



SENTINEL Hub

One Data Format to Rule Them All

Grega Milcinski

Sinergise/Sentinel Hub

grega.milcinski@sinergise.com



**EUROPEAN EO COMPANY
OF THE YEAR 2018**

EARSC
European Association
of Remote Sensing
Companies

An aerial photograph of a coastline, showing a sandy beach and the ocean. A large, bold black question mark is centered over the image. Several text labels are placed around the question mark.

NetCDF

COG

HDF5

?

JP2000

ZARR

MRF



SENTINEL Hub

One Data ~~Format~~ Pattern to Rule Them All

Grega Milcinski

Sinergise/Sentinel Hub

grega.milcinski@sinergise.com



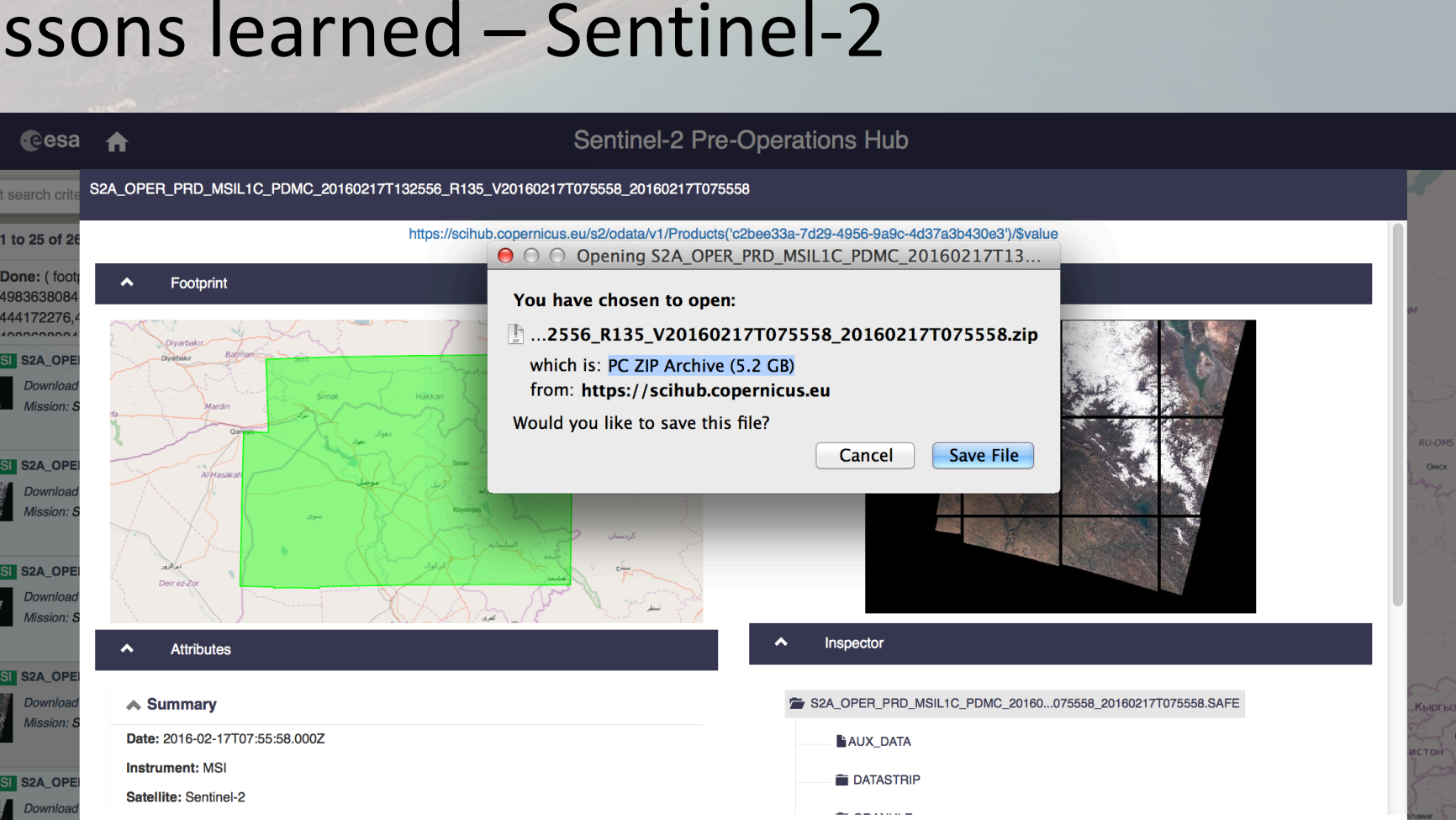
**EUROPEAN EO COMPANY
OF THE YEAR 2018**

