Delivering Spatial Ecosystem Data at Scale: TERN's Cloud Native Approach

eResearch Australasia 2025

23rd of October 2025

Gerhard Weis, Siddeswara Guru, Enzhen Luo, Tiancheng Lan

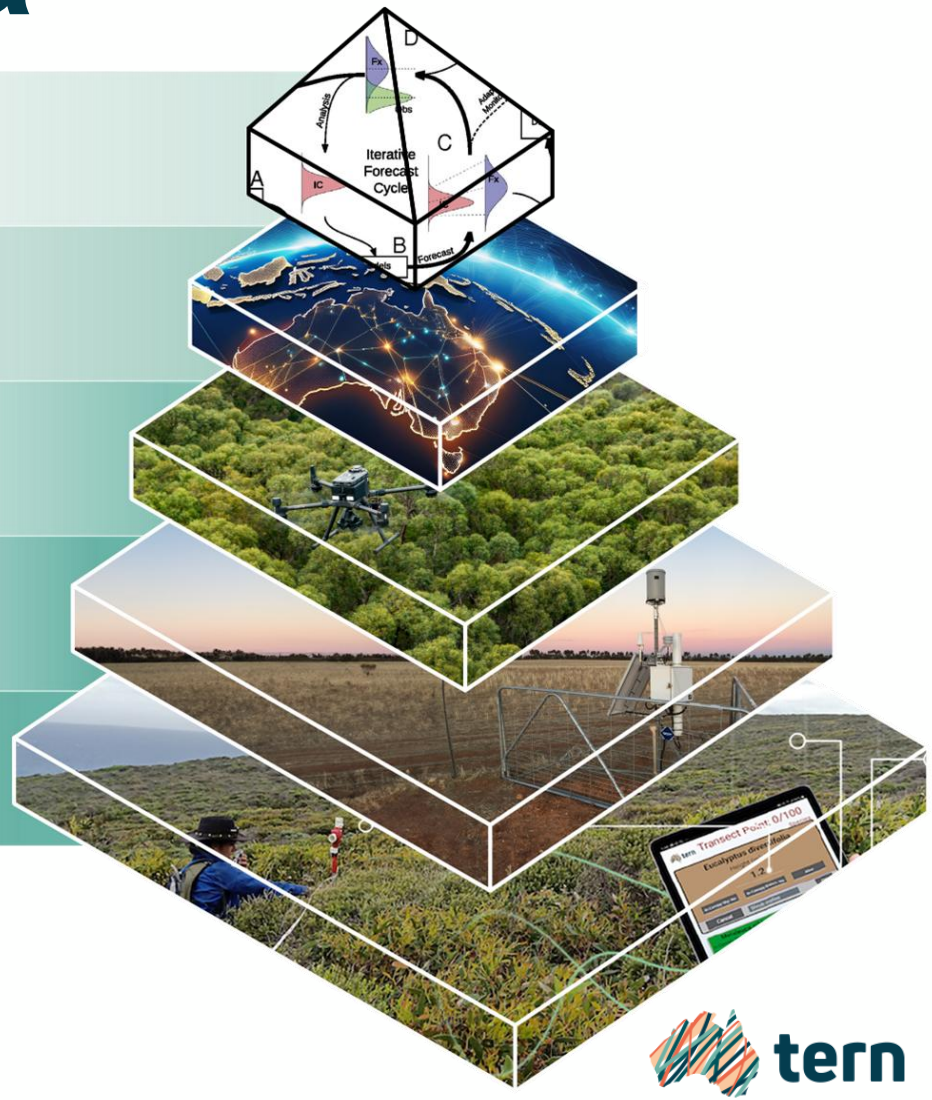
www.tern.org.au

We at TERN acknowledge the Traditional Owners and Custodians throughout Australia, New Zealand and all nations. We honour their profound connections to land, water, biodiversity and culture and pay our respects to their Elders past, present and emerging.

