



The LifeWatch ERIC Biodiversity
& Ecosystem eScience Conference

PLOVDIV, BULGARIA

7-10 JULY 2026

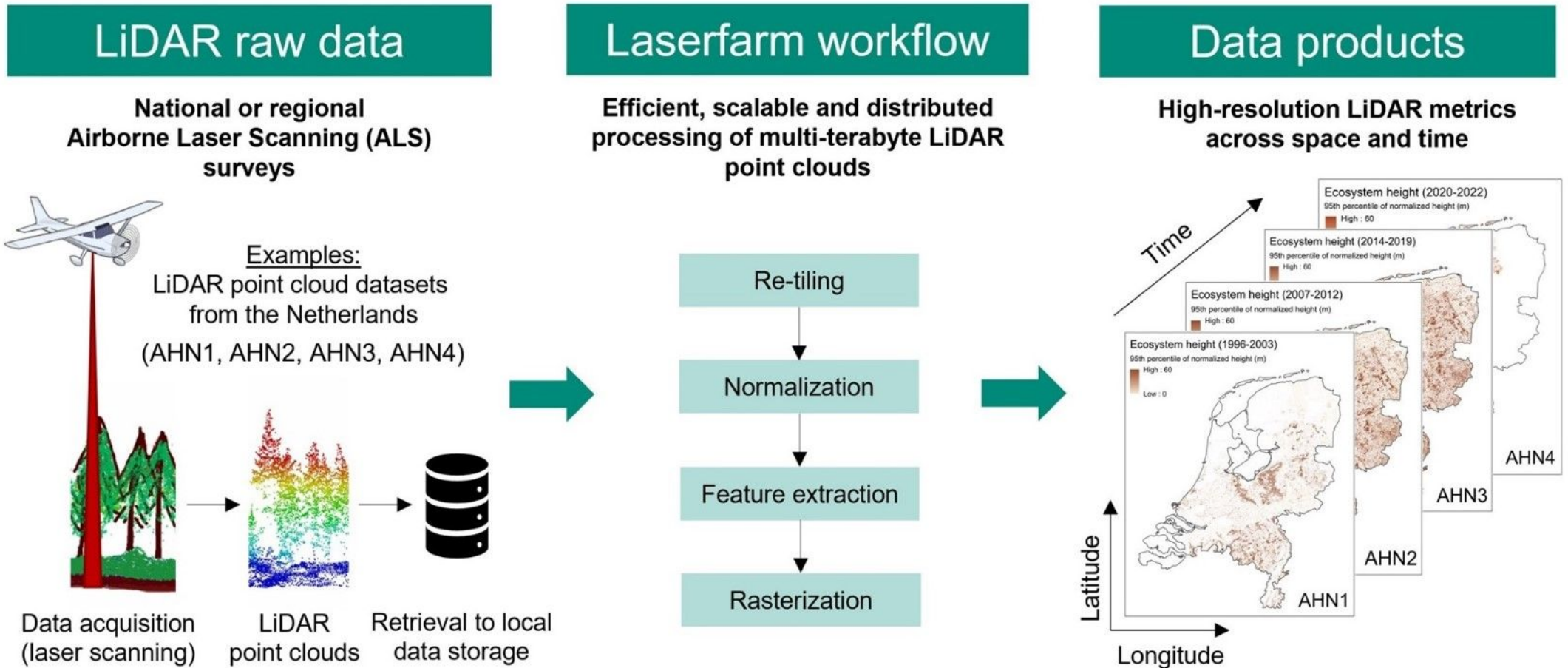
Addressing environmental challenges
through large-scale data analysis in a
shared virtual lab: NaaVRE Demo