EARSC
European Association
of Remote Sensing
Companies

Optimization factors

- open-source drivers (GDAL et al)
- cloud friendly
 - storage
 - speed
 - compute costs
 - access to subset

Lessons learned – Sentinel-2



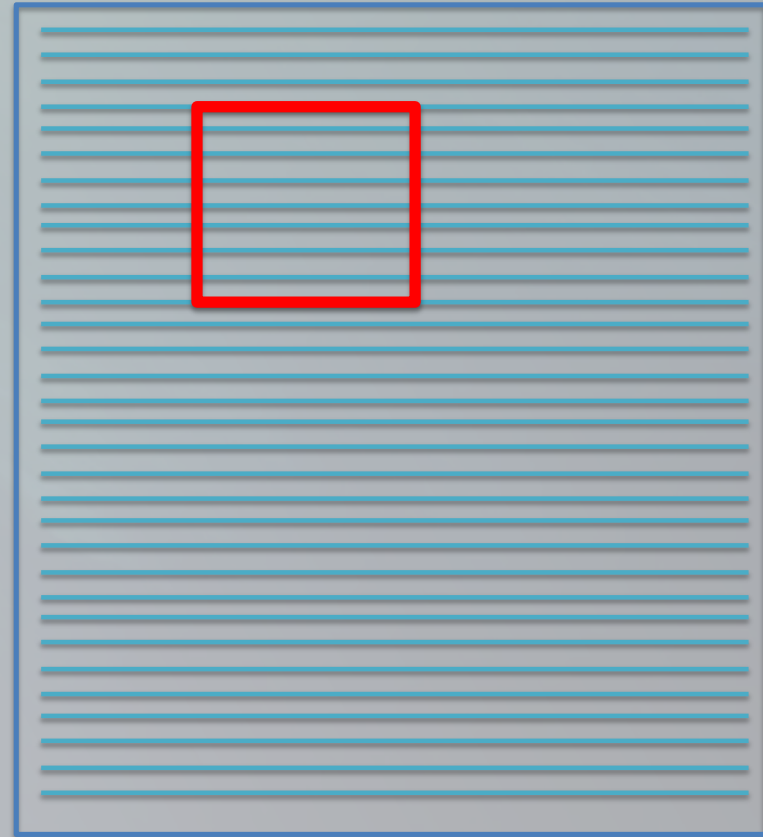
The screenshot displays the Sentinel-2 Pre-Operations Hub interface. A central dialog box prompts the user to save a file named "...2556_R135_V20160217T075558_20160217T075558.zip", which is a 5.2 GB PC ZIP Archive. The dialog includes "Cancel" and "Save File" buttons. The background interface shows a map of the Middle East with a green footprint overlay, a list of products on the left, and a summary/attributes panel at the bottom. The summary panel indicates the date is 2016-02-17T07:55:58.000Z, the instrument is MSI, and the satellite is Sentinel-2. The right panel shows the file structure, including AUX_DATA and DATASTRIP.

Lessons learned – Sentinel-2

- JP2000 vs. COG
- precinct (min. processing size) - 2.048x2.048px vs. 256x256 px
 - (same for tiles in COG)
 - <number of requests> vs. <data transfer cost/bandwidth/speed> vs. <decompression CPU> vs. <memory needed>
- raster meta-data in raster not in XML/GML
- appropriate bit depth (SCL, CLD, SNW in 16 bit)
- SCL in COG not JP2000
- processing baseline should change whenever processing process changes
- introduction of changes only when it makes sense (GRANULES->GRANULE) and with advance notice
 - Is regression testing too much to expect?

Lessons learned – Sentinel-1 GRD

- It is GeoTiff!
- No internal tiling
- No overview files
- Header position at EOF





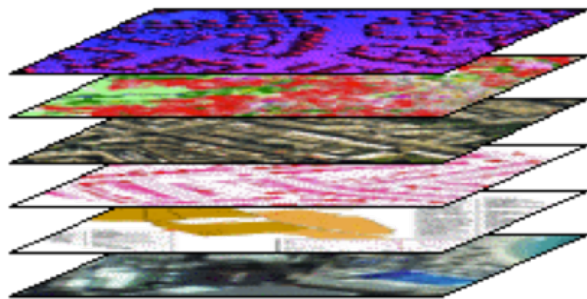
When ZIPs do make sense

- Replication of archives (1 week/month rolling archive)
- Sentinel-1 SLC et al

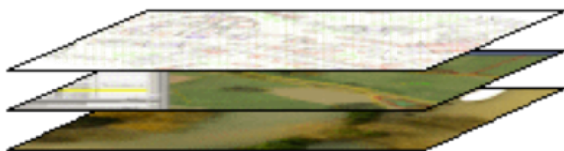
Analysis ready data

- Which projection to use?
- Which processing parameters to use?
- Which digital elevation model to use?
- Who will pay for processing and storage?
- When moving to cloud, only perform compute-intensive steps, which benefit vast majority of users and do not impact the quality