tern
Ecosystem Research Infrastructure

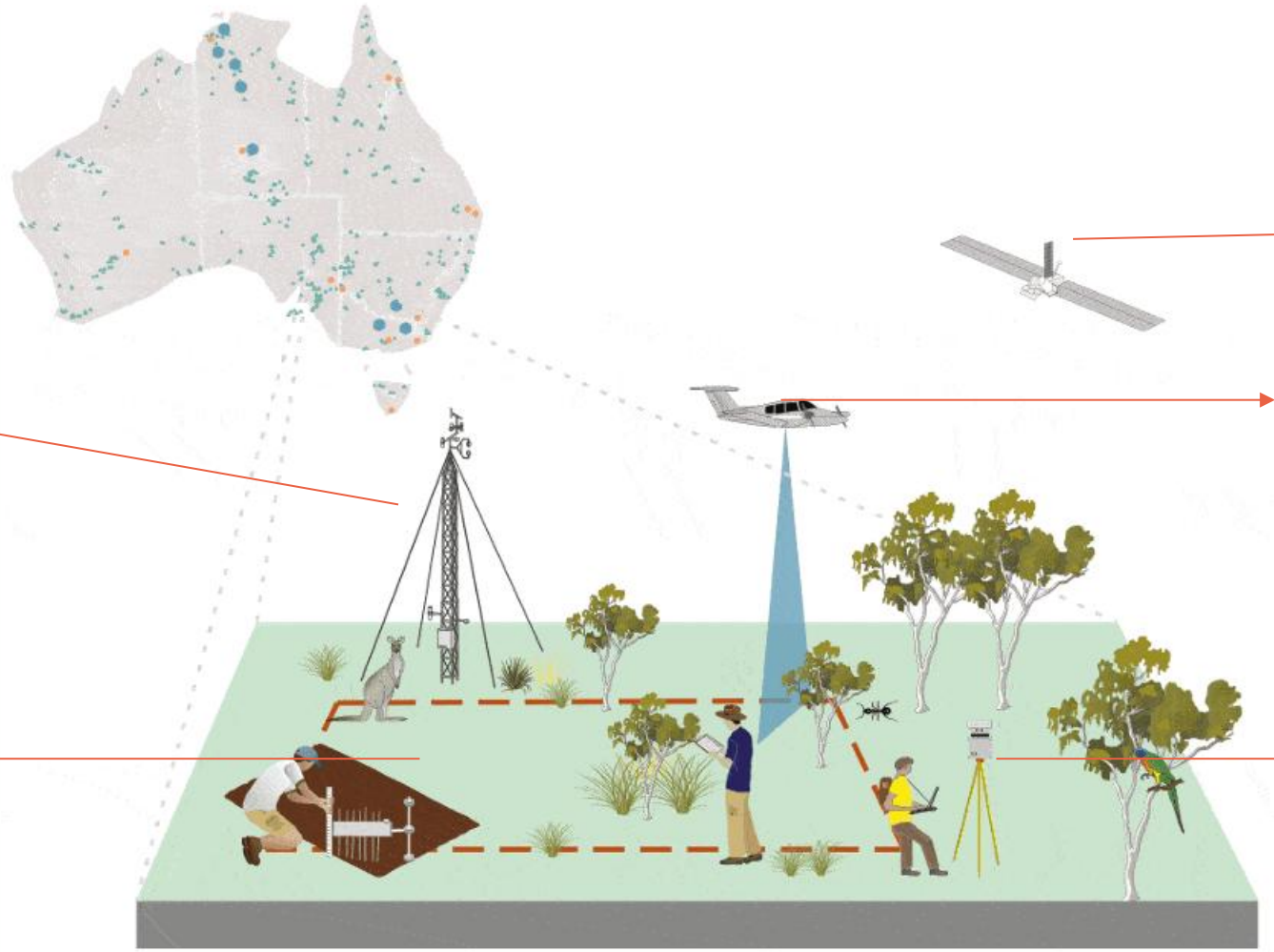
About TERN Australia

- It is national research infrastructure
- Funded by the Australian Government (NCRIS grant) since 2009
- It makes observations, collects and aggregates data that is multi-scale (continental, regional and local)
- Collects physical samples
- The data tools, apps, protocols and field sites are accessible to all
- It provides measures of change in **terrestrial & coastal ecosystems**
- Technical people, instruments and data are defined as research infrastructure