Presenter: Koen Greuell, LifeWatch ERIC

10 July 2026

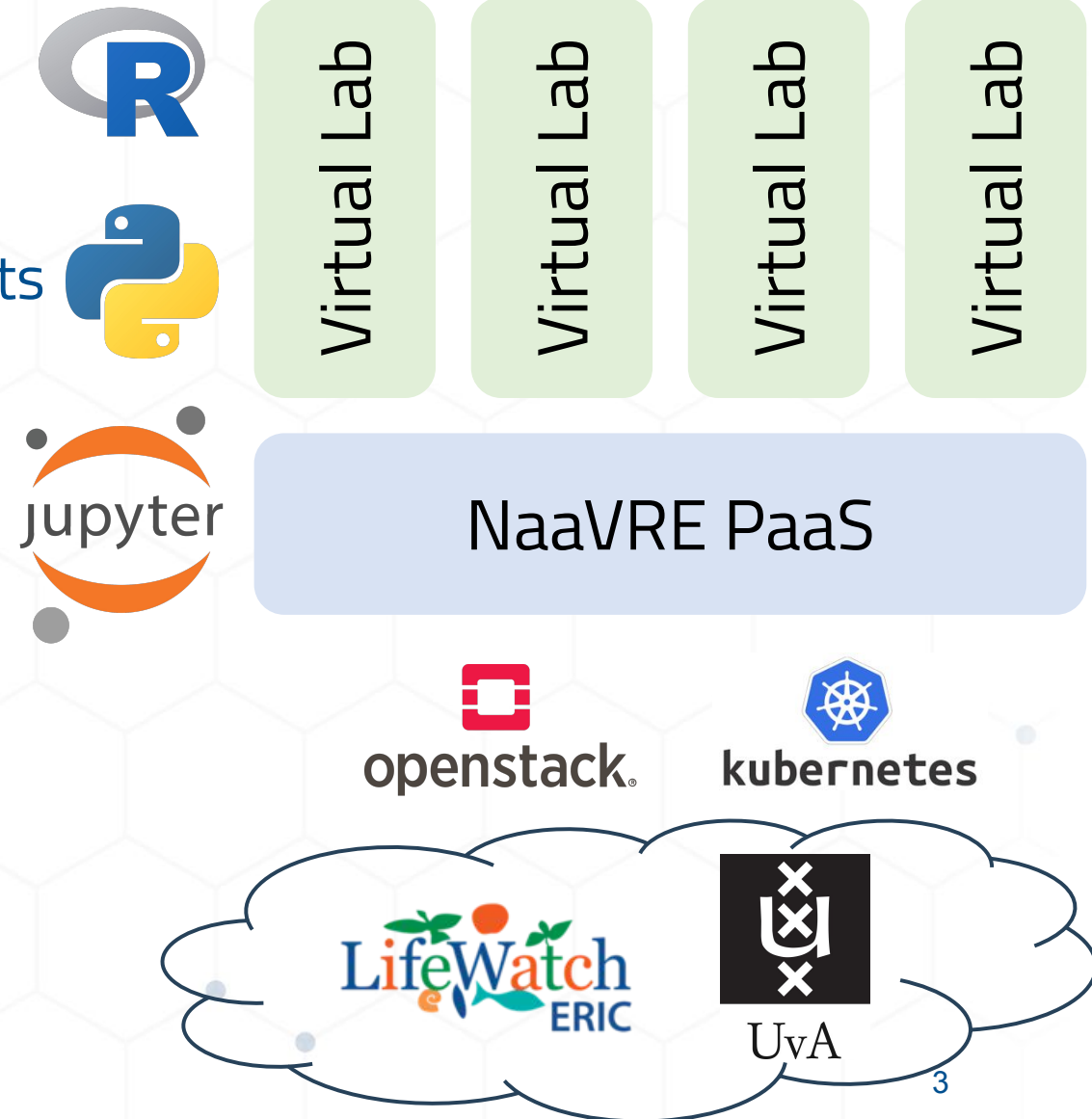
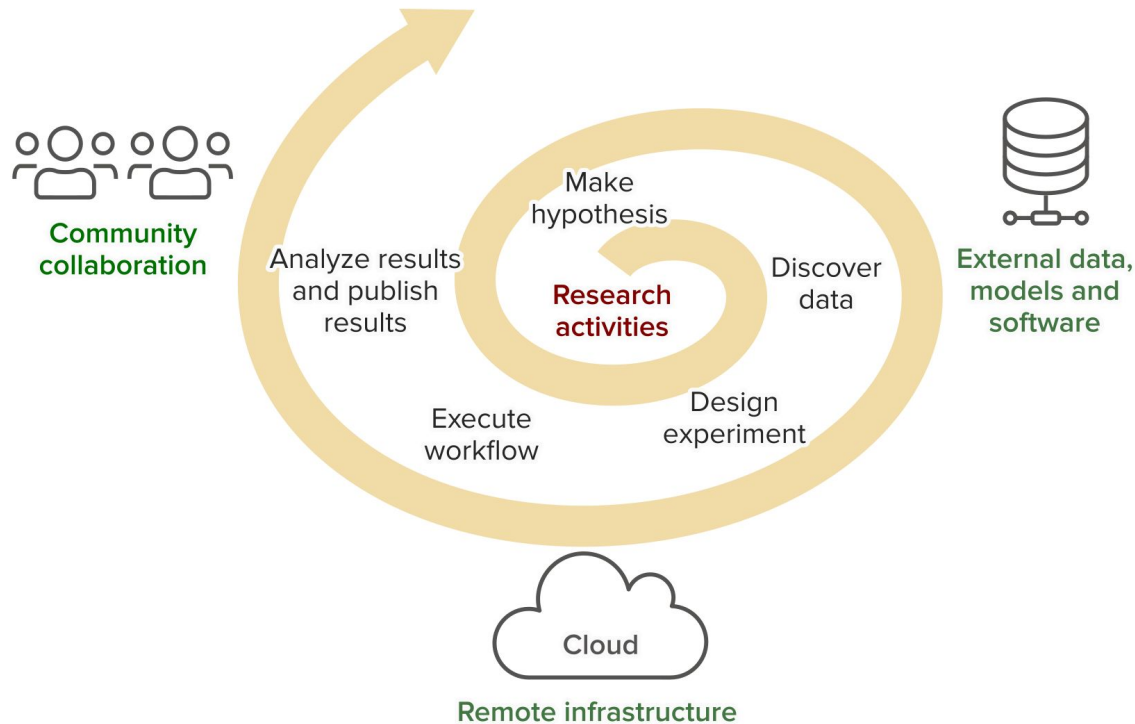


