

OPEN DATA & OPEN KNOWLEDGE Workshop

❖ *Session 1: Open Space Based Data and Data Cubes: Kenya Space Agency efforts*
Open Data Open Knowledge Workshop

Geneva, Switzerland - 15-16 June 2023

Mr. Charles Mwangi, Acting Director Space Sector & Technology Development

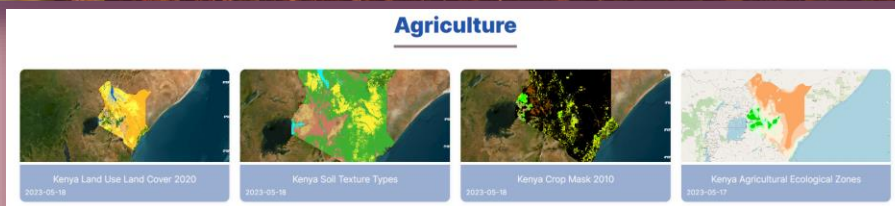
Gaps and Opportunities...

- ▣ **Gaps...**
 - **Hardware:** Compute, storage...
 - **Software:** Analysis, visualization...
 - **Data:** Geospatial, non-geospatial, machine readable...

- ▣ **Aspiration**
 - Leverage on **free** and **open source data** and **systems** to enable policy makers to make decision based on data...
 - Make **Analysis Ready Data** accessible and within reach...

Data Hub (Data Portal/Data Cube/Knowledge Hub)

3

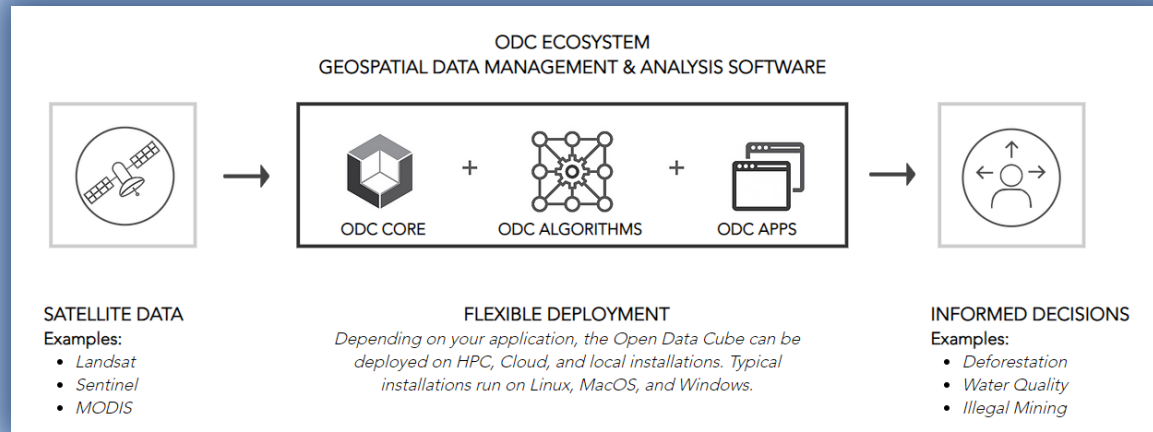


Why a Data Hub?

- ▣ Centralized data repository
 - Geospatial data
 - Knowledge Hub
- ▣ FAIR principles of data
 - Findability
 - Accessibility
 - Interoperability
 - Reusability
- ▣ Connected to a **Data cube**
 - Jupyter Notebooks
- ▣ To be linked to a **HPC**
 - Discussions in progress...

Open Data Cube

4



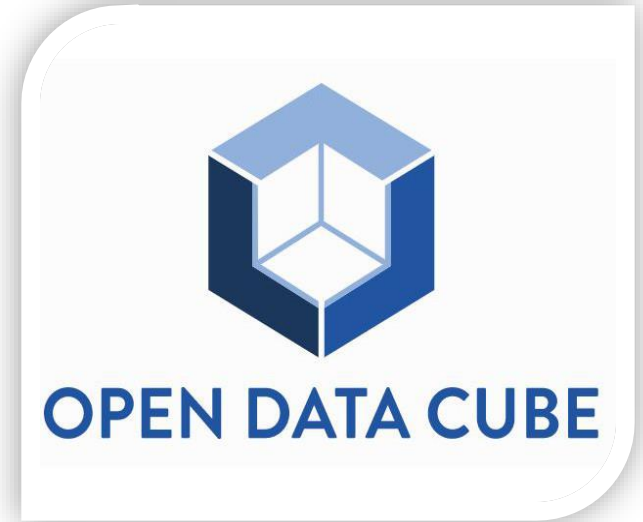
- ❑ Open Data Cube (ODC) is an **Open Source Geospatial Data Management and Analysis Software** project that helps harness the power of **Satellite data** by providing an **open and freely accessible exploitation architecture**.
- ❑ At its core, the ODC is a set of **Python libraries** and **PostgreSQL database** that helps you work with geospatial raster data.

Open Data Cube (ODC) Core

5

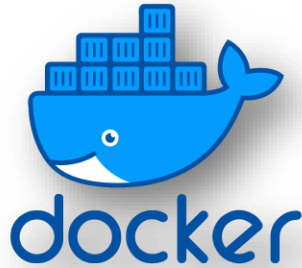
- Data cube core for the platform is installed using the python's package manager **Conda**.