tern

Field-based observatory – Monitoring Approach



Remote sensing satellite

Fluxtowers + phenocam + Acoustic

Aerial Observation Platform/ UAV

Survey site

Terrestrial Laser Scanner

Data in numbers

- 25 million observations -> ~1B triples
- Acoustic 200TB – 3.5M files
- Images 4TB – 1.5M files
- Raster 80TB – 1.5M files (CoG)
- Flux data 10s GB – NetCDF
- Drone / LiDAR 5TB (RGb, Multispectral, PointClouds)
- Physical Samples - plant specimens, leaf tissue, soil (100k+)

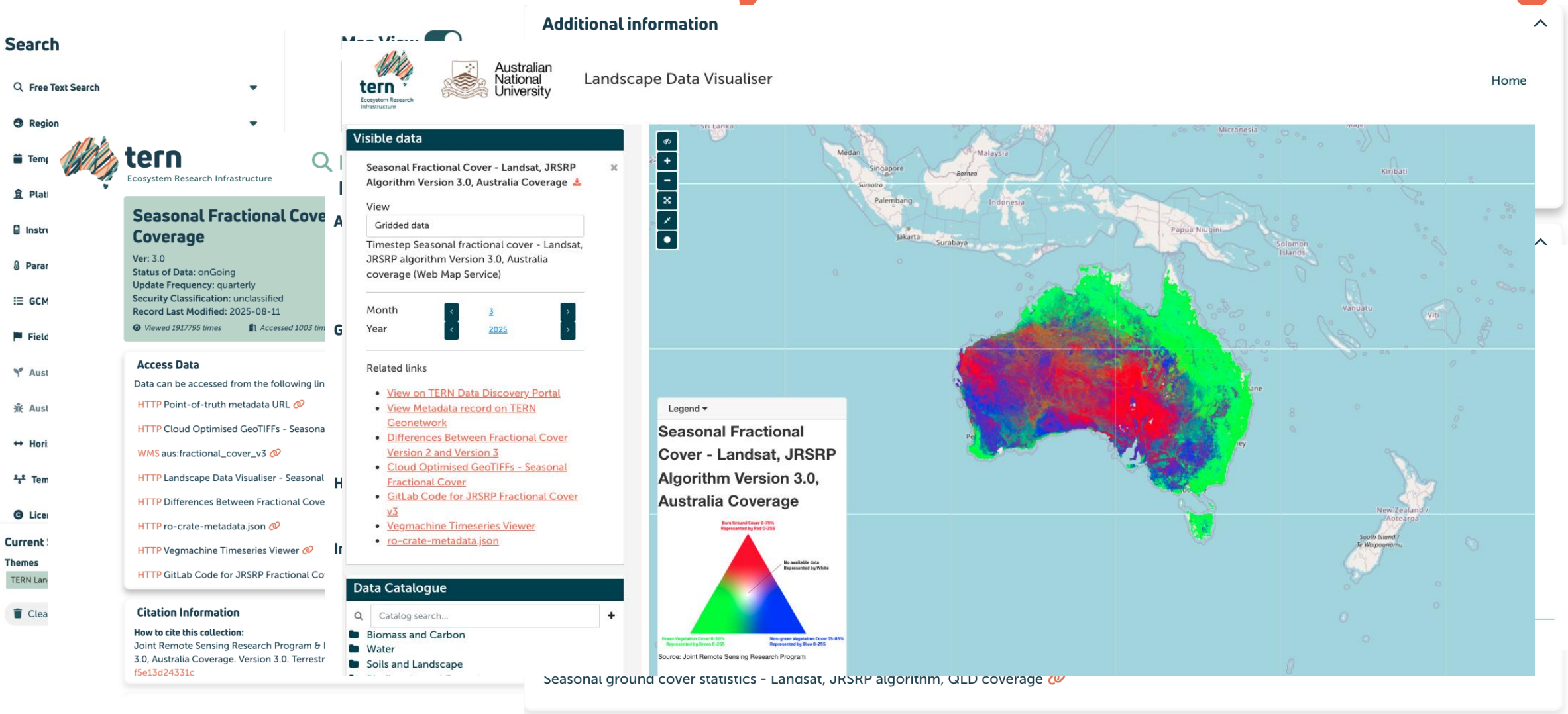
tern

TERN Data Management Practices

- Data and Metadata standards
  - ISO 19115-3 metadata standard, GeoNetwork metadata catalogue
  - DCAT to describe the catalogue
  - CV to describe platform, instruments, parameters, UoM, people and organisations
  - Multiple data standard depending on the type of data
    - Plot data exchange standard, OGC (WMS, WFS, WCS), OGC API, etc…