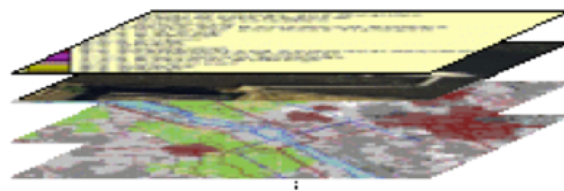
Sentinel-1, Sentinel-2 @ AWS EU-1



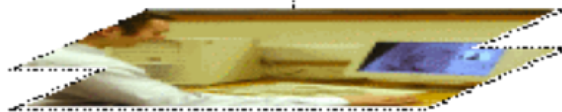
Landsat, MODIS @ AWS US-West



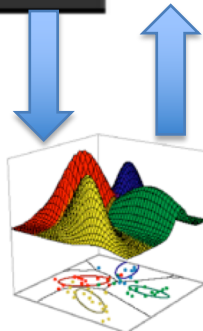
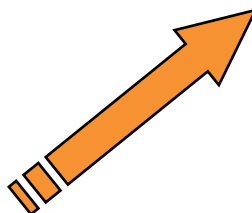
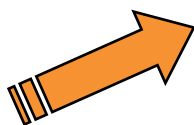
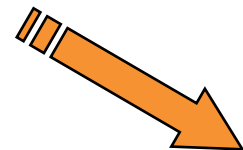
User data in COG, data on DIAS-es, ...



What the future brings?



