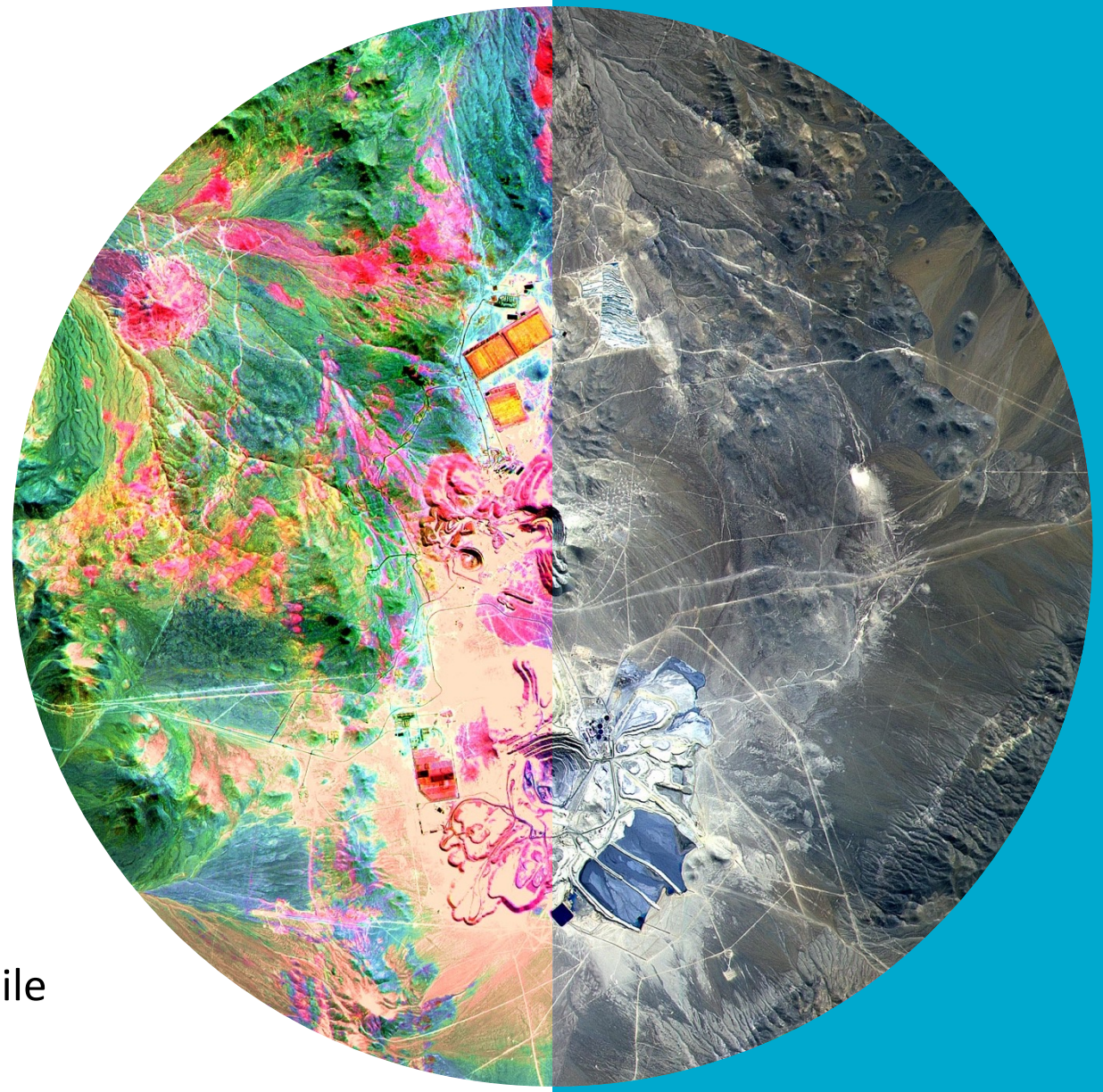
CSIRO

Chile

Earth Analytics Interoperability Lab (EAIL)