- file Formats in use
  - Raster – CoG , NetCDF
  - PointCloud – (CoPC)
  - Human observation – CSV, GeoJSON
  - Images - JPEG
  - Audio - FLAC

tern

TERN Data Discovery Portal – remote sensing



Cloud native approach

- Compute and software is a commodity, you just run it
- Data and configuration needs to be managed

- Data is directly accessible from anywhere (no need to bulk download first, just read the bits required)
- Machine 2 machine oriented

tern

STAC a cloud native approach

Spatio Temporal Asset Catalog

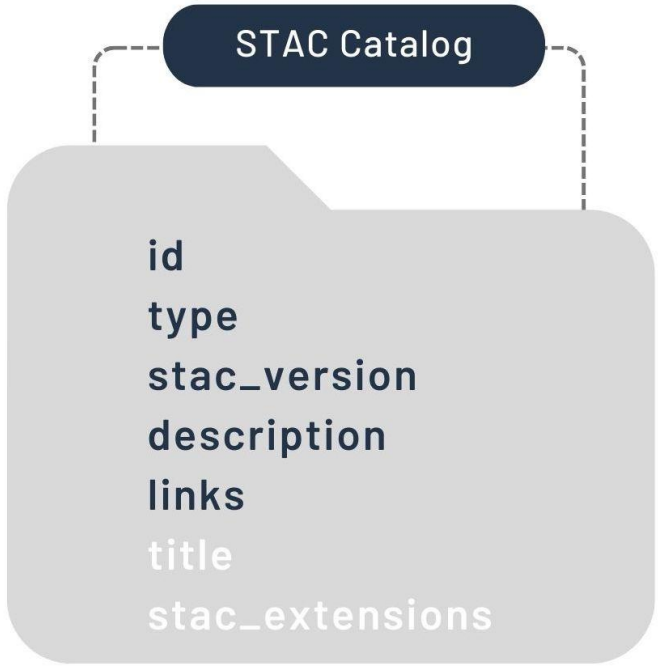
- Simple extensible JSON schema - community driven
- Geolocated assets only
- Static (files) and Dynamic (API)

STAC Elements

- Catalog – Folders
- Collections – Grouping of related items
- Item – Single spatiotemporal asset
- Asset – Actual data file, metadata file etc…