SENTINEL Hub



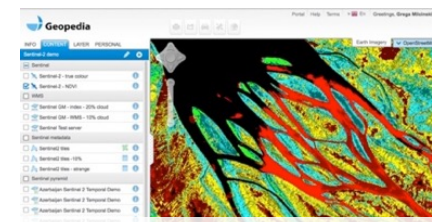
Machine learning

WMTS

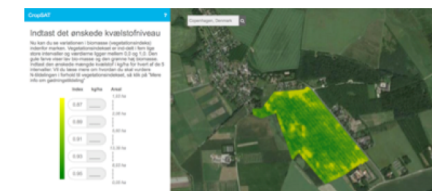
WMS

WCS

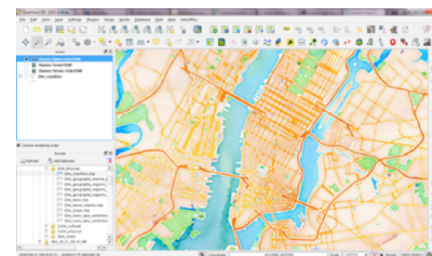
API



Cloud GIS



Web / Mobile apps



Desktop (QGIS,, ArcGIS...)

```

In [1]: path_out = "/workspace_out" if use_main_path else "/workspace_target"

# Subsetting coordinates
sys.path.append("../example_data")
from data_loader import load_img, load_mask
from data_loader import load_classes

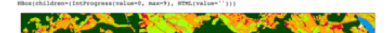
fig, axes = plt.subplots(figsize=(10,10),ncols=2,rows=1)

# Load data
for i in range(len(classes)):
    img = load_img(path_out, classes[i])
    mask = load_mask(path_out, classes[i])
    axes[i].imshow(img)
    axes[i].imshow(mask)
    axes[i].set_title(classes[i])
    axes[i].set_xlabel("X")
    axes[i].set_ylabel("Y")
    axes[i].axis('off')

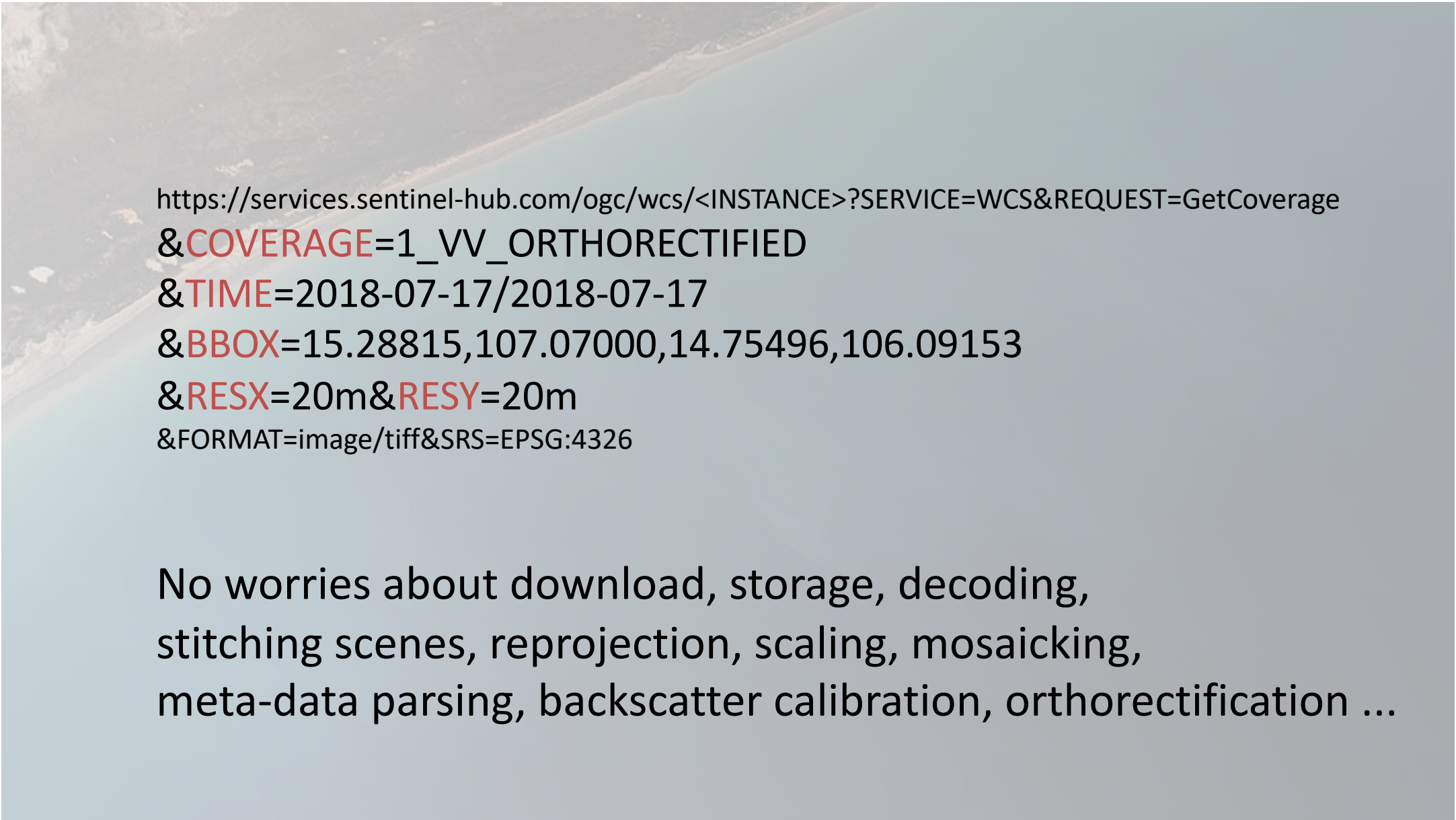
fig.savefig(path_out, format='png')

# Save data
for i in range(len(classes)):
    img = load_img(path_out, classes[i])
    mask = load_mask(path_out, classes[i])
    axes[i].imshow(img)
    axes[i].imshow(mask)
    axes[i].set_title(classes[i])
    axes[i].set_xlabel("X")
    axes[i].set_ylabel("Y")
    axes[i].axis('off')

fig.savefig(path_out, format='png')
    
```



Scripting (R, Python, ENVI...)