Jonathan Hodge
Program Director, CSIRO Chile

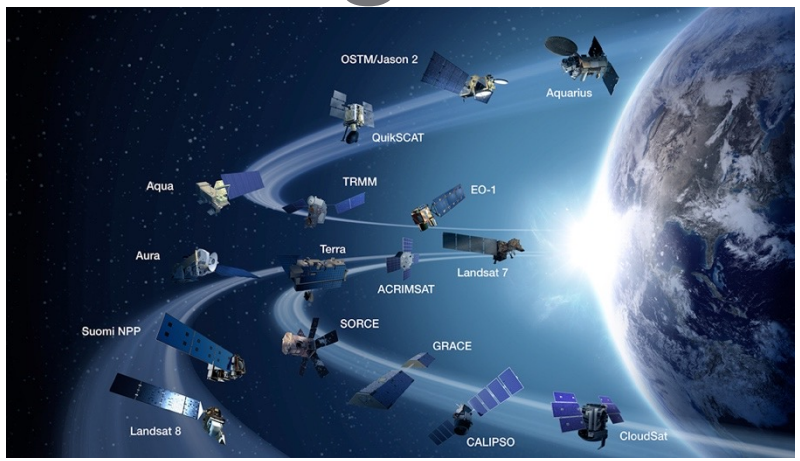


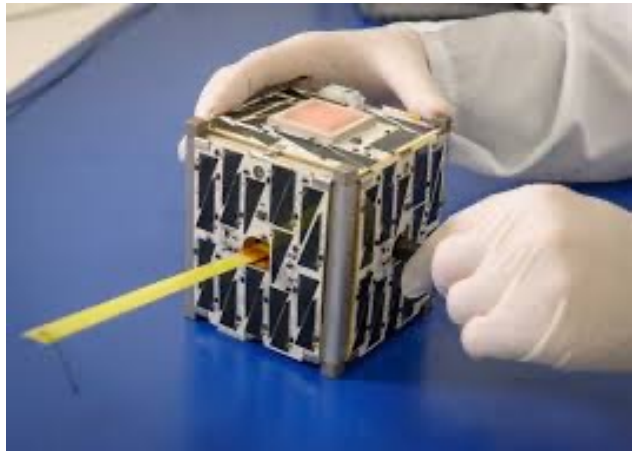
Australia's National Science Agency

Mina Escondida, Región de Antofagasta, Chile

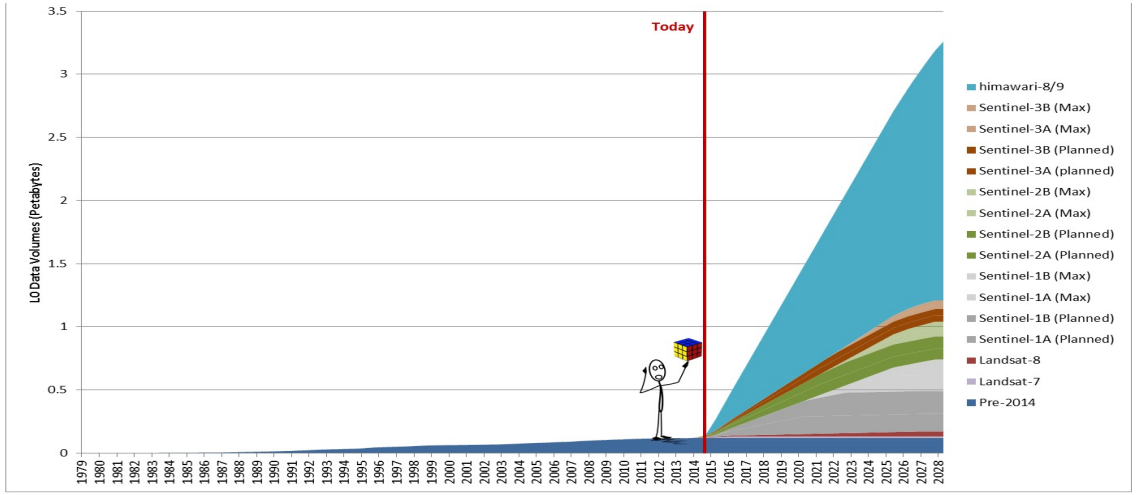
Images courtesy NASA and U.S./Japan ASTER Science Team

Challenge and Opportunity





- Over 400 EO sensors next decade
- New affordable satellites
- Sensor to user



Goal

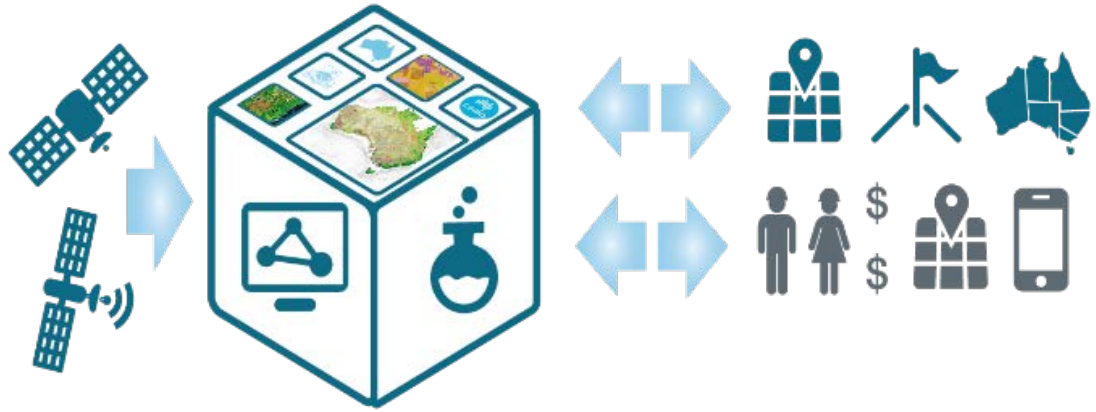
Everyone has access to the best:

1. Data volumes and access
2. Processing and correcting
   - Analysis Ready Data
3. Products/algorithms and validation
   - Science literacy and use

Common and consistent data for everyone

Before: Individuals invest in entire value chain

After: Data supply and big data analysis platform is shared. Individuals focus on delivery of specific value



Data cube

A scalable platform for Data and Science

Data preparation and access	Data cube	Products and users
• National archives • Instrument experts • Atmospheric correction • Data custodians **Output:** • **Analysis Ready Data**	• Prepared data sets (ARD) • Spatially-aligned pixels • Programming interface • Desktop, Cloud and HPC **Output:** • Search and read data • Large data array (cube)	• **Do science** • Combine data • Scale task **Output:** • Data products • Share applications

Spatial data infrastructures

Open Data Cube

Rasdaman

Google Earth Engine

Thematic Exploitation Platforms & Copernicus Hubs

SciDB

Goals

- Increase the impact and value of satellite data
- Provide an EO data analysis architecture
- Foster a community to grow the technology and applications

An Overview of Platforms for Big Earth Observation Data Management and Analysis - https://www.mdpi.com/2072-4292/12/8/1253

- Abstraction (data, processing, infrastructure)
- Scalability (storage, processing}
- Interoperability
- Extensibility
- Infrastructure requirements
- Reproducibility of science
- Governance

OPEN DATA CUBE

Where are the Data Cubes?