STAC Elements - Catalog

- Entry point to link to other things
- Think of it like a directory tree



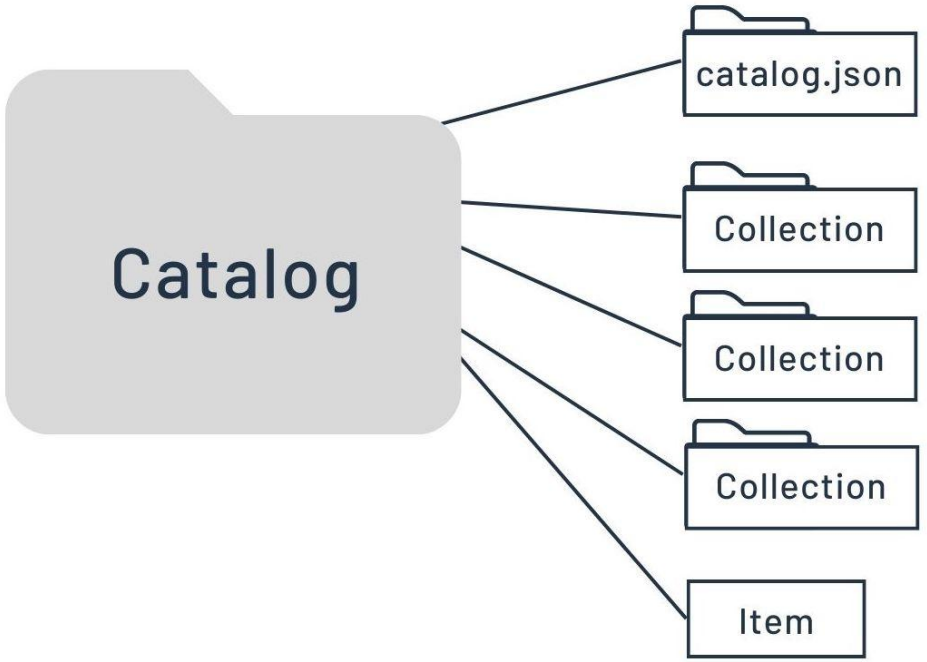


Image Source: https://stacspec.org/en/tutorials/intro-to-stac/

STAC Elements - Collection

- Builds upon catalog
- Describe set of related items
- Common metadata
- Overviews

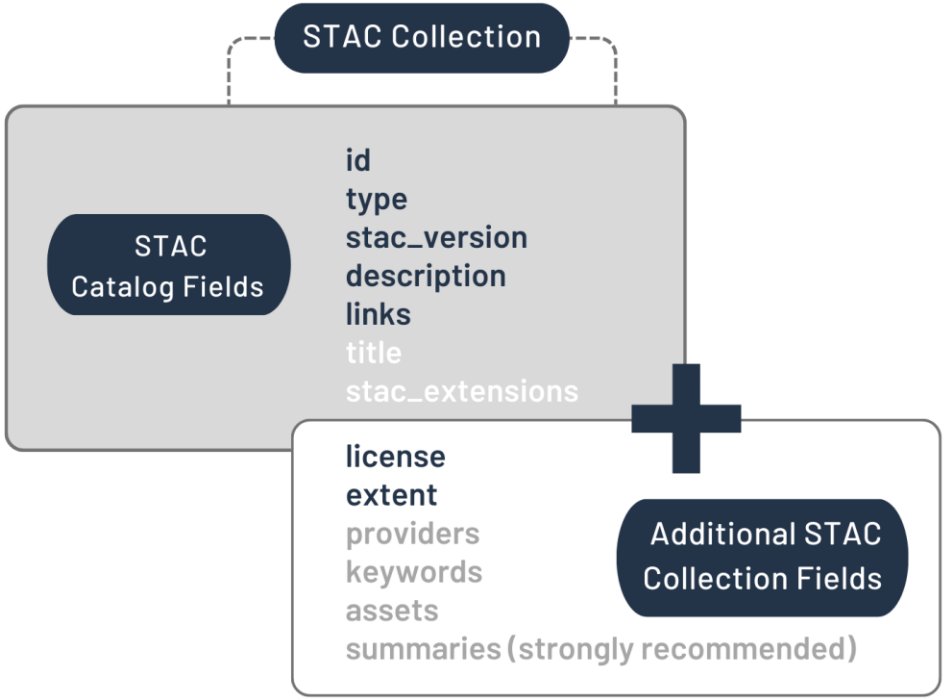


Image Source: https://stacspec.org/en/tutorials/intro-to-stac/

Stack Elements - Item

- GeoJSON feature
- Single spatiotemporal asset
- Asset – Actual data
- E.g.: Multiband raster; each band is one asset (file); one item to represent all bands

