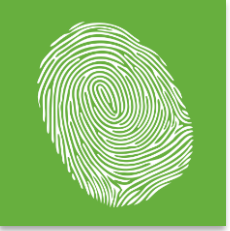

vito
remote sensing

# Copernicus Global Land Mapping, from private to public cloud

Bruno Smets, Marcel Buchhorn, Dirk Daems
bruno.smets@vito.be



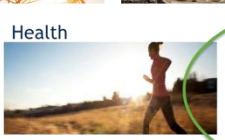
# Outline

- Who are we
- VITO's mandate in Copernicus
- Global Land Cover mapping in private cloud
- Scaling Up to public cloud
- Conclusions

SEE THE BIGGER PICTURE

# VITO Remote Sensing



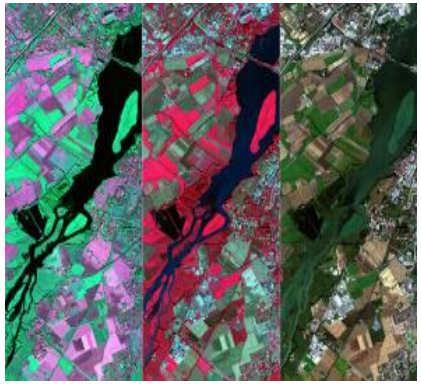


## Our MISSION

We make you **see the bigger picture**.
We take you **high-tech with low risk**.
We leverage your **impact on sustainability**.

## Our VISION

We go **end-to-end to give you the insights you need**. Without the hassle.
We do your **science & research** to stripe off complexity and risk.
We deliver top-notch **operational services and customer oriented solutions**.



vito
remote sensing

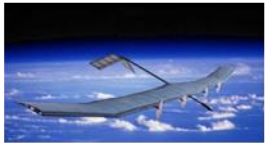
# VITO remote sensing

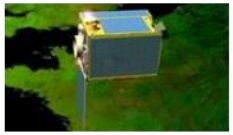
~90 FTE
~16 M€/yr

**Platforms**

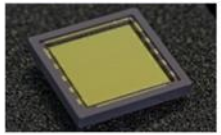

UAV


AIRBORN


HALE UAV


SATELLITE

**Sensors**



**Value Added Serv & Information Products**

**Markets**


Agriculture


Landuse & Biodiversity


Climate


Water & Coast


Infrastructure


Security

vito remote sensing

remotesensing.vito.be

# VITO Remote Sensing Data Center

@2018/06

## Tiered storage capacity:
- ~7 PB Disk storage
- ~5 PB Tape archive

## Server capacity:
- ~650 physical servers (2600 core cluster)
- ~300 virtual servers

## Networking capacity:
- Géant connection of 10 GiB

## Technologies used

KVM, CentOS, NetApp, puppet, ICINGA, openstack, ceph, jupyter, docker, vmware, Hewlett Packard Enterprise, Spark, hadoop, Hortonworks, GeoServer, FOREMAN, Check Point

## Customers

JRC EUROPEAN COMMISSION, belspo, United Nations, esa, Food and Agriculture Organization of the United Nations, Agiv

vito remote sensing

remotesensing.vito.be

# VITO MANDATE IN COPERNICUS

# VITO in Copernicus


- Lead Global Land Service Operations (since 2013) (http://land.copernicus.eu/global )

- Lead Climate (C3S) Land Biosphere ECV (since 2017)

- Member of Copernicus Academy
- Member of Copernicus Relays network
- Collaborative Ground Segment for Sentinels (http://www.terrascope.be )

TERRA SCOPE

# CGLOPS Product Portfolio

## VEGETATION

- Leaf Area Index (LAI)
- Fraction of Absorbed Photosynthetically Active Radiation (FAPAR)
- Fraction of vegetation cover (FCOVER)
- Normalized Difference Vegetation Index (NDVI)
- Vegetation Condition Index
- Vegetation Productivity Index
- Dry Matter Productivity
- Burnt Area
- Greenness Evolution Index
- Phenology metrics
- Moderate Yearly Land Cover

## ENERGY

- Top-of-Canopy reflectance
- Surface Albedo
- Land Surface Temperature
- Radiation Fluxes
- Evapotranspiration
- Active Fires
- Surface soil moisture
- Soil Water Index

## WATER

- Water Bodies
- Coastal Erosion
- Lake surface water temperature
- Lake and river water level*
- Lake surface reflectance
- Lake turbidity
- Lake trophic state

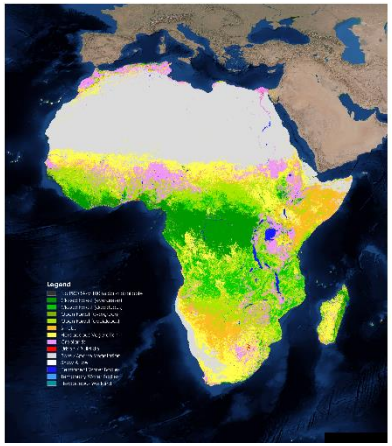
* non-gridded product

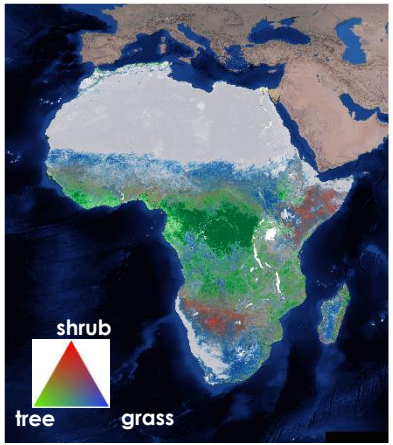
## CRYOSPHERE

- Lake Ice Extent
- Snow cover extent
- Snow water equivalent

# Copernicus Dynamic Global Land Cover Map

Continuous Covers (0-100%) Discrete Map (18 class)