Notebook-as-a-Virtual-Research-Environment used for country-wide vegetation metrics analysis using LiDAR



Notebook-as-a-Virtual-Research-Environment (NaaVRE)

- Researchers code their own workflows
 - Write code in a Jupyter Notebook
 - Turn the code into workflow components
 - Run workflows on the cloud



Roles involved in virtual lab development

Virtual lab researchers:



Domain scientist: Research question & scientific exploration.



Modelers: Develop mechanistic, statistical, or machine learning models.



Workflow developer: Turn research ideas into code.

Support roles:



Virtual research environment engineer: Advises virtual lab researchers.



Training and support staff: Organize training activities and support the user community.

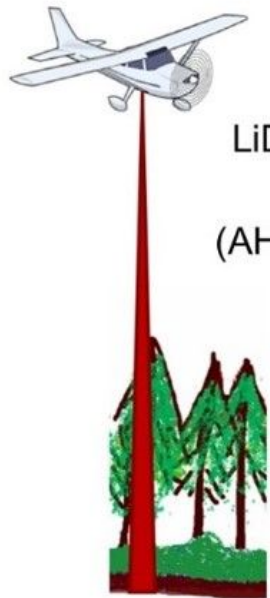


Infrastructure operation supporter: Provision computing and storage resources.

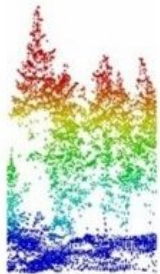
Development journey of a virtual lab for country-wide vegetation metrics analysis

LiDAR raw data

National or regional
Airborne Laser Scanning (ALS)
surveys



Examples:
LiDAR point cloud datasets
from the Netherlands
(AHN1, AHN2, AHN3, AHN4)



Data acquisition
(laser scanning)

LiDAR
point clouds

Retrieval to local
data storage

Laserfarm workflow

Efficient, scalable and distributed
processing of multi-terabyte LiDAR
point clouds

Re-tiling

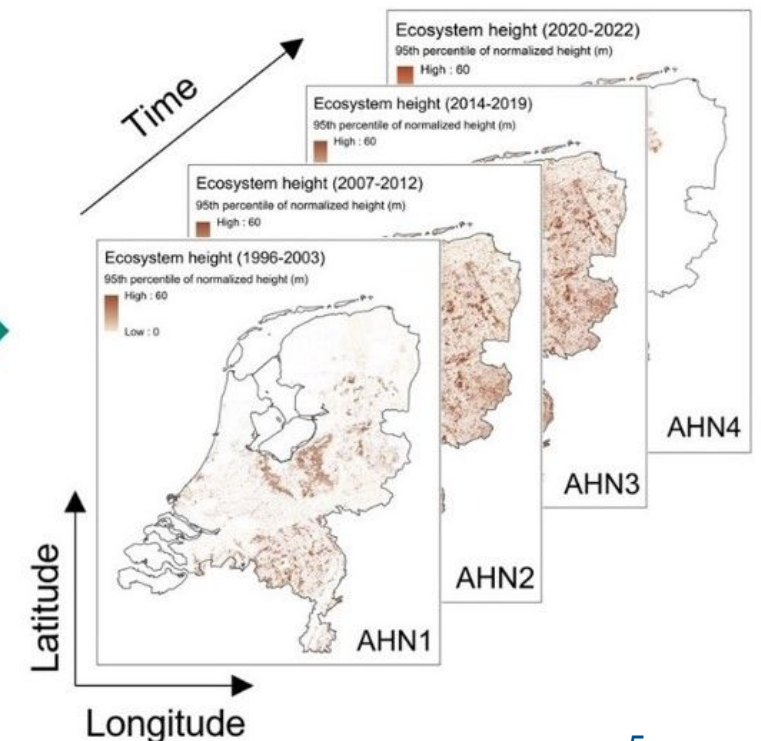
Normalization

Feature extraction

Rasterization

Data products

High-resolution LiDAR metrics
across space and time



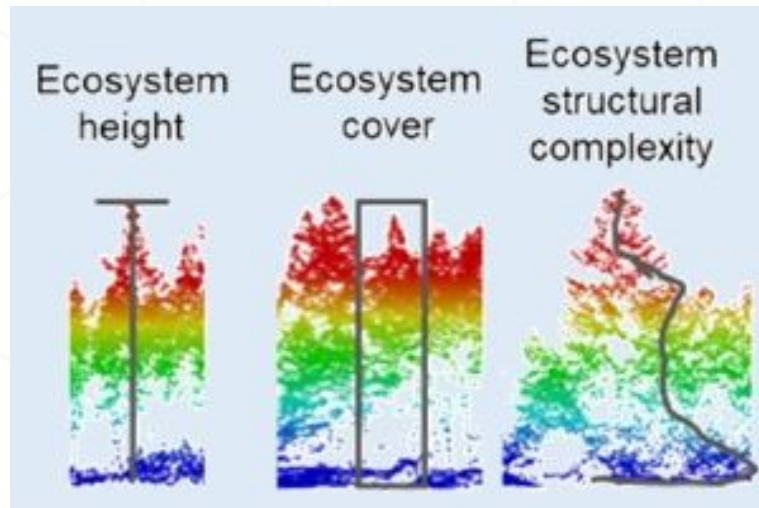


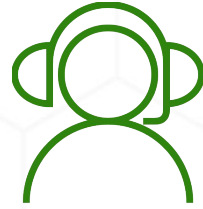
Level 0

The initial proposal of a virtual lab



A domain scientist started the virtual lab with a goal: Map the ecosystem structure of the Netherlands.