OpenDataCube.org

10 operational cubes (Australia, Colombia, Switzerland, Taiwan, Vietnam, Kenya, Tanzania, Ghana, Sierra Leone, Senegal)
67 data cubes in development – includes 49 additional countries in the Digital Earth Africa initiative
28 countries expressing interest or reviewing data cubes

Partners and platforms

Partners and Associates	Platforms and Initiatives
NASA (CEOS SEO)	CEOS Earth Analytics Interoperability Lab
Geoscience Australia	Digital Earth Australia Digital Earth Africa
CSIRO	Earth Analytics Science Innovation Hub
USGS	Landsat cloud archives
UK Catapult	Common Sensing
Switzerland	Swiss Data Cube
VNSC	Vietnam Data Cube
IDEAM	Colombia Data Cube

Open Earth Alliance
https://www.openearthalliance.org

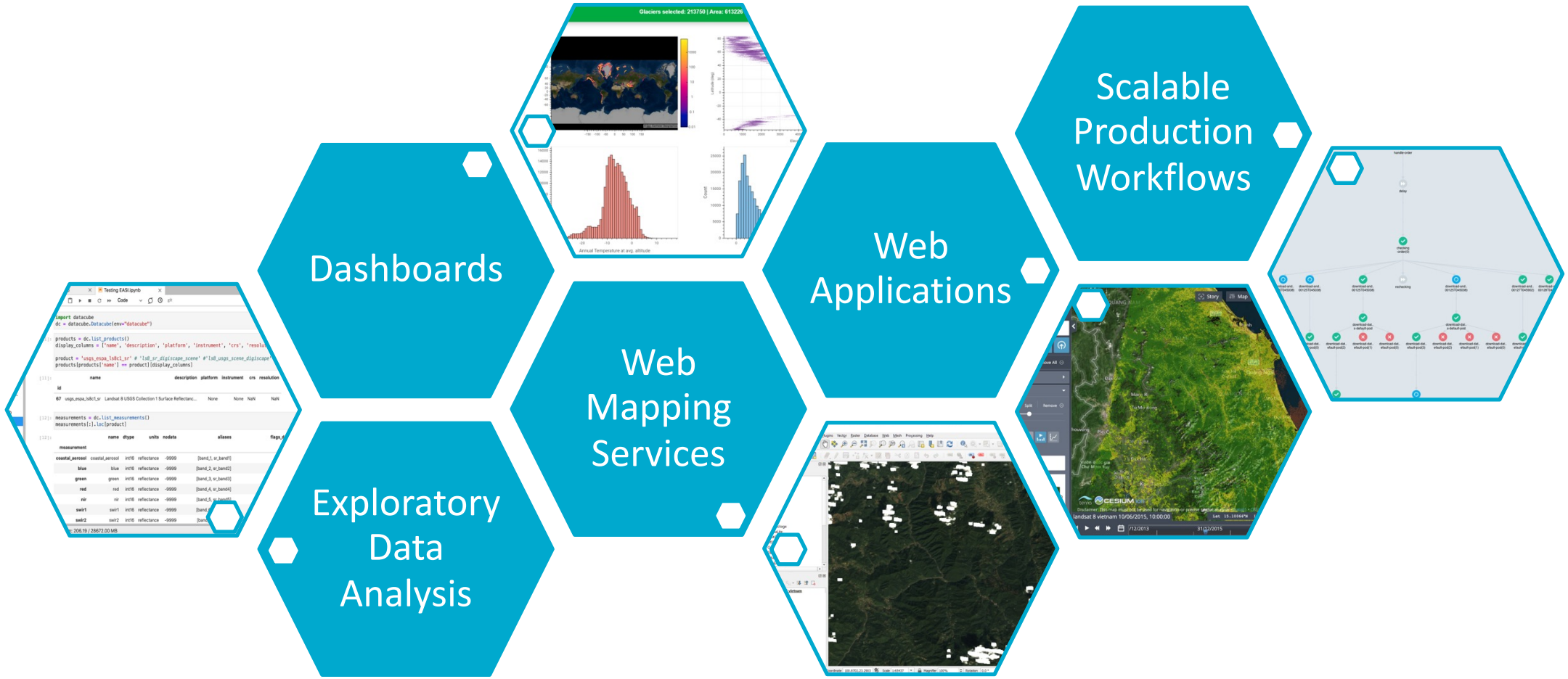
Open Data Cube
https://www.opendatacube.org

What is the Open Data Cube?
https://medium.com/opendatacube

Powered by OPEN DATA CUBE

CSIRO

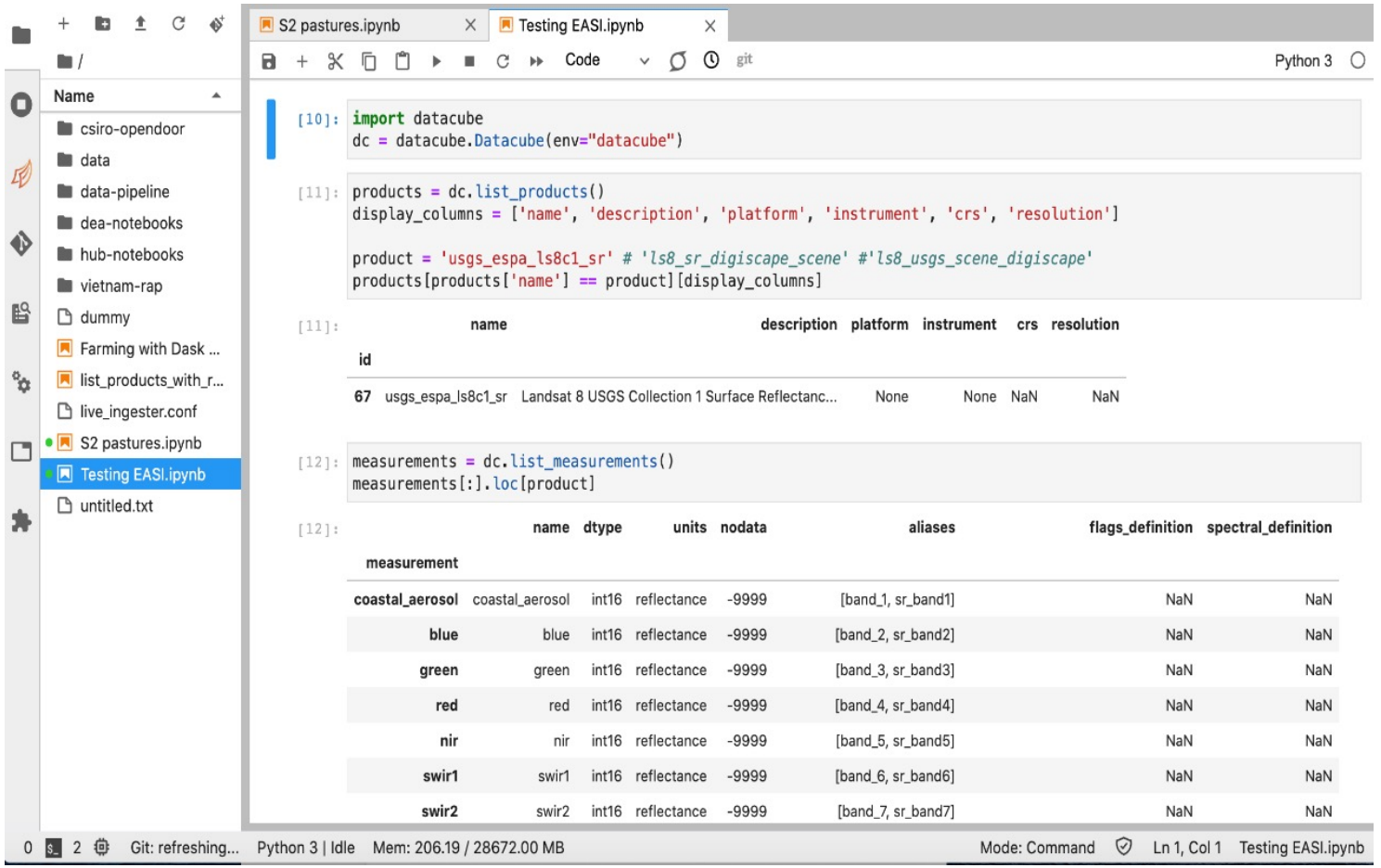
EASI is an ecosystem with a range of interfaces