» Copernicus The European Earth Observation Programme

## *Dynamic Global Land Cover Map*

A systematic SERVICE providing a DYNAMIC, YEARLY,
USER- ORIENTED at GLOBAL scale
@ 100m resolution from 2015 onwards

*Complementary to Pan-European, with less thematic details*

» ***Available collections***

| Algorithm version | Spatial | Temporal | Sensor |
|---|---|---|---|
| 1 | Africa | 2015 | PROBA-V |
| 2* | Global | 2015 - present | PROBA-V, Sentinel-2 |

(*) under preparation, expected release 2019

» ***Contributors***

- » **Marcel BUCHHORN**, Luc BERTELS, Ruben VAN DE KERCHOVE, Bruno SMETS
- » Martin HEROLD, Nandika TSENDBAZAR, Jan VERBESSELT, Dainius MASILIUNAS
- » Steffen FRITZ, Myroslava LESIV, Martina DUERAUER

vito remote sensing

WAGENINGEN UNIVERSITY & RESEARCH

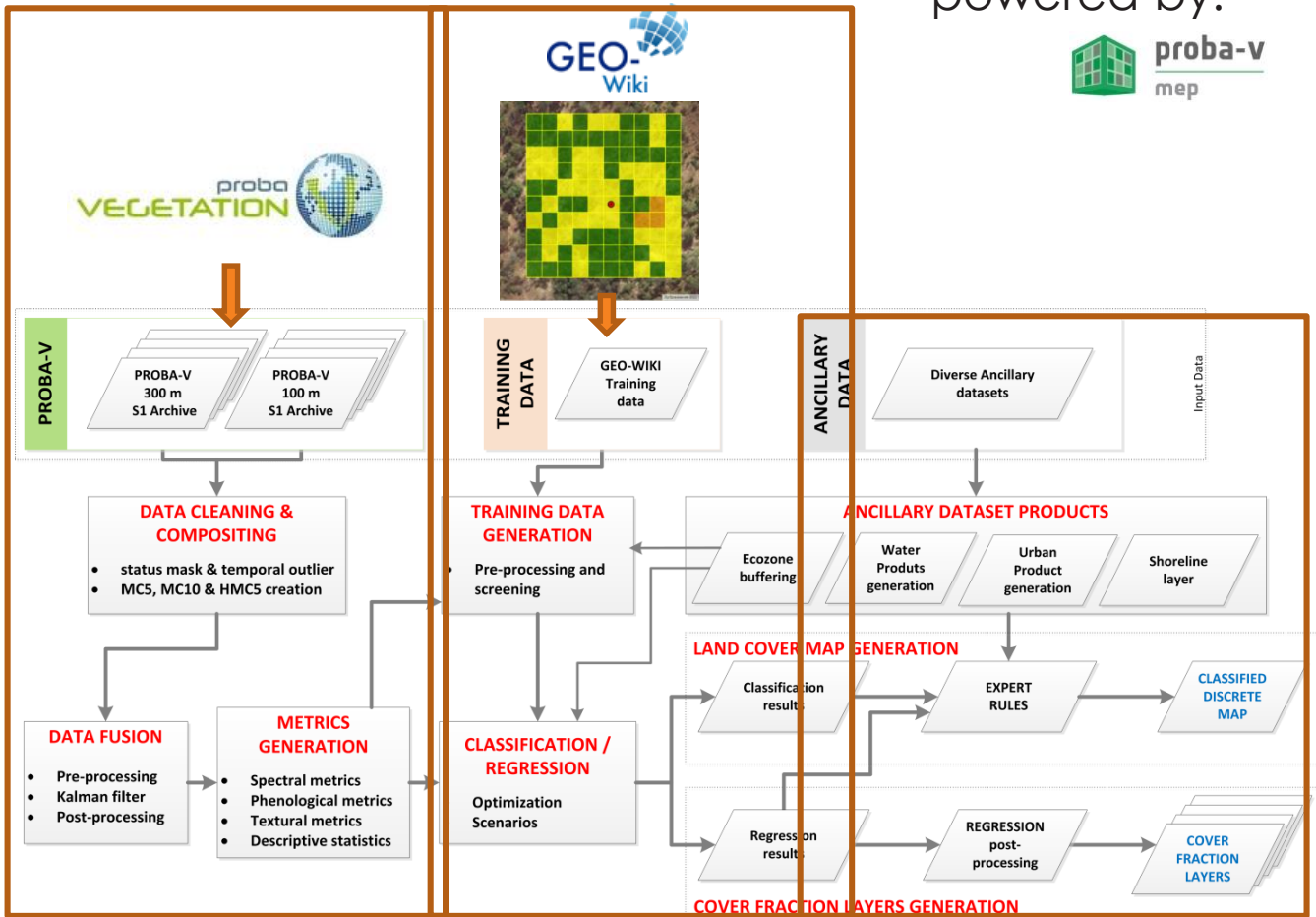
IIASA

# Copernicus LC100 v1

powered by:

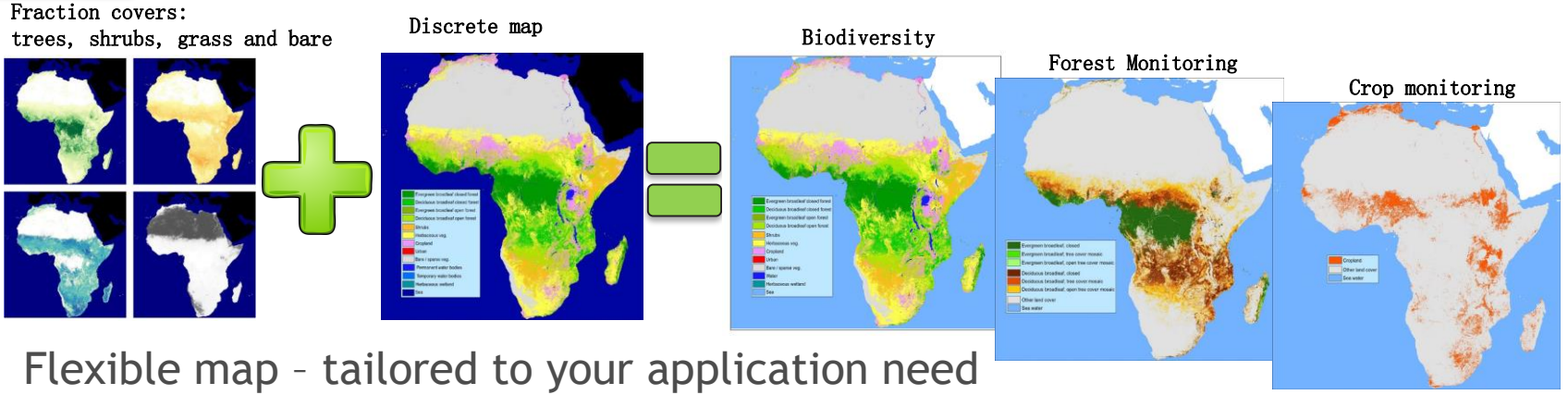
proba-v mep

- Pre-processing
  - Data cleaning & fusion
  - Generation of TS metrics
- Classifier/Regressor
  - RF optimized, 5 folded CV
  - 400 metrics, best band selection per ‘eco’-zone
  - 25K training points
  - 5 results + 4 fraction maps
- Expert Rules
  - Decision tree
  - Incorporate areas of agreement from GLC maps
  - Imprint JRC GSW/GHSL, DLR GSL (GUF+)




vito remote sensing

# Copernicus LC100 v1




- Flexible map – tailored to your application need
  https://blog.vito.be/remotesensing/lcm-tailored-to-your-needs
  - No ‘fixed’ mixed classes
  - Continuous cover fractions
- *CEOS-LPV ‘independent’ validated, incl. spatial distributions*
  - *Overall accuracy*: 72.3% +/- 1.8% to 80.4% +/- 1.1%
  - *MAE and RMSE*: 6% to 14% on cover fractions