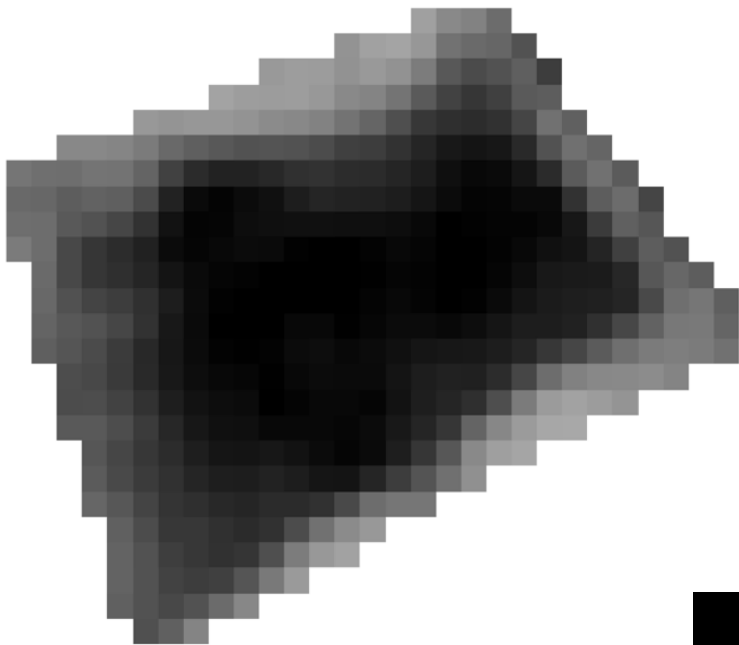


[https://services.sentinel-hub.com/ogc/wcs/<INSTANCE>?SERVICE=WCS&REQUEST=GetCoverage
&COVERAGE=1_VV_ORTHORECTIFIED
&TIME=2018-07-17/2018-07-17
&BBOX=15.28815,107.07000,14.75496,106.09153
&RESX=20m&RESY=20m
&FORMAT=image/tiff&SRS=EPSG:4326](https://services.sentinel-hub.com/ogc/wcs/<INSTANCE>?SERVICE=WCS&REQUEST=GetCoverage&COVERAGE=1_VV_ORTHORECTIFIED&TIME=2018-07-17/2018-07-17&BBOX=15.28815,107.07000,14.75496,106.09153&RESX=20m&RESY=20m&FORMAT=image/tiff&SRS=EPSG:4326)

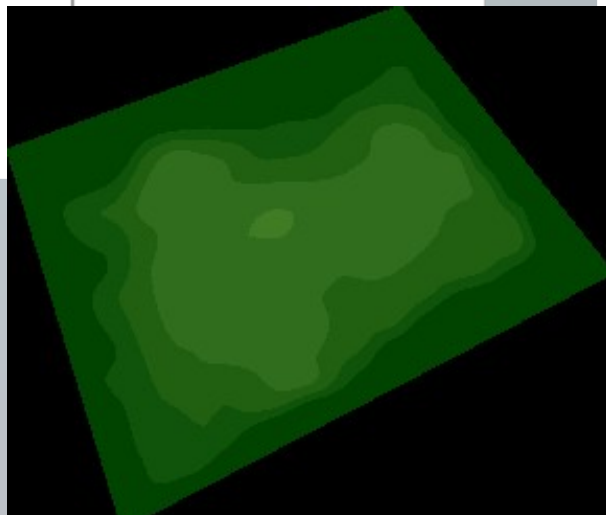
No worries about download, storage, decoding,
stitching scenes, reprojection, scaling, mosaicking,
meta-data parsing, backscatter calibration, orthorectification ...

Processing options

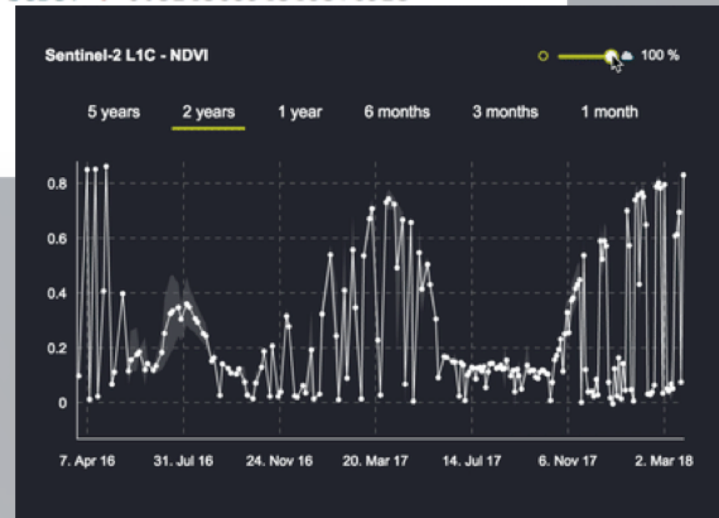
- Pre-defined product (true color, false color, etc.)
- Custom band combination
 `return [2.5*B08, 2.5*B04, 2.5*B03];`
- Various indices
 `var val = (B08-B04)/(B08+B04);`
 `return [val];`
- More complex algorithms
 - LAI, FAPAR – (basic) neural networks
 - <https://github.com/sentinel-hub/custom-scripts/blob/master/sentinel-2/lai/script.js>
- <https://github.com/sentinel-hub/custom-scripts/>
- Ortorectification



Identify Results	
Feature	Value
▼ 0	Sentinel-2 L1C fr...
▼ Sentinel-2 L1C from 201...	
Band 1	0.707848
► (Derived)	