Level 0

The initial proposal of a virtual lab

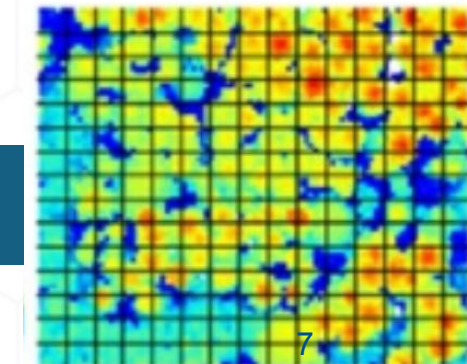
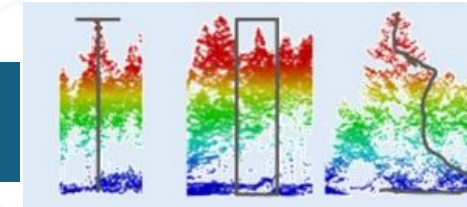
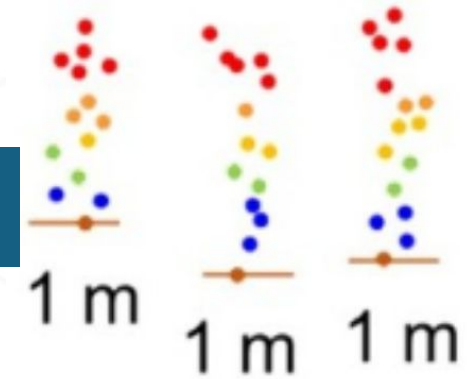
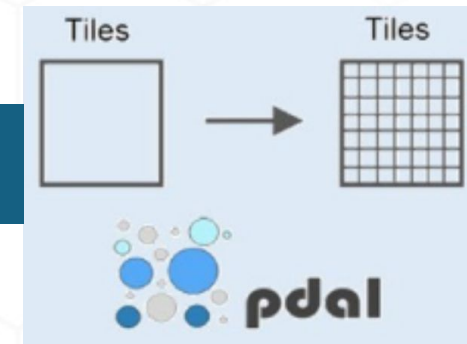
Domain scientist & VRE engineer: co-designed a draft workflow to get from: LiDAR data to: Vegetation metrics

Re-tiling

Normalization

Feature extraction

Rasterization





Level 0

The initial proposal of a virtual lab



Domain scientist, modeler & VRE engineer: Assessed the computational challenges and requirements.

- Data

- Source: Public repository with API
- Size:
 - Total 12 TB
 - Files between 100MB and 500MB

- Computing

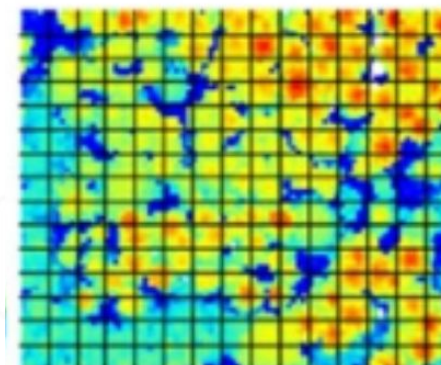
- Algorithms: existing Laserfarm library
- Parallelization: Workflow steps must be sequential.
- RAM: Rasterizing many files at once reduces noise.

Re-tiling

Normalization

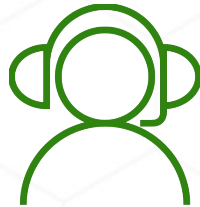
Feature extraction

Rasterization





Domain scientist: Arranged a team to develop the virtual lab.



VRE engineer: Created a new virtual lab.

Level 0
The initial proposal of a virtual lab



Home
About

Logged in as Koen Greuell
[Logout]

NaaVRE: Notebook as a Virtual Research Environment

We have simplified the sharing of notebook files, workflow components, and workflow files. Check out [the assets catalogue documentation](#)

If you encounter an issue or need help with anything, [get in touch](#) or submit a bug report on [GitHub](#).