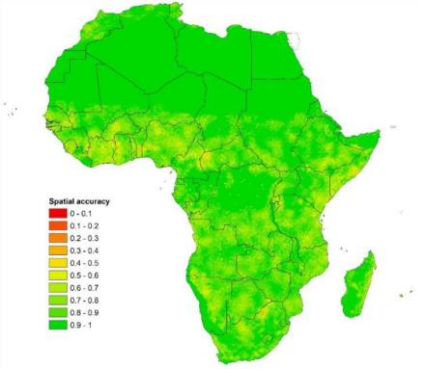



*Figure courtesy of N. Tsendbazar*

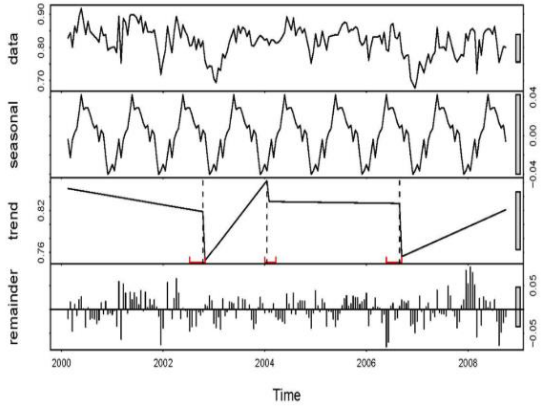
vito remote sensing

# COPERNICUS LAND COVER MAPPING IN PRIVATE CLOUD

# Copernicus LC100

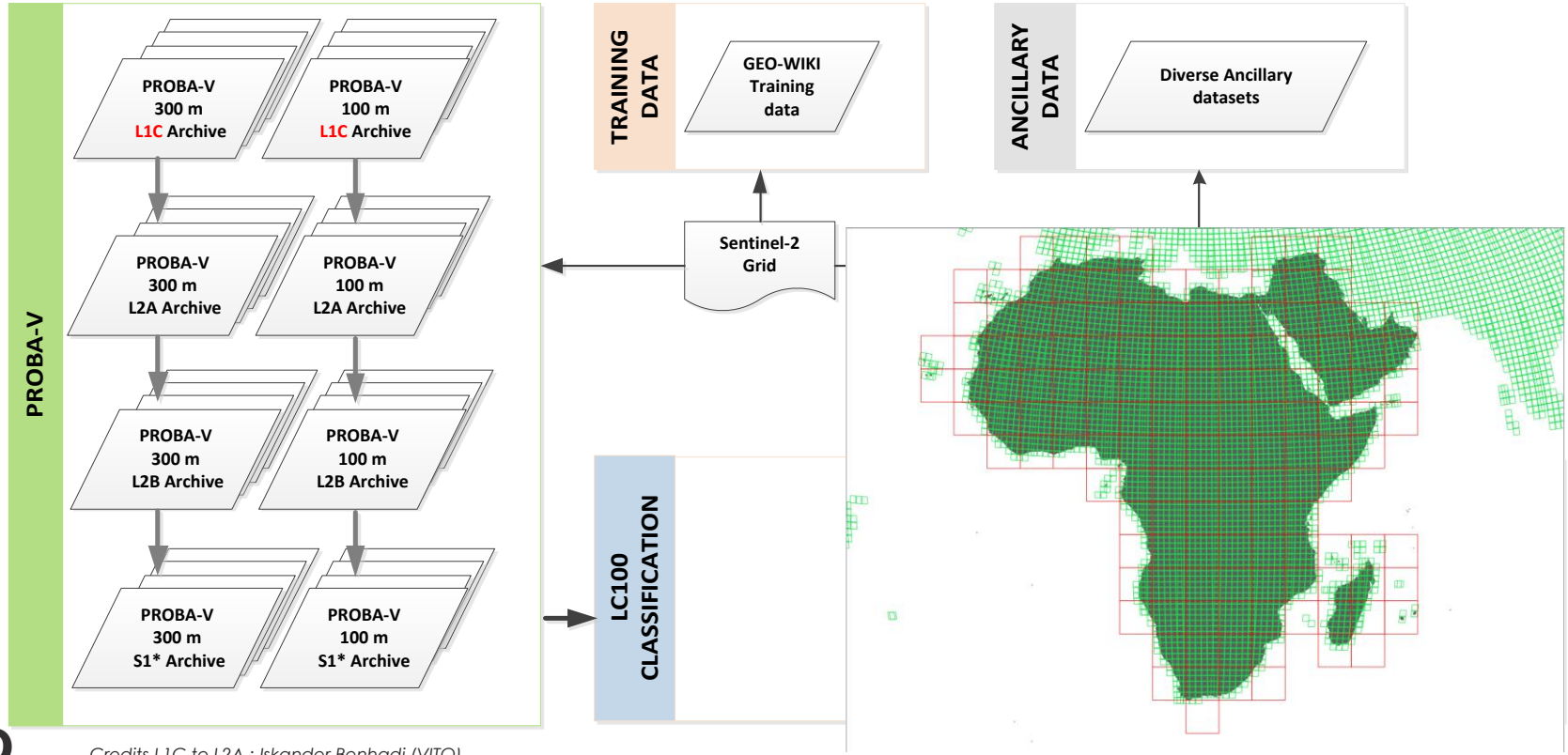
- For V1 of the *Africa* product
  - Fully automated workflow, with ~20K training points over Africa continent

- For V2 of the *Global* product -> less geometric distortion & better interoperability – Landsat and Sentinel-2
  - Generation of PROBA-V UTM 100m ARD
  - Training data and validation data was also switched to UTM (120K human qualitative training points with 10x10 box of 10m resolution)
- For V2 we integrate *change detection*
  - BFAST
    method for detecting and characterizing change within time series
  - NAUC
    indicator on stability of pixels




vito
remote sensing

# PROBA-V UTM ARD for LC100 V2




*Credits L1C to L2A : Iskander Benhadj (VITO)*

vito remote sensing

remotesensing.vito.be

# PROBA-V to UTM

- geometric correction including pixel alignment to the Sentinel-2 grid
- cloud and cloud shadow detection PLUS snow/ice detection (using additional ancillary data) – NEW: ***temporal cloud detection***
- Improved atmospheric correction using CAMS NRT data for AOD550, O3, and TCWV
- image compositing to generate time series stacks plus **image clipping** to the Sentinel-2 tiling grid (resolving issues in overlapping areas)
- generation of final status masks – improved to usable data mask
- **gap-filling** in the time series of a pixel via a Kalman-filter approach which uses the PROBA-V 300m corresponding side cameras information, compositing to 5-daily observations

proba-v
mep

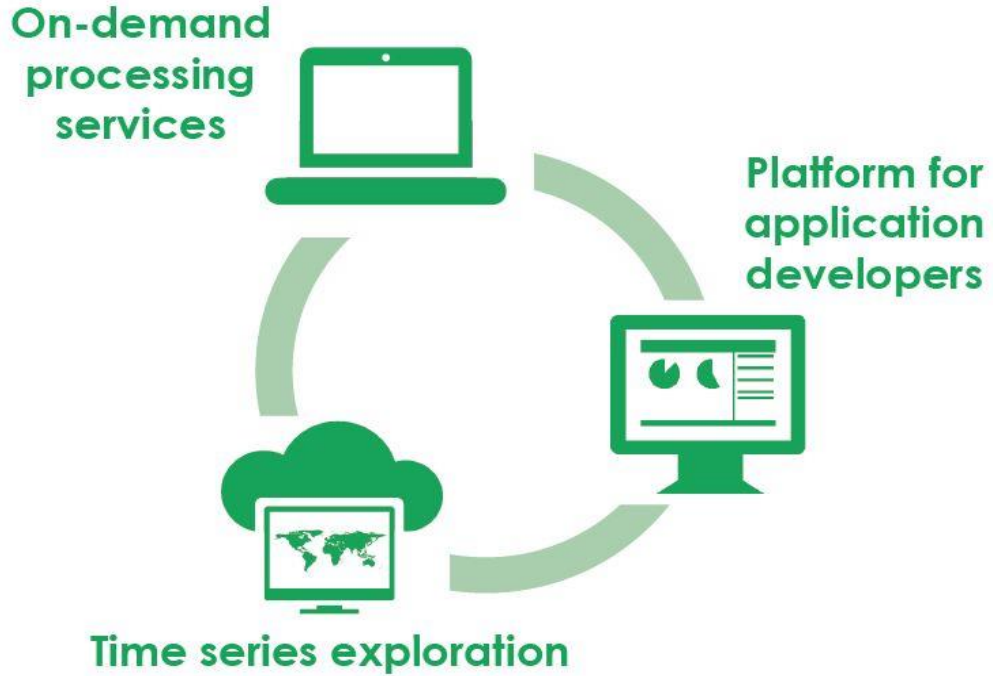
# *A node in a federation of platforms*

VEGETATION

proba VEGETATION

Copernicus Global Land Service
*Providing bio-geophysical products of global land surface*