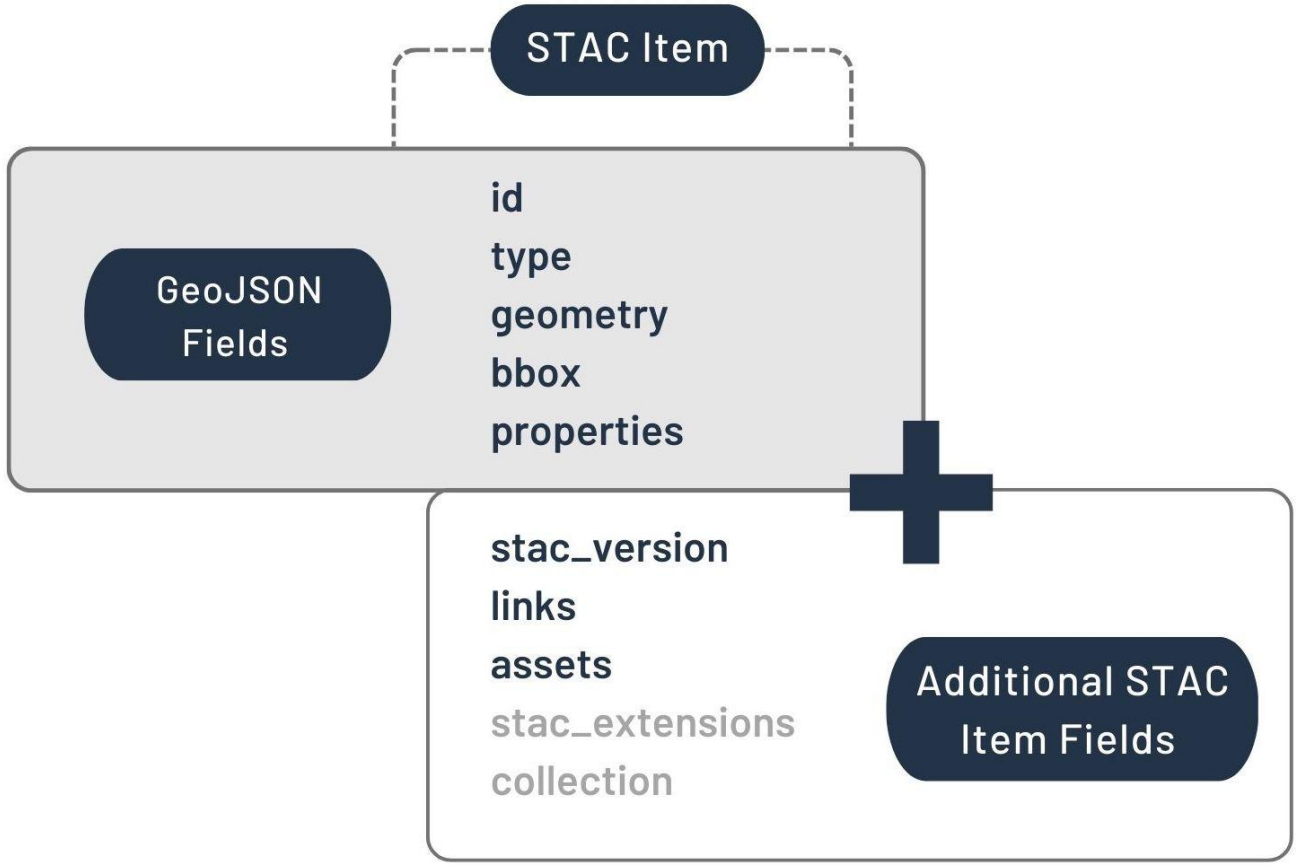


Image Source: https://stacspec.org/en/tutorials/intro-to-stac/

tern

STAC general structure

- Catalog <-> Collections <-> Item

- Link to other things (metadata, items, WMS, data assets)
- Geo location
- Properties
- Data structure (Raster bands, dimensions, variables, …)