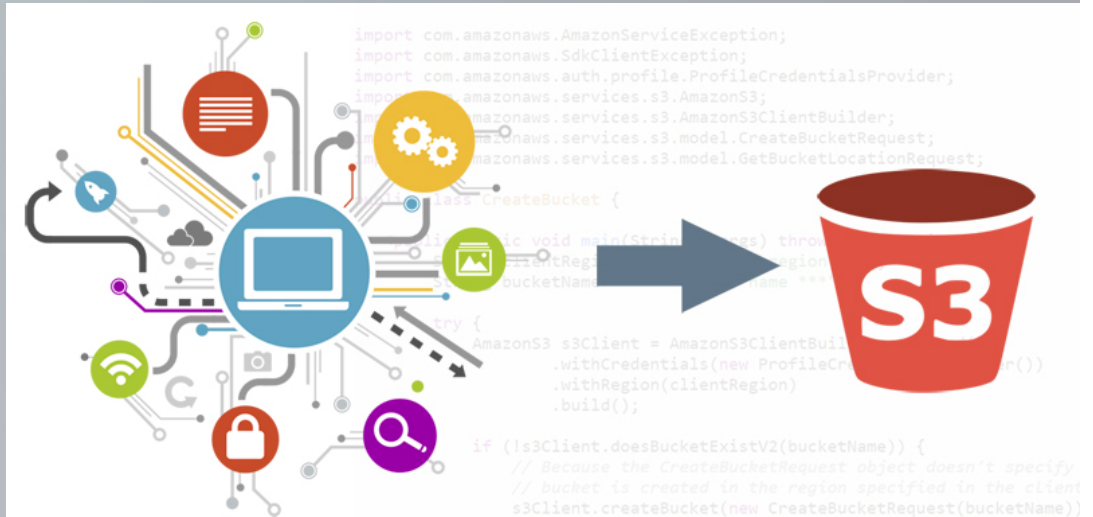


```
{
  "NDVI": [
    {
      "date": "2015-08-30",
      "basicStats": {
        "min": -0.5478424429893494,
        "max": 0.7815912365913391,
        "mean": 0.147320137875888,
        "stDev": 0.35443419609590726
      }
    },
    {
      "date": "2015-07-11",
      "basicStats": {
        "min": -0.5127978920936584,
        "max": 0.8115044236183167,
        "mean": 0.20168528533031557,
        "stDev": 0.31436594348376923
      }
    }
  ]
}
```



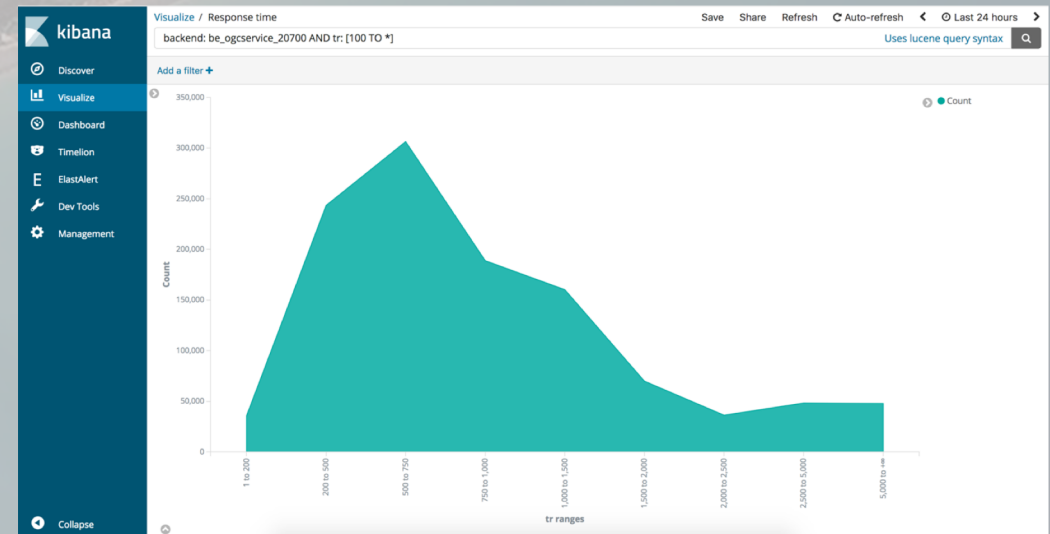
Bring your own data to Sentinel Hub

- Self-serve for raster data in COG (and ZARR in a few months)
 - Data residing on AWS, Mundi, CreoDIAS
 - Full Sentinel Hub functionality
 - Data remain under your control
-
- Satellite imagery
 - Derived products



For naysayers

- Yeah, but this is slow.



- Yeah, but this is expensive.