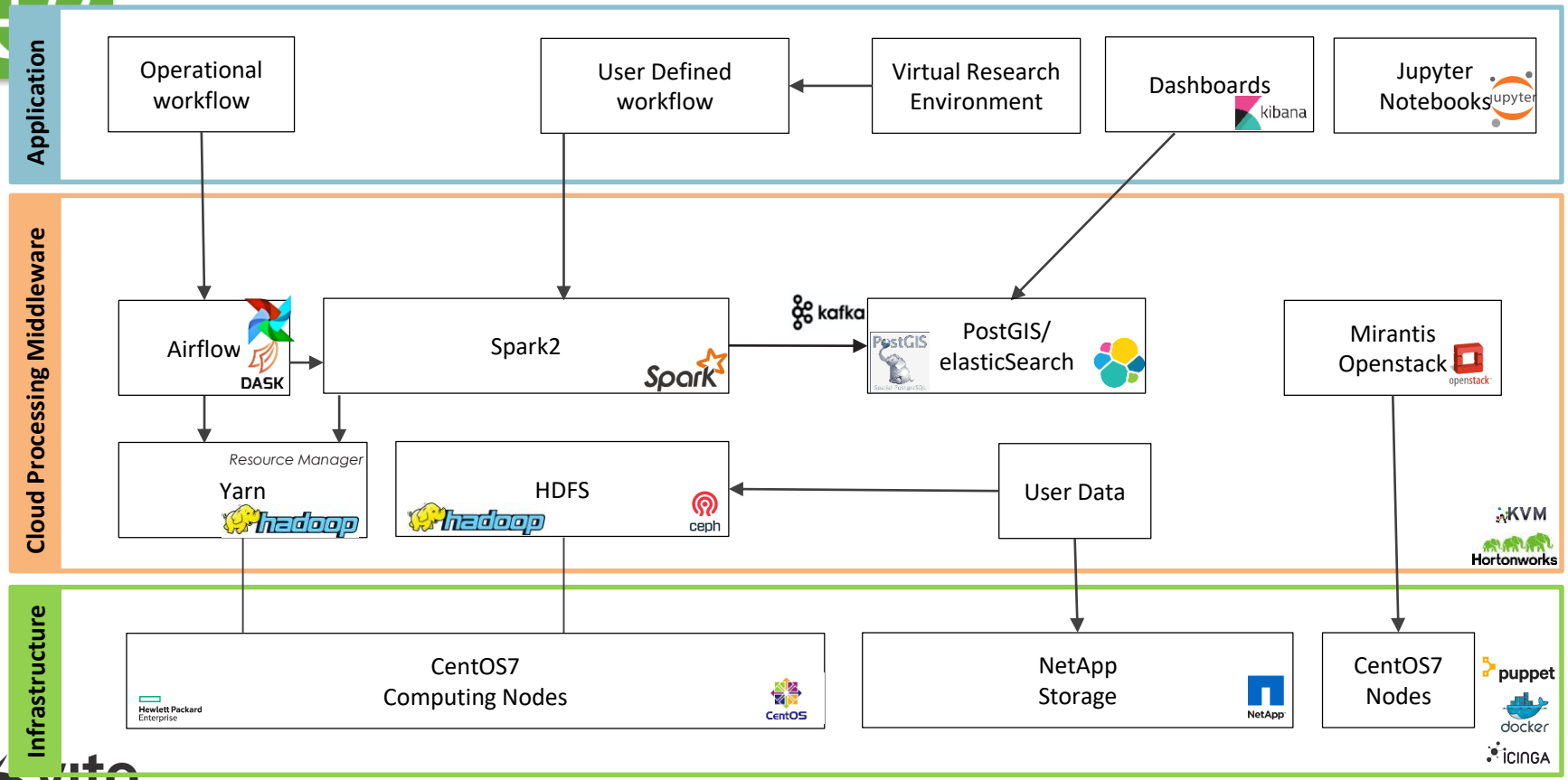
SENTINEL 1
SENTINEL 2
SENTINEL 3
TERRASCOPE




Develop → Operate → Support users

vito
remote sensing

# MEP Software Stack



# PROBA-V to UTM

- Over **50 million** single acquisitions were processed to generate a global PROBA-V ARD at 100m resolution spanning over 5 years … in 2 weeks using less 30% of resources
- Fully **aligned** to the **Sentinel-2** tiling grid and naming

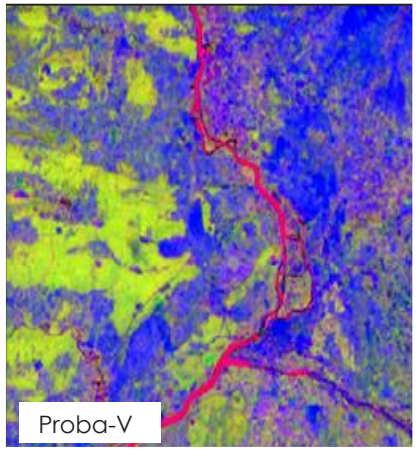
- Additional metadata are injected to enable a SpatioTemporal Asset Catalog (STAC), based on an ElasticSearch database to allow easy search through json
- Geotrellis based backend, allowing on the fly extraction of time-series cubes for specified areas or whole tiles – used in OpenEO

