cox

TERN Linked Data: <https://linkeddata.tern.org.au/>

Research Vocabulary Australia <https://vocabs.ardc.edu.au/>