- https://stacspec.org

tern

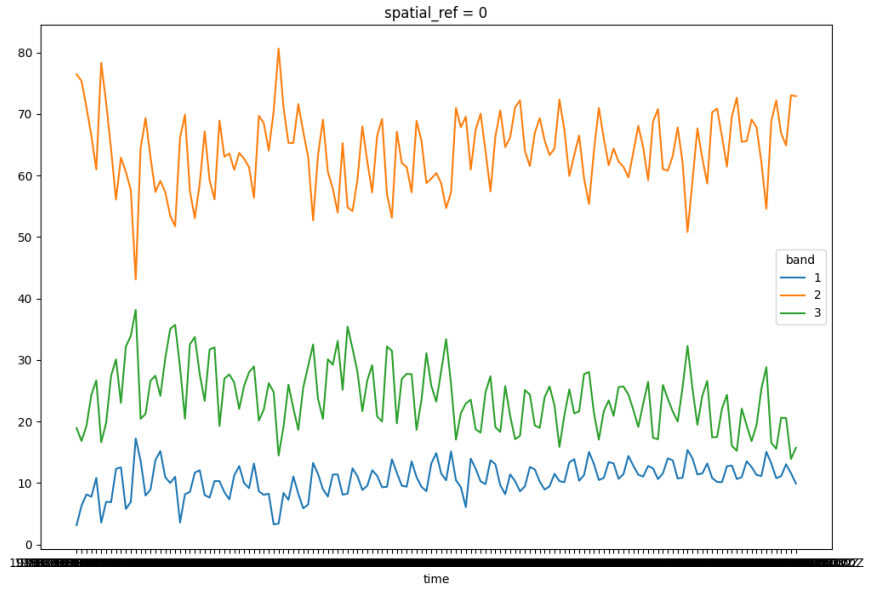
STAC benefits

- Easy to generate
- Growing support via applications tools and libraries (QGis, GDAL, numpy, xarray, pandas, OpenDataCube, etc… - there is R support as well)
- Community driven governance
- Can be managed statically
- STAC API for dynamic access and programmatic discovery
- Link to data, metadata, related items … (HTTP, OGC-WMS, citation, ISO19115-3, DCAT)

tern

Example – Cloud native fzractional cover

- https://github.com/ternaustralia/TERN-Data-Skills
- Seasonal Fractional Cover Dataset – 30m resolution (Landsat)
  - There is also a 10m version based on Sentinel
- 1987 -> now
- 147 time steps, 3 bands, 135k x 140k pixels
  - 34TB data.
- Steps
  - STAC collection to find assets
  - Load structural file metadata,
    slice to Brisbane area (6 seconds)
    - 147 time steps, 2k x 3k pixels (10GB)
  - Average over area and plot graph (3 minutes)



Resources

- Data Discovery : https://portal.tern.org.au
- Data Submission Tool: https://shared.tern.org.au
- EcoPlots: https://ecoplots.tern.org.au
- EcoImages: https://ecoimages.tern.org.au
- Data Visualiser: https://maps.tern.org.au
- CoESRA VDI: https://coesra.tern.org.au
- Field Sample Library: https://www.tern.org.au/field-sample-library/
- Field Survey Protocol and App: https://www.tern.org.au/field-survey-protocols-apps/
- AusplotsR: https://cran.r-project.org/web/packages/ausplotsR/ausplotsR.pdf
- Controlled vocabularies: https://linkeddata.tern.org.au/viewer/tern/
- TERN Data Skills: https://github.com/ternaustralia/TERN-Data-Skills
- User Support: https://ternaus.atlassian.net/wiki/spaces/TERNSup/overview
- STAC: https://data.tern.org.au/catalog.json

tern

tern
Ecosystem Research Infrastructure

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

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.**

Acknowledgement: TERN Data Services and Analytics Team, NeCTAR Cloud Infrastructure