# LC100 Classification

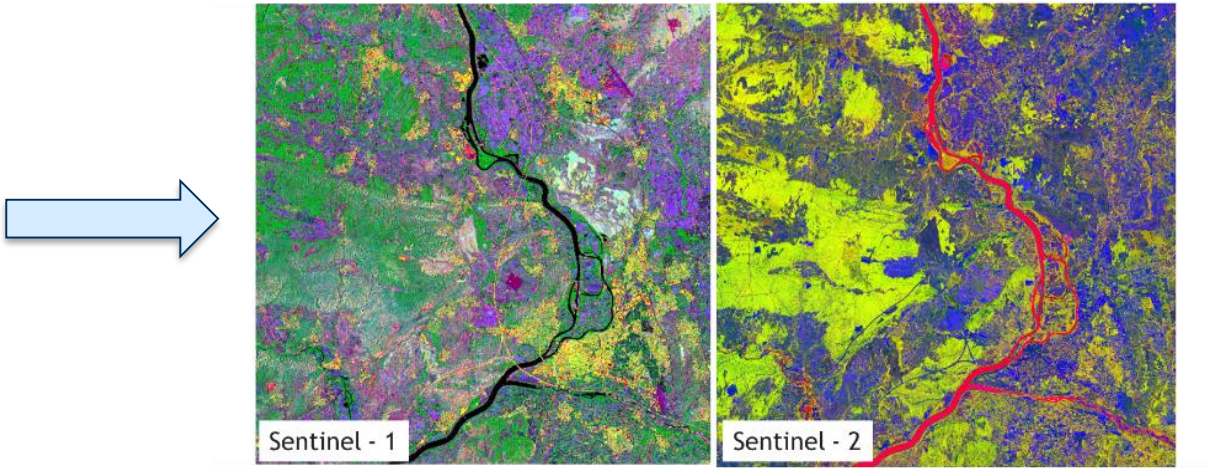
- Implemented in **15 python** modules
- 3 years data in ‘base’ mode
- **270 metrics** per tile
- 18962 tiles:
  - 5 folded CV
  - 6 classifiers + 7 regressors (+urban, water, crop)
- Up till now **31 continental** runs performed
- Using 500 executors, within 1GB memory
  - Data preparation takes 2-3 days
  - Classification < 1 day