Powered by Open Data Cube and the Python data science ecosystem

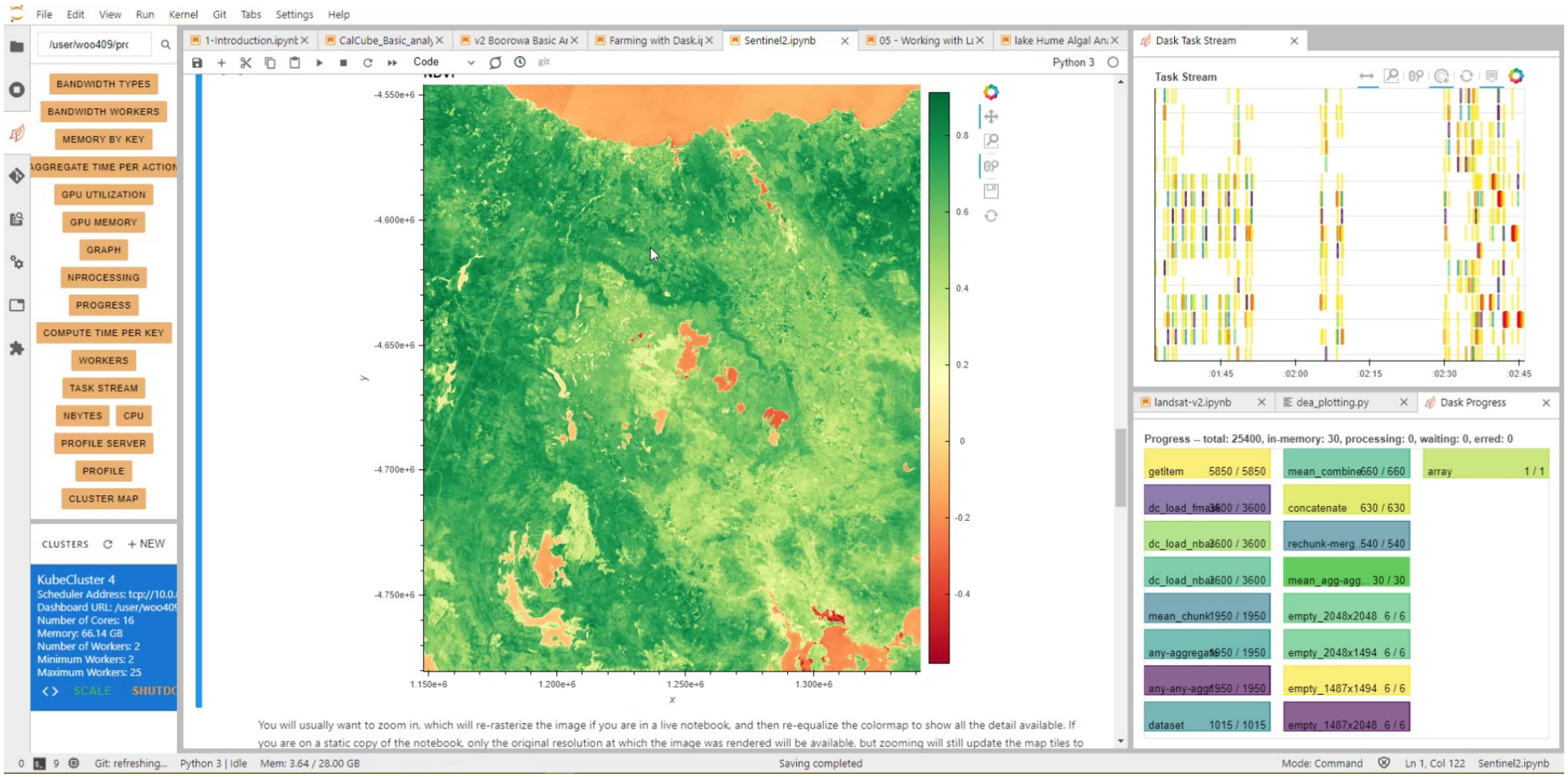
Exploratory data analysis with Jupyter Lab

- Interactive analysis with python notebooks
- User's have own project space (backed-up)
- Pre-configured for use with system datacube
- Copy and save notebooks via Git
- Per User Computing Clusters
- Interactive User Interfaces in a few lines of code