	GET STARTED	INDIVIDUAL NON-COMMERCIAL USE	INDIVIDUAL COMMERCIAL USE	APP DEVELOPERS & ENTERPRISE
Price	0 €	13.59 € / month (billed as 163.11 € / year, + VAT)	83.25 € / month (billed as 999.00 € / year, + VAT)	from 500 € / month (Click for details)
APP DEVELOPERS & ENTERPRISE				
Enterprise package - basic	500 € / month (5.000 € / year, 17 % discount)		200.000 processing units / month	
Enterprise package - enlarged	1.000 € / month (10.000 € / year, 17 % discount)		500.000 processing units / month	
Reserved capacity package 10	2.000 € / month		10 proc. unit / second (5 million processing units per month, unspent carried on up to 50 million)	
Reserved capacity package 20	3.000 € / month		20 proc. unit / second (10 million processing units per month, unspent carried on up to 100 million)	

Wishlist for Santa

- USGS to move Landsat 5 to AWS
- Move entire production to the cloud and provide access to intermediate steps
- Status dashboard
- Handle the process as an operational, not a research one

More info

- <http://sentinel-hub.com/>
- <http://apps.sentinel-hub.com/eo-browser/>
- <https://sentinel-hub.github.io/custom-scripts/>
- <https://github.com/sentinel-hub/eo-learn>
- <https://education.sentinel-hub.com>

Thanks





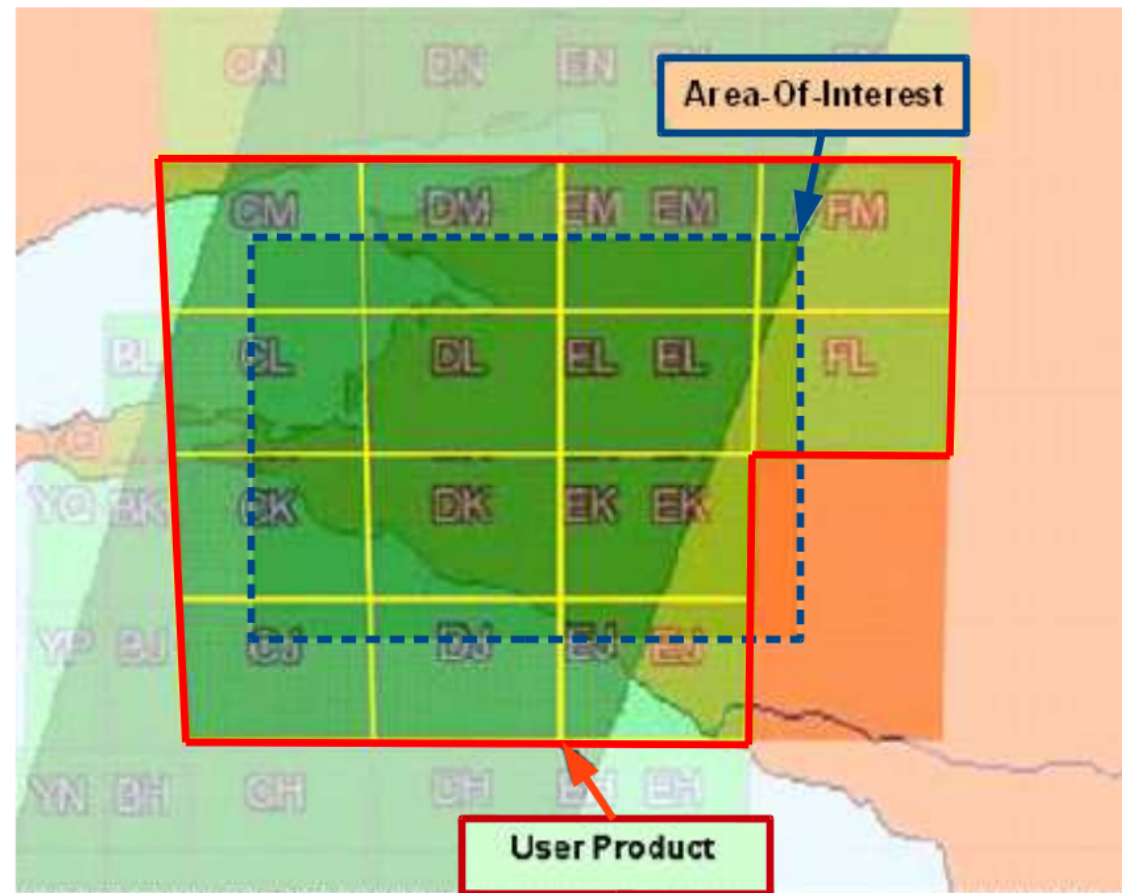
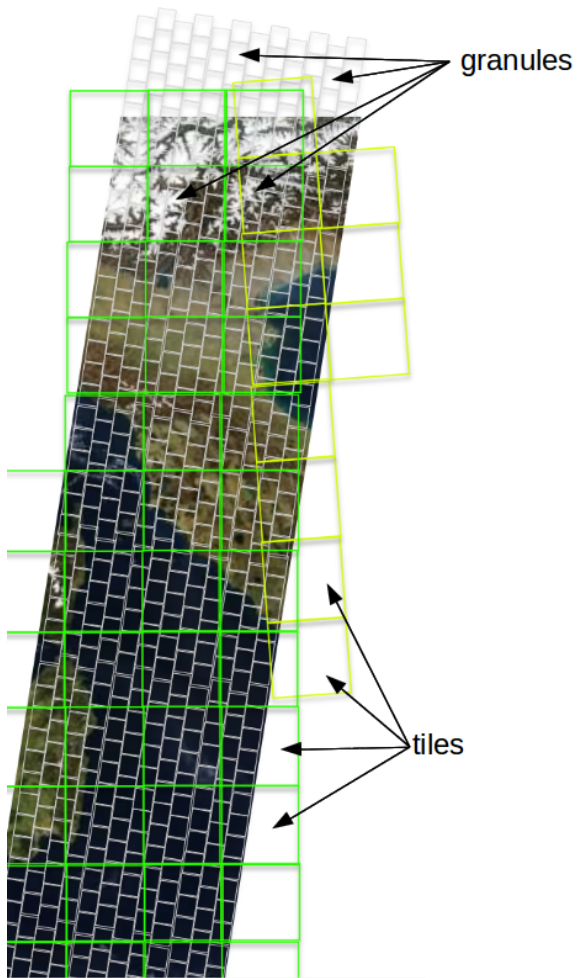
An aerial photograph of a coastline, showing a sandy beach and the ocean. The image is overlaid with a semi-transparent blue gradient that fades from the top left towards the bottom right.