[Documentation](#)

Pinned virtual labs



NaaVRE Tutorial

NaaVRE Open Lab

Open experimentation with virtual labs



ECVs

Essential Climate Variables Virtual Lab

A virtual lab to access ECVs. The workflows in the VLab were developed by ENVRI-Hub-NEXT project.

All virtual labs

laserfarm

ANERIS

ECVs

Experimental features

LTER-LIFE

Marco-Bolo

NaaVRE Tutorial

Tutorial



Tutorial

Laserfarm collaborative virtual labs

LiDAR point cloud analysis for macro-ecology



Level 0

The initial proposal of a virtual lab

Laserfarm Virtual Lab

- ☐ Development plan
- ☐ Virtual lab development team



Level 1

Implementation of the virtual lab functionality

Domain scientist & workflow developer:

Level 1

Implementation of the virtual lab functionality

Jupyter Notebooks

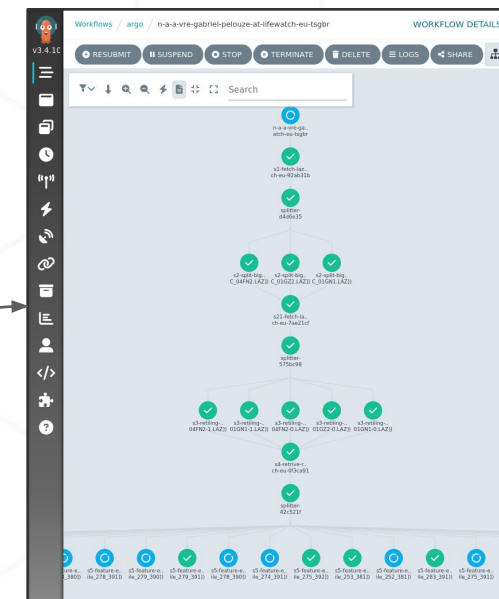
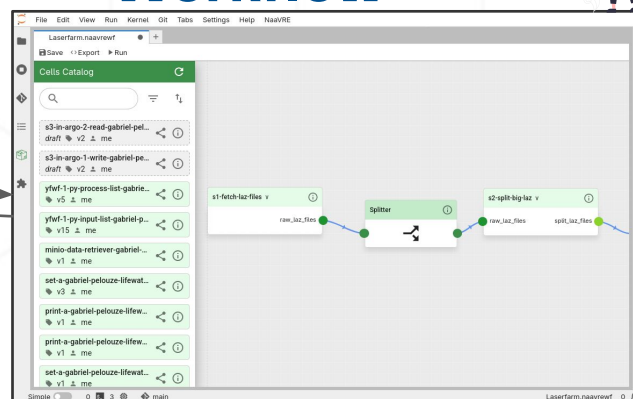
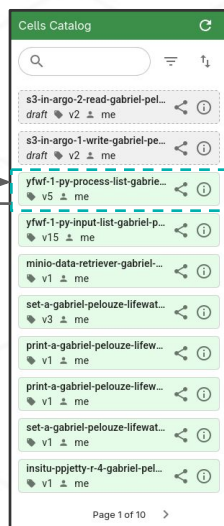
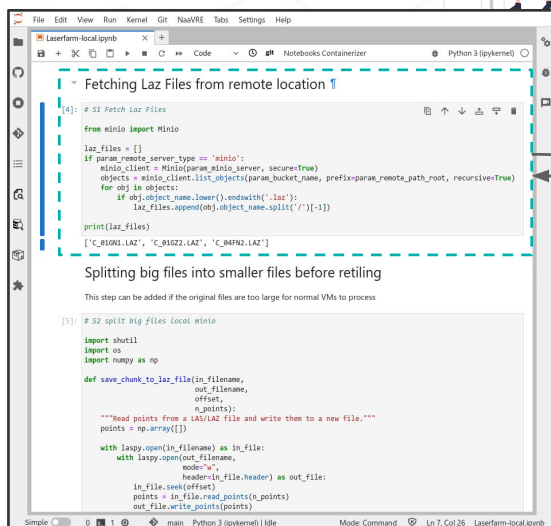
Workflow components

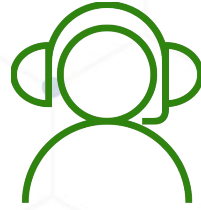
Workflow

Results

Data storage

VRE assets catalogue





Level 1

Implementation of the virtual lab functionality

VRE engineer & infrastructure operation supporter: Prepare infrastructure for large scale experiments.