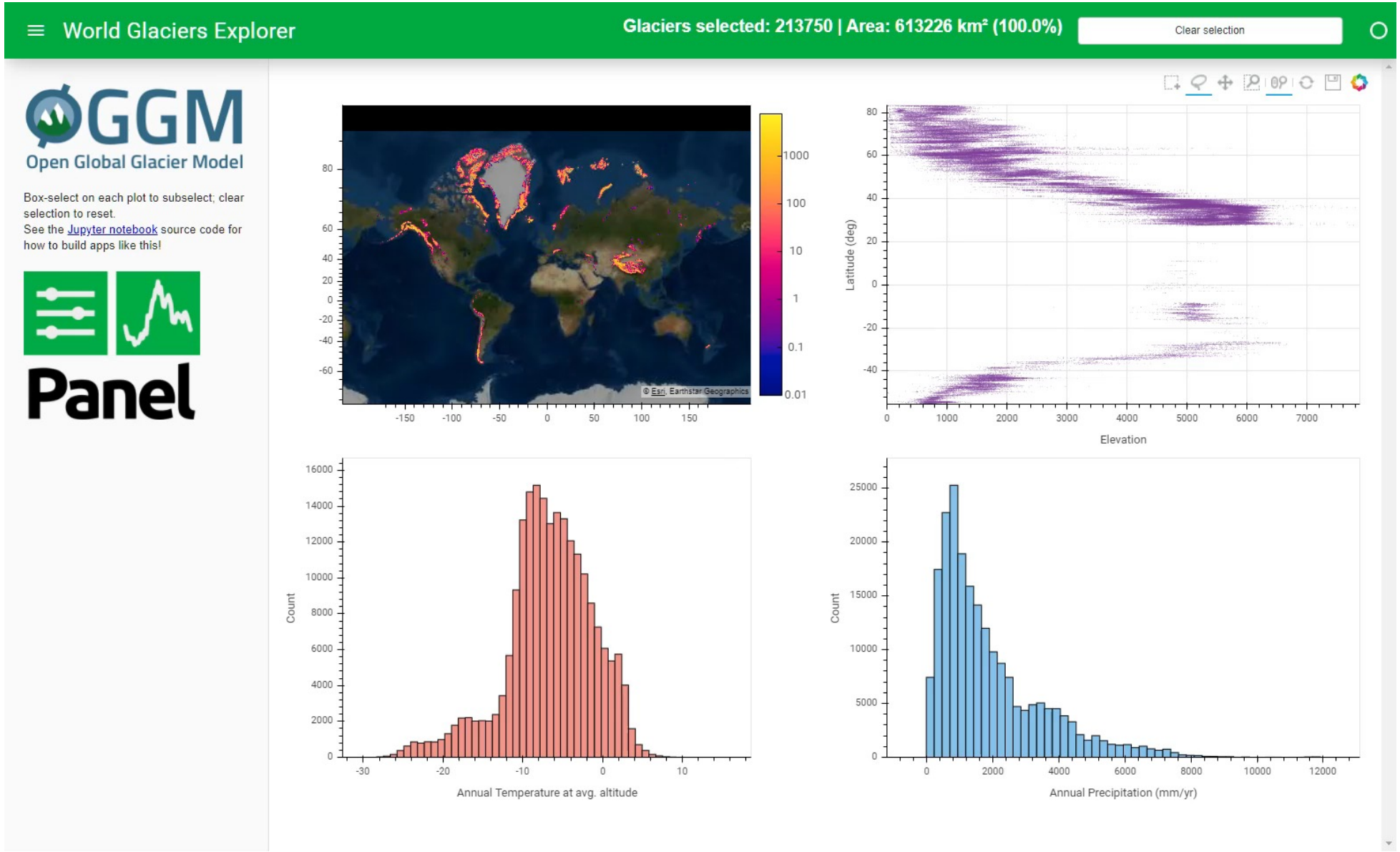


https://jupyterlab.readthedocs.io/en/latest/user/interface.html

EDA with Jupyter Lab – Scales with Cluster

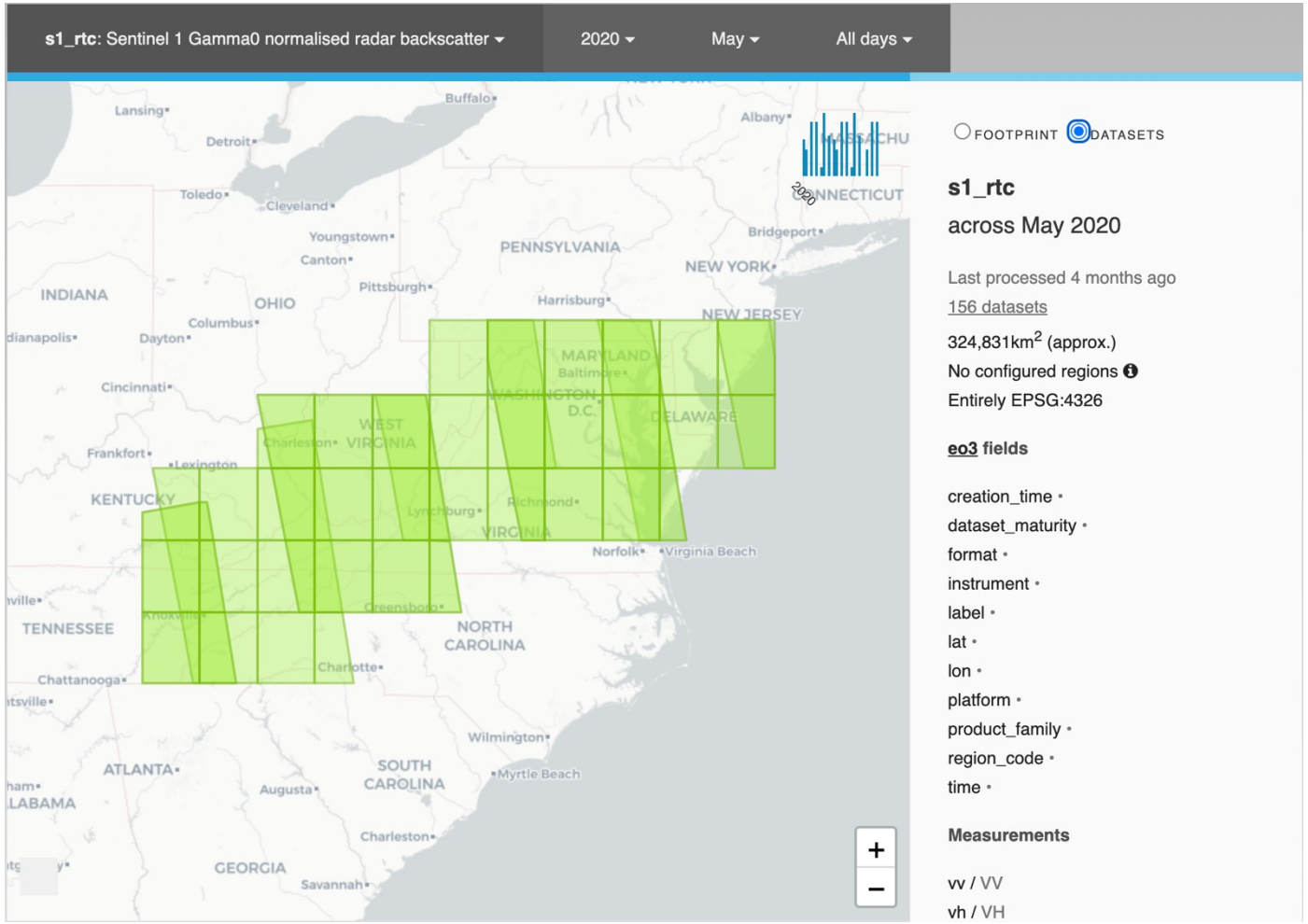


Evolution to Dashboards



