ISRO Data Cube Initiatives for Regional CB

By
Debajyoti Dhar
Group Director
Signal and Image Processing Group
Space Application Centre(ISRO)
Ahmedabad/INDIA

CEOS WGCapD-8 Annual Meeting
Agenda Item #39
Working Group on
Capacity Building & Data Democracy
Indian Institute of Remote Sensing
Indian Space Research Organisation
Dehradun, India
March 06th – 08th, 2019
Why EO Datacube?

EO data can be considered as Big data that have volume, velocity and pseudo-variety which makes it complicated to analyze and visualize.

Expertise, Infrastructure, or Network bandwidth requires to efficiently and effectively access, process, and utilize EO data.

**Data Processing**

Both Pre & Post-processing EO data introduces challenges when trying to integrate or analyze.

Challenges associated with serving Big EO data.
Why EO Datacube?

- Another challenge is working with data from multiple Missions
- Data processing servers keep processed count data, TOA and SR data from multiple missions archived at different locations.

The archive not accessible to application scientists due to many constraints -
- Lack of unified storage location
- Security (directly exposing the archive library is not feasible)
- Difficult to categorize products for private use and for public access kept at same location.

Inefficient use of resources leading to under-utilization of data
Vision for ISRO Datacube

**E-governance (Unified Framework for all applications ...)**
- Application of EO data for crop monitoring, urban and rural infrastructure planning, disaster management, weather forecasting, forest preservation, pollution awareness etc.

**Public Consumption of data for societal benefits**
- Data services easily available and accessible to users
- Reduce burden of data-handling and post-processing at user end.

**Growth of R&D on EO data applications**
- Datacube to serve as powerful tool for data interoperability for EO Missions, in tandem with advances in collaboration among ISRO and other agencies on Analysis Ready Data (ARD)

**Decentralizing by Regional Capacity Building**
- To provide software toolset for deploying local, regional or national time-series of multiple spatially aligned datasets as per user needs (based on region, time period, layers, grid projection)
**Technology development**

Data storage, middleware, APIs and UIs in web and mobile applications.

**User Engagement**

- Common understanding of the issues in different phases and mitigation plan.
- Focus on the user experience and feedback.

**Data processing and Re-formatting**

Standardization in Pre-processing, database definition, ARD generation, formatting, tiling, post-processing etc.

**Prototype Deployment**

Emphasis on deployment of Data Cubes to ISRO users as beta site and populate more ARDs into the framework.

---

**Data Preparation**

- Analysis Ready Data (ARD)
- Satellite products RS1/RS2/2A/OCM
- Bio-geophysical data products

**Core Technology**

- Backend Software (Pre-processing Library) for data ingestion and data loading
- API Interfaces
- Mosaic and composite software
- Visualization backend

**Analysis & Visualization**

- Historical Data Analysis
- Urban Planning
- Resource Monitoring and Prediction
- Change Detection
- Trend Analysis
- Visualization tools

**Prototypes**

- Web Data Analysis Portal
- Data Visualization
- Time-series visualization
- Jupyter based online framework
• APGI is a framework that helps automating EO data discovery and (pre-)processing using service chains for transforming observations into information products.

• Uses large storage capacities and high performance computers and interoperable standards to develop a scalable and efficient analysis system.
1. Independent components based Software Architecture
2. Multiple technology assimilation
3. Ingesting high volume, variety and velocity raster and corresponding metadata
4. Integration of multiple on-line and standalone applications for data analysis.
5. Data Visualization components (web-based)
6. Data Analytics components (Jupyter Hub based)
Architecture

Satellite data stores could be co-located with data cube infrastructure
30+ TB data and growing !!!!

**Present Data Stack**
OCM-2 -> RS-1/2/2A (AWiFS -> LISS-3 -> LISS-4) -> CARTO-2S

**Application Infrastructure**
(Python/Jupyter, over web browser)

**API Module (Python)**
Numpy, Xarray & Geopandas for Multi-Temporal Datasets (Latitude/Longitude/Time)

**DataCube “Ingester”**
Database Indexes

**Postgres Database**
XML/YAML/JSON Metadata

**Populating Metadata Database**

**Organization Files**

**Polls DB Data Layers**

**NAS Storage Server**
Pulls Pixel Layers or Pixel Time-Series Stacks
Serves pixels to Application via OGC complaint (WMS/WFS/WCS)

**Server**

**Application/User Interface**
Prototype Development

A small part of West India animated with multi-temporal band-8 of OCM-2

Vegetation fraction and land surface water monthly composite prod.

Horizontal time-series profile featuring pixel drilling operation over multi-temporal VF product stack.
Example Program Linkages

Prototype Deployment for Gujarat, Goa and neighboring Regions

Pre and post rainfall RS2A LISS-3 BAND-5 data featuring change in reservoir levels in Karnataka. Reference reservoir boundary shown in red color.
Visualization of Multi-resolution Image

IRS-R2A
AWIFS image (Resolution: 56.0 m)
Gridsize: 10x10 km

Gandhinagar (Gujarat)
Role of Agencies

• SAC and NRSC will spearhead the ISRO datacube initiative - SAC in software design, development and optimization role - NRSC as data enabler, data assimilation and finally Platform and Data as service Distributer.
• All ISRO and State centers to facilitate EO data enrichment by enabling collection and dissemination of field survey and ground truth data.
• Foster the ingestion of International EO data in tandem with ISROs EO data with collaborations from Other Space Agencies.
Thank you!