- Modules/ packages installed
 - Open data cube core and command line tools
 - Libraries: **Scientific (Scipy)**, **numerical (numpy)** and **plotting (matplotlib)**
 - Geospatial Data Abstractions Library (GDAL)
 - eodatasets3 library

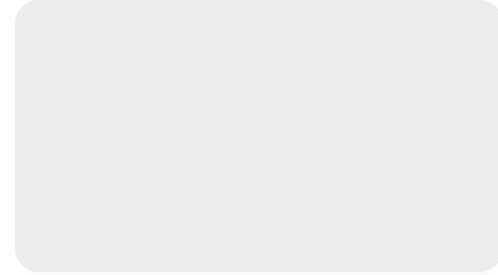


Docker & JupyterHub

6



- ❑ Docker platform installed on KSA server environment to enable hosting and deployment of software services as independent entities (containers)
- ❑ **KSA data hub** and **data cube jupyterhub** are deployed as docker containers.



- ❑ JupyterHub platform installed to enable the creation and management of users accounts accessing the data cube platform.
- ❑ JupyterHub platform creates separate jupyter notebook servers for each signed up user of the KSA datacube platform as docker containers.

PostgreSQL, STAC & EO3 Standard

7



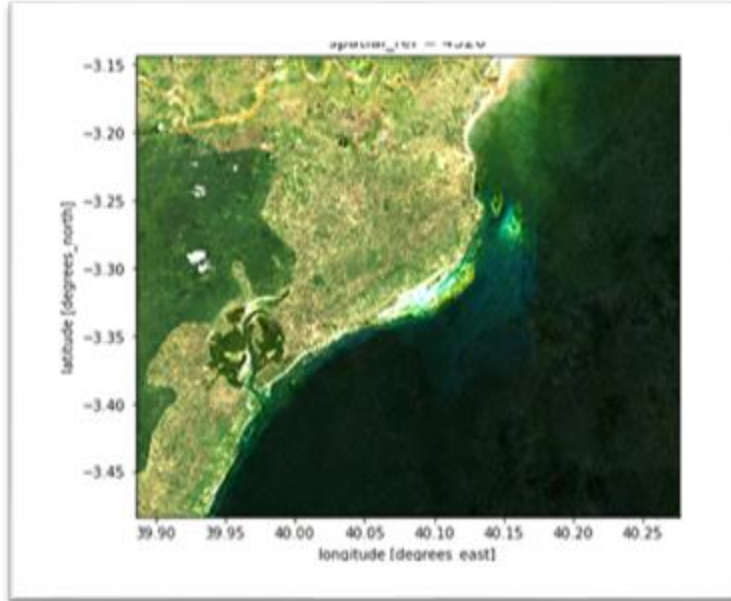
- PostgreSQL database was installed for storing details about the satellite imagery in the datacube
- Database deployed and managed as a docker container accessible by through users' jupyter notebook server containers



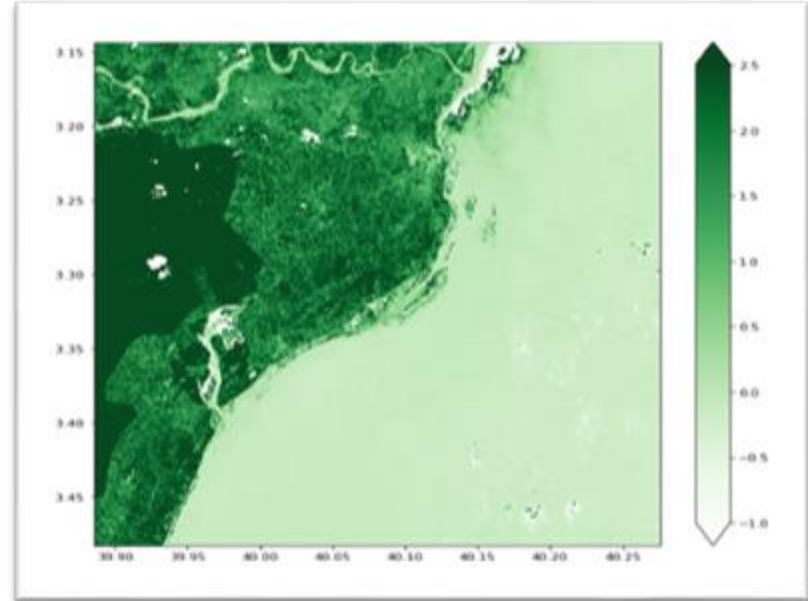
- Landsat and Sentinel 2 datasets indexed from AWS into open data cube instances using STAC API tools
- Satellite datasets for Kenya regions indexed
 - Sentinel data (2015 to 2022)
 - Landsat data (2000 to 2022)

- **EO3 standard** used to generate YML file for indexing GeoTIFFs into the open data cube instance

Sample ODC Products



■ **True Color image (RGB)**
 Landsat – Lamu, Kenya



■ **Enhanced Vegetation Index (EVI)**
 Landsat – Lamu, Kenya

Challenges...

9

- ▣ **Learning curve for ODC Deployment**
 - ❖ Need for dedicated effort by blended team (Geospatial/IT/Computer)
- ▣ **Loading Landsat data hosted on AWS on Jupyter Notebooks**
 - ❖ Challenge in loading AWS indexed Landsat into the jupyter notebooks
- ▣ **Deploying jupyterhub on the Kubernetes platform**
 - ❖ Deploying jupyterhub platform on the KSA kubernetes environment unsuccessful
- ▣ **Indexing Sentinel-2 imagery using EO3 standard**
 - ❖ Challenge in indexing Sentinel 2 imagery from the downloaded Geotiffs
- ▣ **Commercialization of the ODC**
 - ❖ Organizations providing ODC services with certain limitations and cost implications

Thank You



Email: charles.mwangi@ksa.go.ke
Website: www.ksa.go.ke