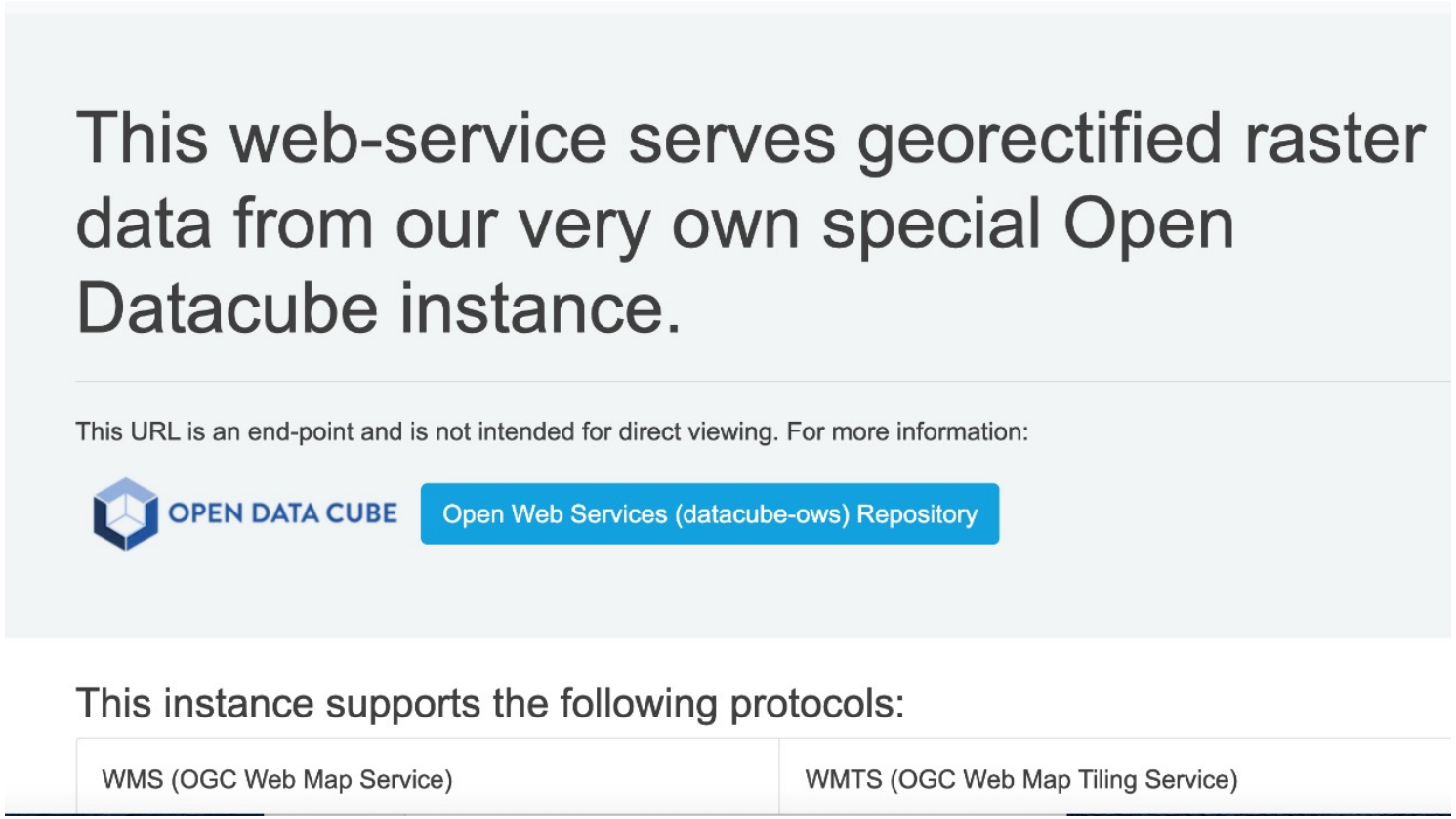
Explorer

- Browse available datasets
- View specific scene coverage
- Discover product names and measurement names
  - important for requesting data in datacube.load()