# SCALING UP TO PUBLIC CLOUD

# Continuity & Higher spatial resolution

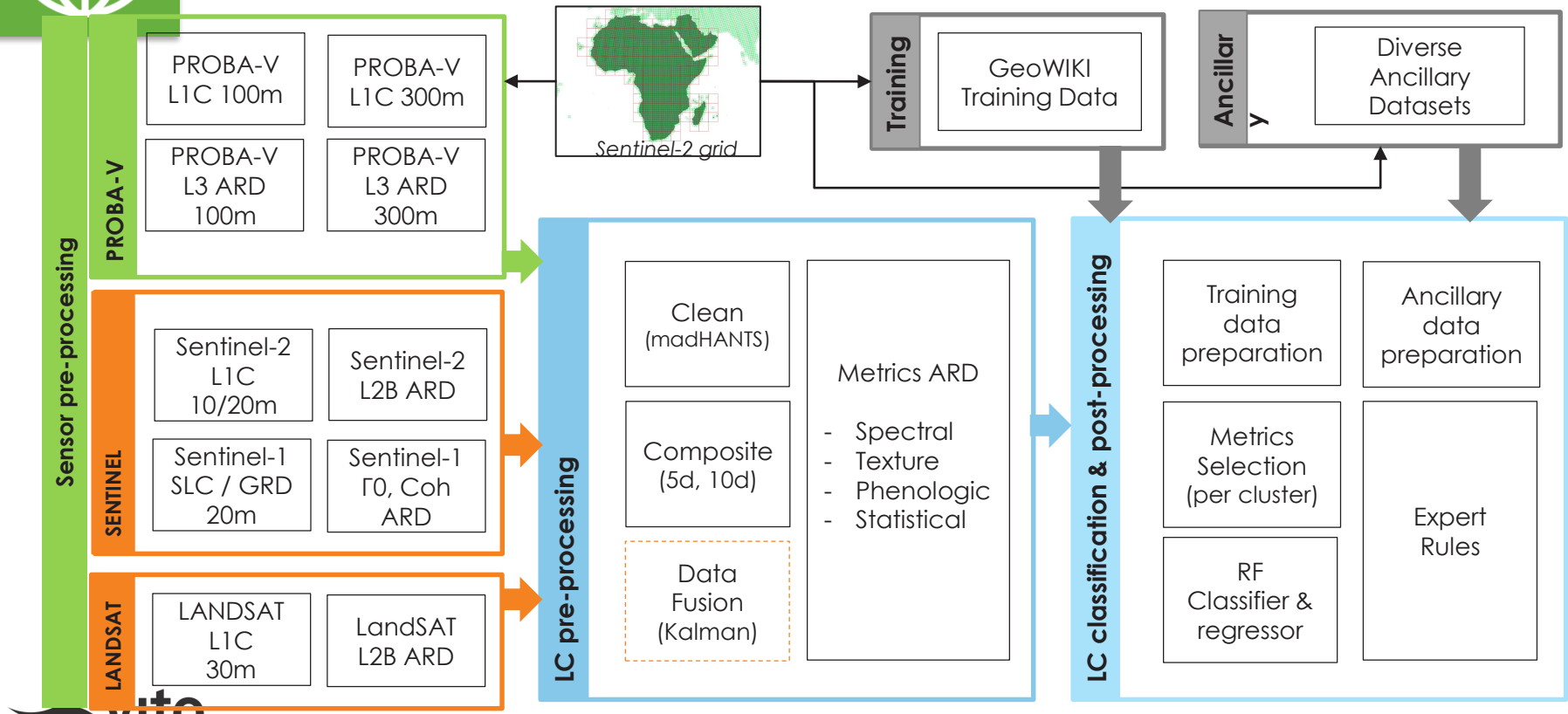


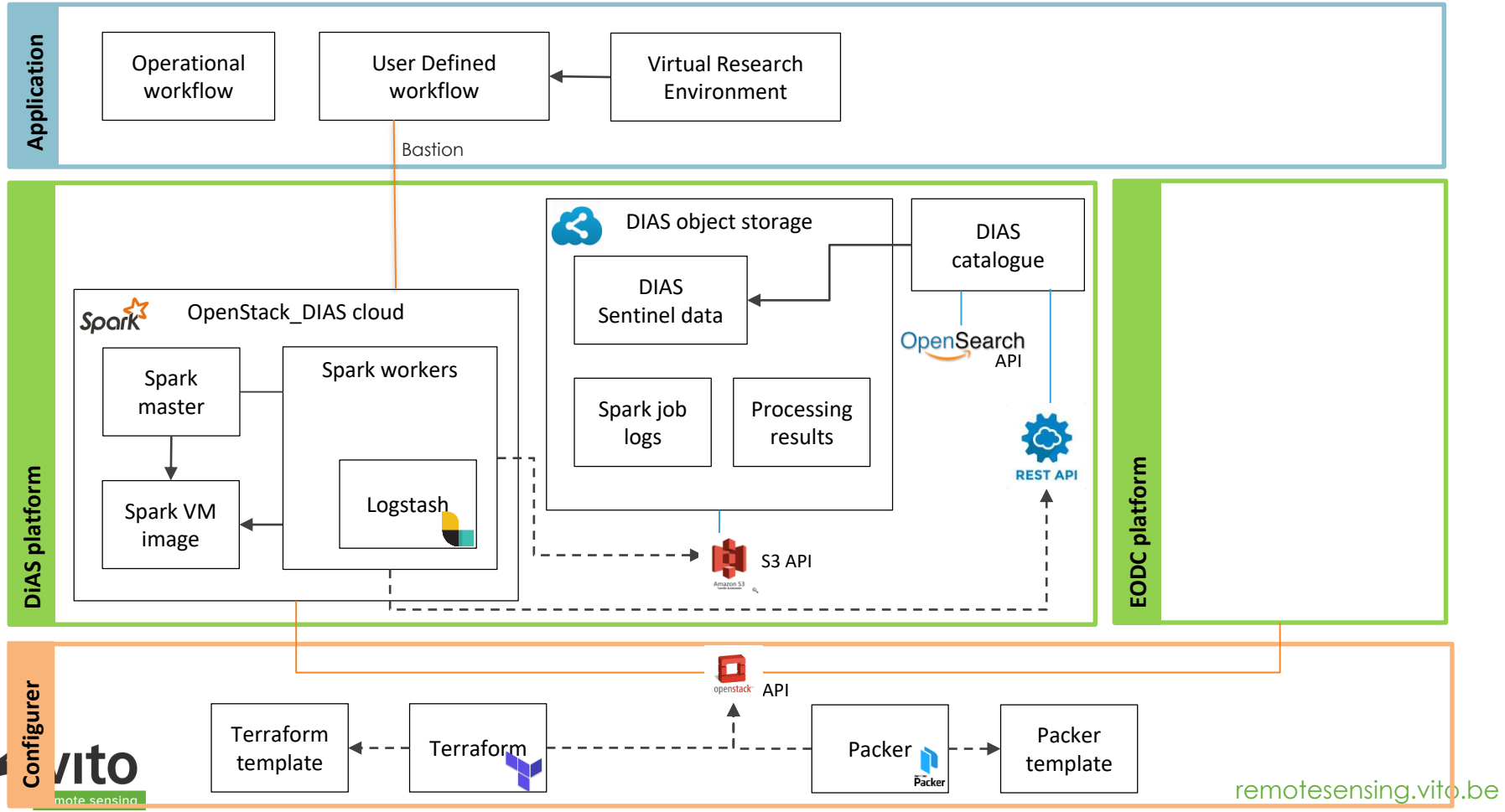

100m

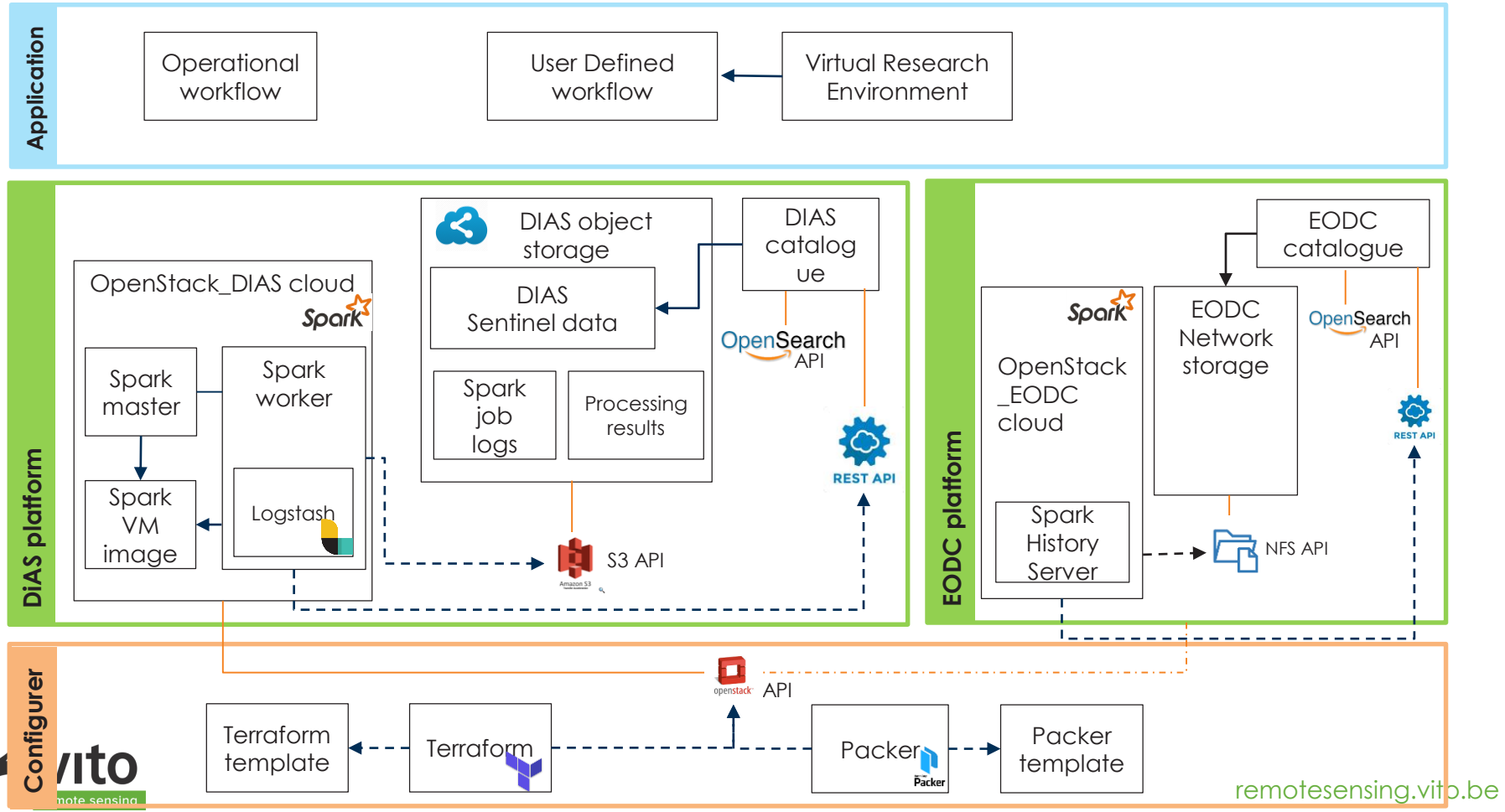



10m

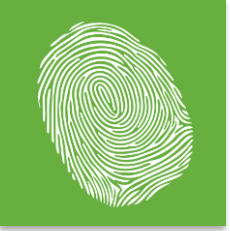
# Sensor agnostic (fusion), re-use LC workflow




Application
Operational workflow
User Defined workflow
Virtual Research Environment
Bastion
DiAS platform
DIAS object storage
DIAS Sentinel data
DIAS catalogue
OpenStack_DIAS cloud
Spark
Spark master
Spark workers
Spark VM image
Logstash
Spark job logs
Processing results
OpenSearch API
REST API
S3 API
EODC platform
Configurer
openstack
API
Terraform template
Terraform
Packer
Packer template
vito
remotesensing.vito.be


Application
Operational workflow
User Defined workflow
Virtual Research Environment
DiAS platform
OpenStack_DIAS cloud
Spark
Spark master
Spark worker
Spark VM image
Logstash
DIAS object storage
DIAS Sentinel data
Spark job logs
Processing results
DIAS catalog ue
OpenSearch API
REST API
S3 API
EODC platform
EODC catalogue
OpenStack _EODC cloud
EODC Network storage
OpenSearch API
REST API
Spark History Server
NFS API
Configurer
openstack API