Level 1

Implementation
of the virtual lab
functionality

- ☐ Source code & workflow created.
- ☐ Verified on a small testing dataset.
- ☐ Infrastructure ready for large scale experiments.



Level 2

Run large scale
scientific
experiments and
write scientific
publications



Domain scientist & workflow developer: Run large scale experiments & analyze results



Level 2

Run large scale scientific experiments and write scientific publications

Run Workflow

File Edit View Run Kernel Git Tabs Settings Help NaaVRE

Cells Catalog

- s3-in-argo-2-read-gabriel-pe-draft v2.0 me
- s3-in-argo-1-write-gabriel-pe-draft v2.0 me
- ywf-1-pp-process-list-gabriel-pe v5.0 me
- ywf-1-pp-input-list-gabriel-pe v15.0 me
- minio-data-retriever-gabriel-pe v1.0 me
- set-a-gabriel-pelouze-lifewat v3.0 me
- print-a-gabriel-pelouze-lifewat v1.0 me
- print-a-gabriel-pelouze-lifewat v1.0 me
- set-a-gabriel-pelouze-lifewat v1.0 me

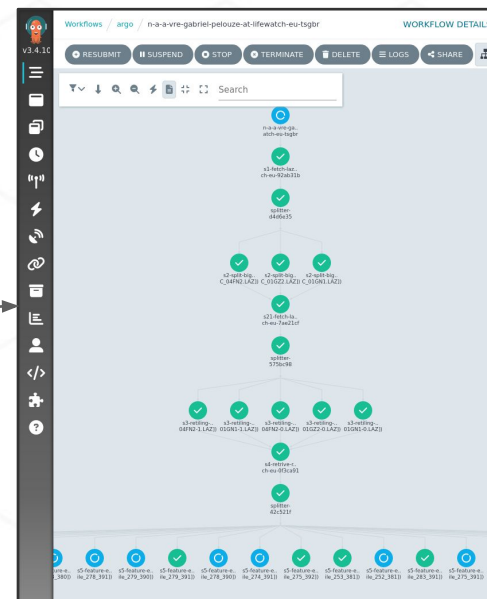
Workflow graph showing tasks: x1-fetch-laz-files v, splitter, x2-split-tyg-laz v, x3-metacat, x4-metacat, x5-metacat, x6-metacat, x7-metacat, x8-metacat, x9-metacat, x10-metacat.

param_use_own_data: 0

param_validate_precision: 0.001

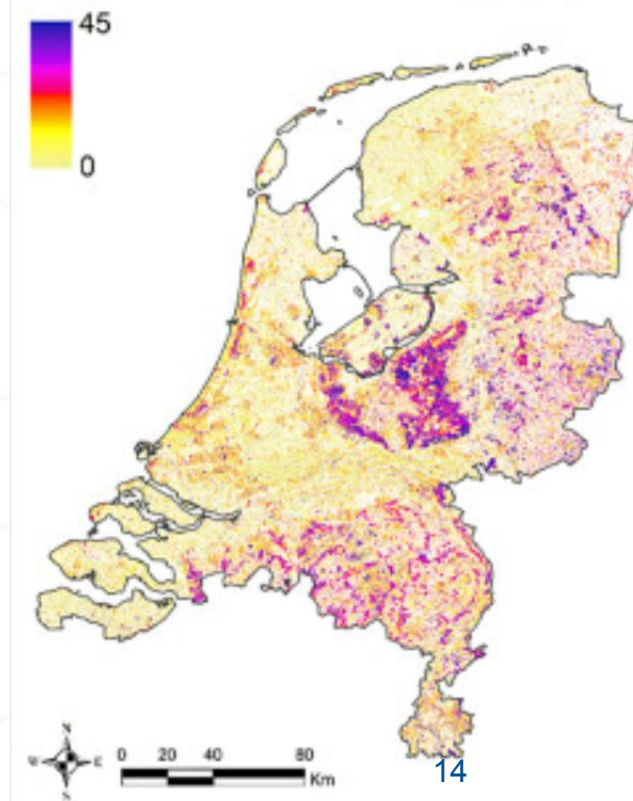
param_workshop_folder_base: vl-laserfarm/Workshop/