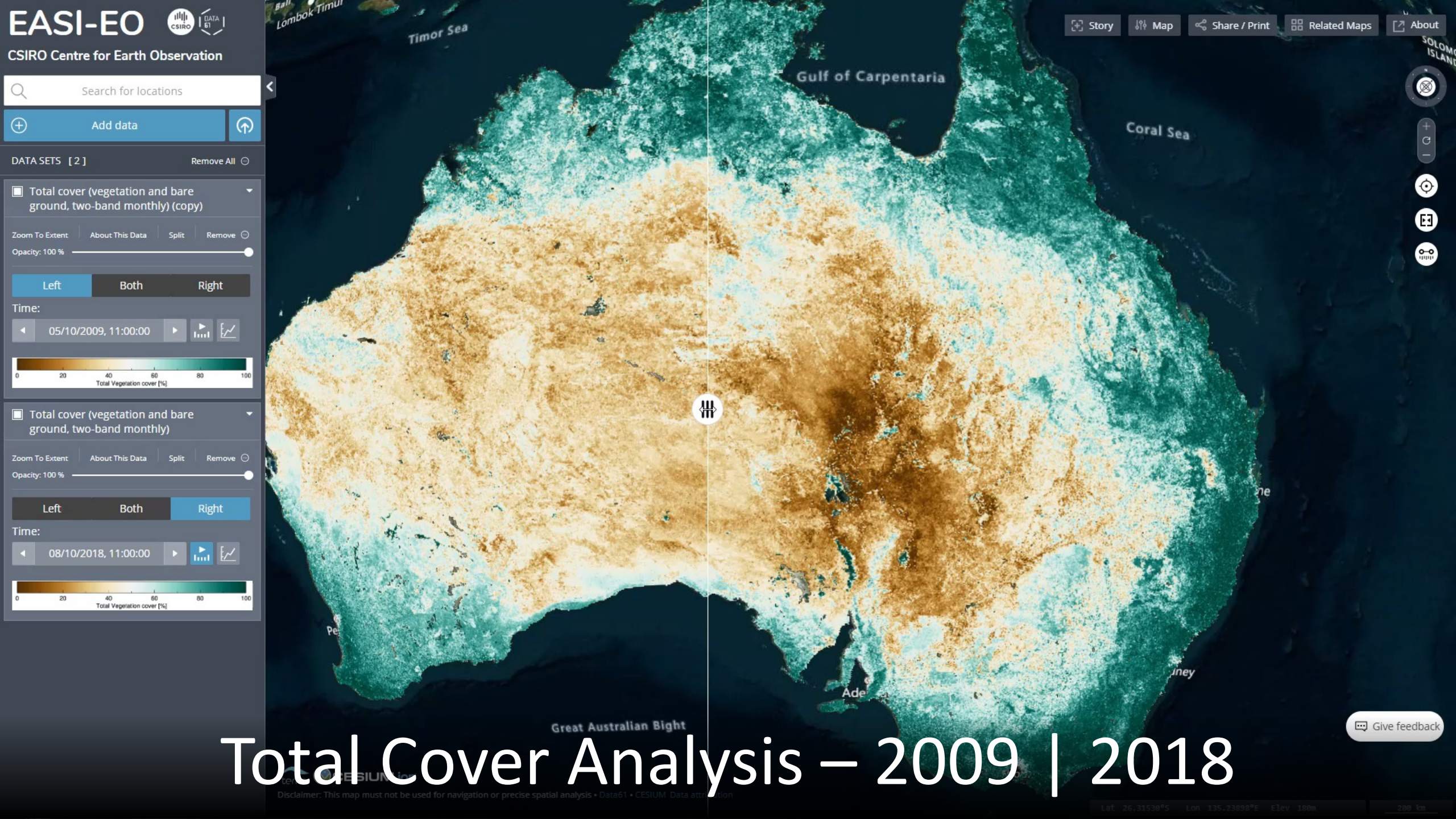
Web Map Services for GIS Applications

- Provides an OGC interface to the ODC API
- All you need is the URL
- Can perform some on-the-fly product generation e.g. NDMI, Cloud masking



TerriaJS Web Map Application (Demonstrator)