A few years ago...

ENVI, ArcGIS and alike can open any file.

Climate scientists play with NetCDF

Lessons learned – Sentinel-2



Source: ESA, EOX

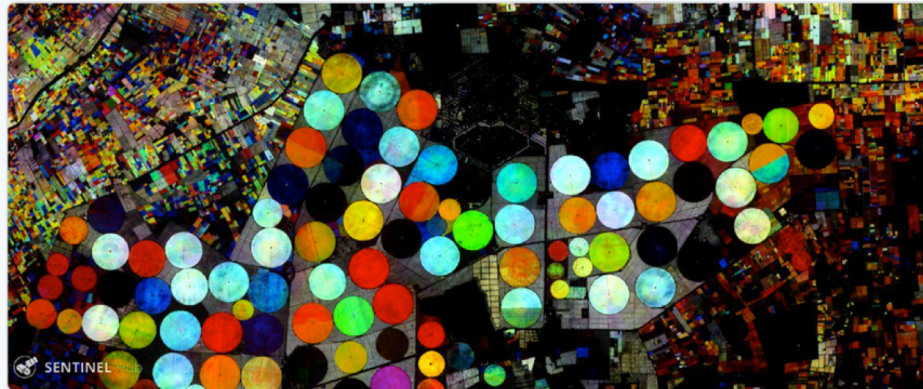
Lessons learned – Sentinel-2

- JP2000 vs. COG
- consistency L1C and L2A
- precinct (min. processing size) - 2.048x2.048px vs. 256x256 px
 - (same for tiles in COG)
 - <number of requests> vs. <data transfer cost/bandwidth/speed> vs. <decompression CPU> vs. <memory needed>
- raster meta-data in raster not in XML/GML
- appropriate bit depth (SCL, CLD, SNW in 16 bit)
- SCL in COG not JP2000
- processing baseline should change whenever processing process changes
- introduction of changes only when it makes sense (GRANULES->GRANULE) and with advance notice
 - Is regression testing too much to expect?



HD @HarelDan · 24 Oct 2017

Tip: Blue field growing, Green fields maturing, Yellow Fields ripe, Red fields reaped/drying. Same place, 3 days ago apps.sentinel-hub.com/sentinel-playg...



4 5 20



Stef Lhermitte

@StefLhermitte

Following

Replying to @HarelDan @sentinel_hub and 4 others

Wow! The moment even my mom can classify petabytes in seconds on her very old computer is getting closer. Just need to teach her Javascript

8:14 AM - 25 Oct 2017

1 Retweet 9 Likes

