### Evolution of data infrastructure for effective integration and management of environmental and ecosystem data

Siddeswara Guru, Gerhard Weis, Wilma Karsdorp, Andrew Cleland, Jenny Mahuika, Edmond Chuc, Javier Sanchez Gonzalez, Mosheh Eliyahu, Tina Parkhurst

### s.guru@uq.edu.au







# Outline

- Introduction to TERN
- Requirements
- Approach
- Summary





# TERN Purpose<sup>1</sup>

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





<sup>1</sup>TERN 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







# Users Requirements

- One place to submit and access data
- Need robust search capability
- Need to visualize data before downloading
- Need programmatic access to data
- Improve usability of data
- Ease of access and use
- Analytical capabilities



# TERN data landscape analysis

- Continental scale gridded data products : Remote sensing, Soil and landscape products
- Plot-based surveillance monitoring: Soils, vegetation
- 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 Observation



# Approach

- Objective 1: Build an integrated platform for plot-based ecological data
  - Build a scalable platform to host all plot-based ecological data
  - Provide an integrated view of data and ability to query parameters across different surveys and institutes.
  - Support data exchange between different initiatives.





- Objective 2: spatial and sensor data infrastructure to host TERN platform data
  - Effective management of Spatial data with Cloud-native solutions.
  - Data are described in a metadata standard
  - Use standard web services for data access
  - Storage resource to host large volume of data





- Objective 3: Improve discovery and access of data
  - Use ontology and controlled vocabs to describe platform, data providers, observed parameters, taxonomy.
    - Align to Semantic Sensor Network Ontology, extended to represent ecology plots and associated data collection.
  - Access to metadata authoring tool
    - ISO 19115-3 compliant metadata
  - Data search and indexing
  - Data visualization
  - Data access via API
  - Build a trust model for TERN data



- Objective 4: compute access closer to data
  - Make data cloud-enabled (data accessible to compute infrastructure)
- Objective 5: data skills program
  - Develop tutorials on accessing and using of data
  - Proactive engagement with users to address their needs
- Objective 6: data impact and user engagement
  - Measure the usability of data



# Data integration platform for plot-based ecology data



Different source databases



#### Spatial Data Infrastructure







### High level Architecture of Overall System







high-level infrastructure for information flow







#### **Engagement with different initiatives**





# Harmonisation

- Ontology and Vocabularies are key
  - Platform -- based on SOSA (align with GCMD)
  - Instruments based on SOSA (align with GCMD)
  - 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)
  - Project based on schema.org
  - UoM QUDT ontology
  - Observed properties Building our terms (align with EnvThes)
  - Methods/procedures Building Vocabulary
  - Organisations based on schema.org
  - People based on schema.org



### Information model should represent all the aspect data



![](_page_18_Picture_0.jpeg)

# **Analytical Platform**

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

### Cloud-based virtual desktop (coesra.tern.org.au)

- A Cloud-based virtual desktop environment accessible via a web browser Collaborative Environment for Scholarly Research and Analysi +61 (07) 3365 909
- Tools available
  - Data Science Jupyter labs, RStudio
  - Programming tool Canopy
  - Geospatial QGIS
  - Scientific Workflows: Kepler, KNIME
  - NetCDF viewer Panoply
  - Data cleaning Google refine
  - Ecology related Biodiverse, MacroecoDesktop
  - Data sharing Dropbox, ownCloud

![](_page_19_Picture_11.jpeg)

![](_page_19_Picture_12.jpeg)

Welcome to CoESRA

✗ What is this?	å Who can use?	Why CoESRA?
A free cloud-based virtual desktop workbench to	Researchers worldwide seeking a portable and	Easily accessible virtual desktop environment with a

![](_page_19_Picture_15.jpeg)

## In Summary

![](_page_20_Picture_1.jpeg)

![](_page_20_Picture_2.jpeg)

# **Overarching Operational Goal**

### For every data collected and collated by TERN

- Where it was collected place and platform
- How it is collected procedure
- Who collected organization/people
- What was collected observed property, UoM
- How can user access web access
- How can user use data Data Skill
- platform for analysis managed cloud analysis platform

![](_page_21_Picture_9.jpeg)

![](_page_21_Picture_10.jpeg)

### Acknowledgement: TERN Data Services Team

TERN data Access: <u>https://portal.tern.org.au</u> TERN Vocabs: <u>https://linkeddata.tern.org.au</u> CoESRA VDI: <u>https://coesra.tern.org.au</u> TERN GitHub: <u>https://github.com/ternaustralia/</u> TERN Skills: <u>https://ternaus.atlassian.net/wiki/spaces/TERNSup/overview</u>

![](_page_22_Picture_2.jpeg)

In partnership with

![](_page_22_Picture_4.jpeg)

![](_page_22_Picture_5.jpeg)

![](_page_22_Picture_6.jpeg)

![](_page_22_Picture_7.jpeg)

![](_page_22_Picture_8.jpeg)

![](_page_22_Picture_9.jpeg)