- Uses the Web Map Services
- Can be extended for visualisation, analysis, etc



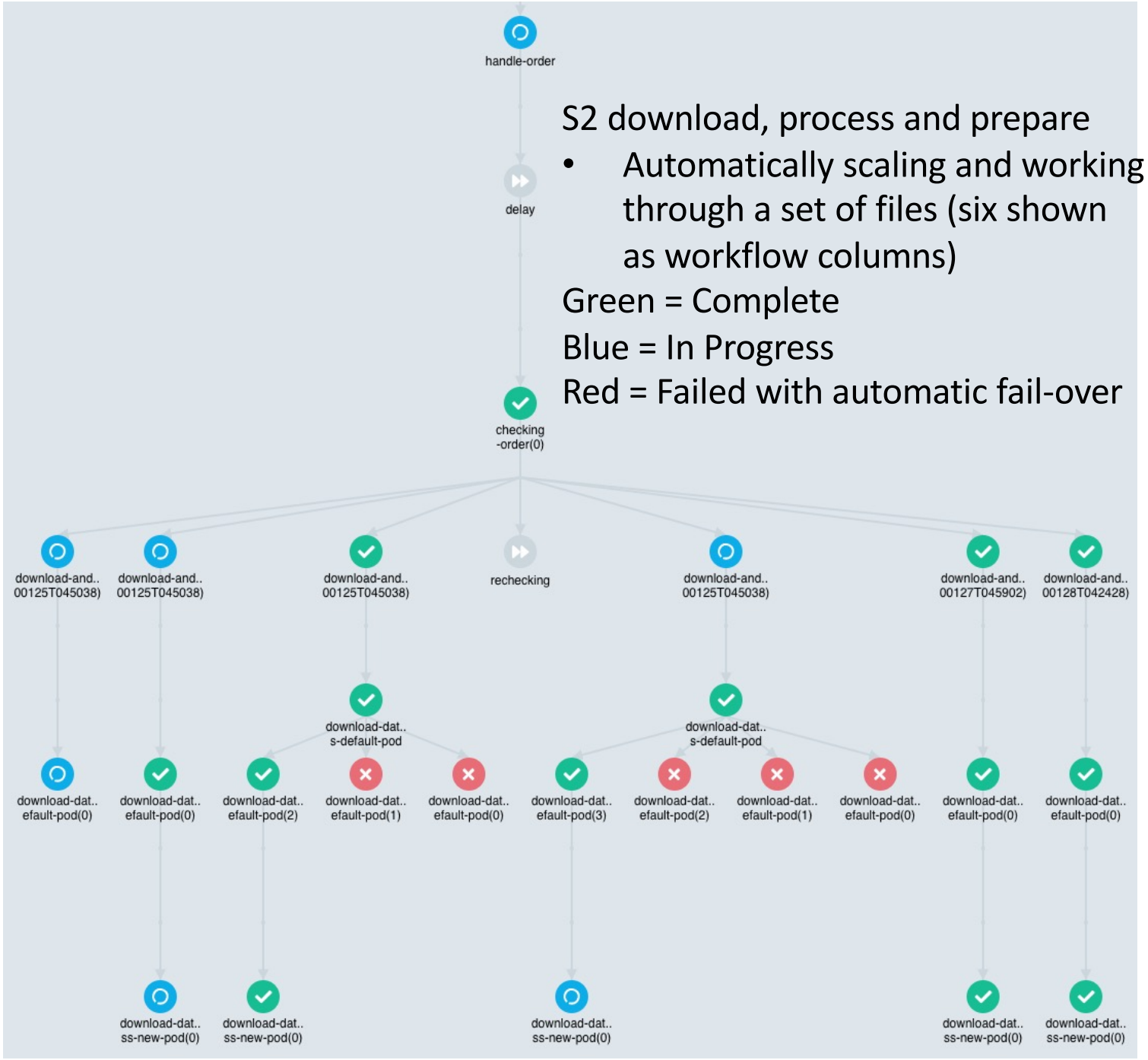
Total Cover Analysis – 2009 | 2018

Argo Workflows

- EASI's data-pipelines
- Custom production workflows

Features

- Scaling, Archive, Logging
- Web API to trigger workflows and monitor
- Single Sign On
- Resource tailoring for stages
- Resource Failover



S2 download, process and prepare

- Automatically scaling and working through a set of files (six shown as workflow columns)

Green = Complete

Blue = In Progress

Red = Failed with automatic fail-over

Committee on Earth Observation Satellites

CEOS Earth Analytics Interoperability Lab Update

Dr. Robert Woodcock, CSIRO, WGISS Chair

Dr. Brian Killough, NASA, CEOS Systems Engineering Office

Mr. Jonathan Hodge, CSIRO Chile, Program Director

WGISS 51 Virtual Meeting

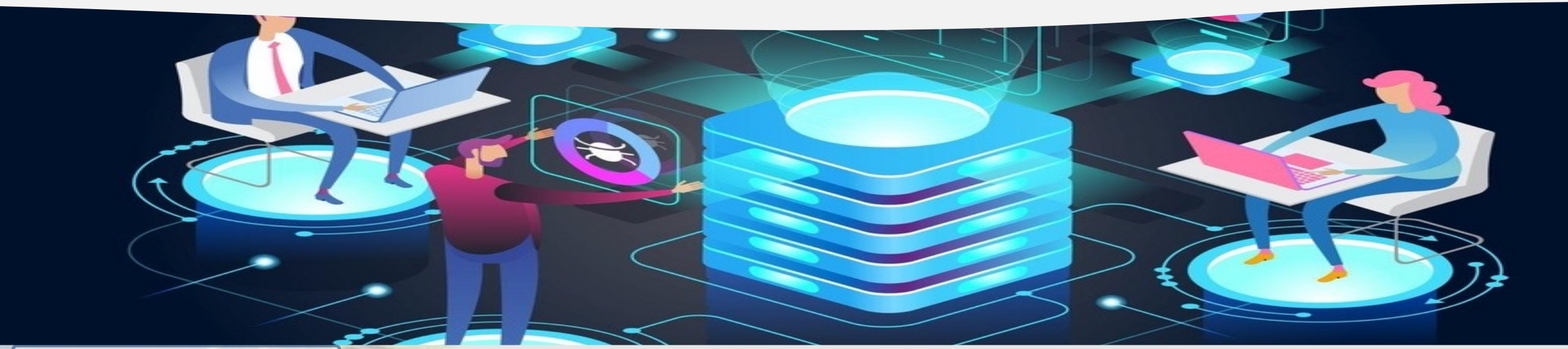
20-22 April 2021

Why do we need an Interoperability Lab?

- **Validating interoperability** between multiple CEOS organizations and working groups is complex
- **ARD and FDA are here** and there are WGISS technologies and Best Practices that need to change

The CEOS EAIL is intended to help CEOS WG & VCs explore these challenges *together*

EAIL is jointed operated by WGISS, CEOS SEO, CSIRO (Aus) and CSIRO Chile



Current Status
Engagement

- **15 registered users – more to come**
  - Large numbers is *not* the goal
- **CEOS engagement:**
  - WG Disasters Flood Pilot
  - CEOS COAST
  - DEMIX
  - Asia-RiCE
  - *In all cases EAIL is working with the groups on preparing data and regions of interest and assisting new users*
- **Open Earth Alliance community forum for EAIL**
  - *https://forum.symbios.space/c/ceos-eail/8*
    - *Very empty now but ready to use*

Current Status
Interoperability

- **Fully deployed and operational**
- **For training purposes US Southeast states and coastline have been indexed** (https://explorer.eail.easi-eo.solutions)**:**
  - USGS Collection 2 – full landsat archive, used-in-place, requester pays
  - Element 84 Sentinel 2 COGS, used-in-place, public
  - Sinergise S1 ARD service – ordered-and-stored, fee for service (but quick)
  - All of these are ARD (confirmed or under assessment for CARD4L)
- **CSIRO EASI Training notebooks**: customised for direct use in EAIL
- **Dask Gateway:**
  - The EAIL supports scalable computing
  - Customised based on project demand with larger nodes, GPU for ML, etc.