Terraform template
Terraform
Packer
Packer template
vito
remotesensing.vito.be

# CLOSING



# Where are we ?

- Full automated Land Cover workflow
  - User-customization in classes possible through 4->7 covers
  - Include Spatial accuracy maps (& Change maps)
  - Large 10m database of high quality training (&validation) points
  - Sensor agnostic
    - Global on PROBA-V UTM
    - AOI on Sentinel & Landsat

- Highly optimized workflow in Spark
  - Stable in private cloud environment
  - Infrastructure agnostic (Spark standalone)
  - Tests in public cloud environment ongoing

# Next steps

- LC100 v2 (Global 2015 + Africa change)
  - Release at Living Planet Symposium


*+ 7 Cover Fractions*
*+ Data Density Indicator*
*+ Spatial Accuracy Map*

- LC100 v3
  - Global Sentinel (1+2)
  - Explore/integrate new AI/ML techniques (reCNN)


**Bruno Smets**
R&D Project Manger Land Use
**VITO NV**
Remote Sensing | Vegetation
Boeretang 200 | 2400 Mol | Belgium
T. +32 14 33 68 39
remotesensing.vito.be | blog.vito.be/remotesensing | @VITO_RS


LOOKING AHEAD, TOGETHER

vito remote sensing

# THANK YOU

remotesensing.vito.be

*Sentinel-2 image Copernicus Sentinel data (2016)*