Make recurring RUN



Results

75th percentile of normalized height (m)





Level 2

Run large scale scientific experiments and write scientific publications

- ☐ The workflow is run on a scientific dataset.
- ☐ The results are analyzed.
- ☐ Workshop / tutorial material has been created.
- ☐ The workflow can be run on different datasets.
- ☐ Source code, workflow and workflow components are publicly available.



Level 3

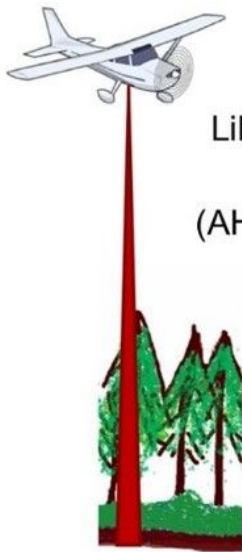
Engage more users through training and hakton events to drive community building

Laserfarm virtual lab for country-wide vegetation metrics analysis

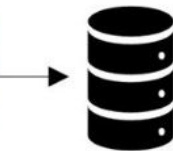
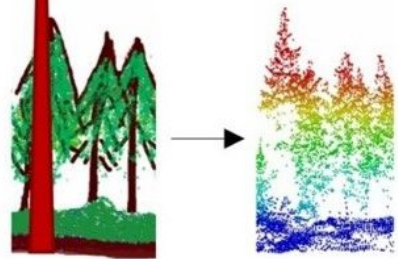


LiDAR raw data

National or regional
Airborne Laser Scanning (ALS)
surveys



Examples:
LiDAR point cloud datasets
from the Netherlands
(AHN1, AHN2, AHN3, AHN4)



Data acquisition
(laser scanning)

LiDAR
point clouds

Retrieval to local
data storage

Laserfarm workflow

Efficient, scalable and distributed
processing of multi-terabyte LiDAR
point clouds

Re-tiling

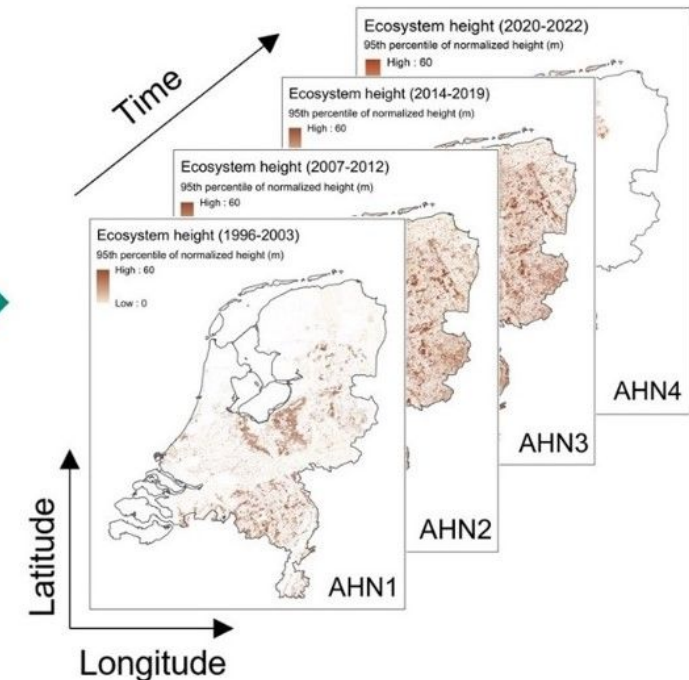
Normalization

Feature extraction

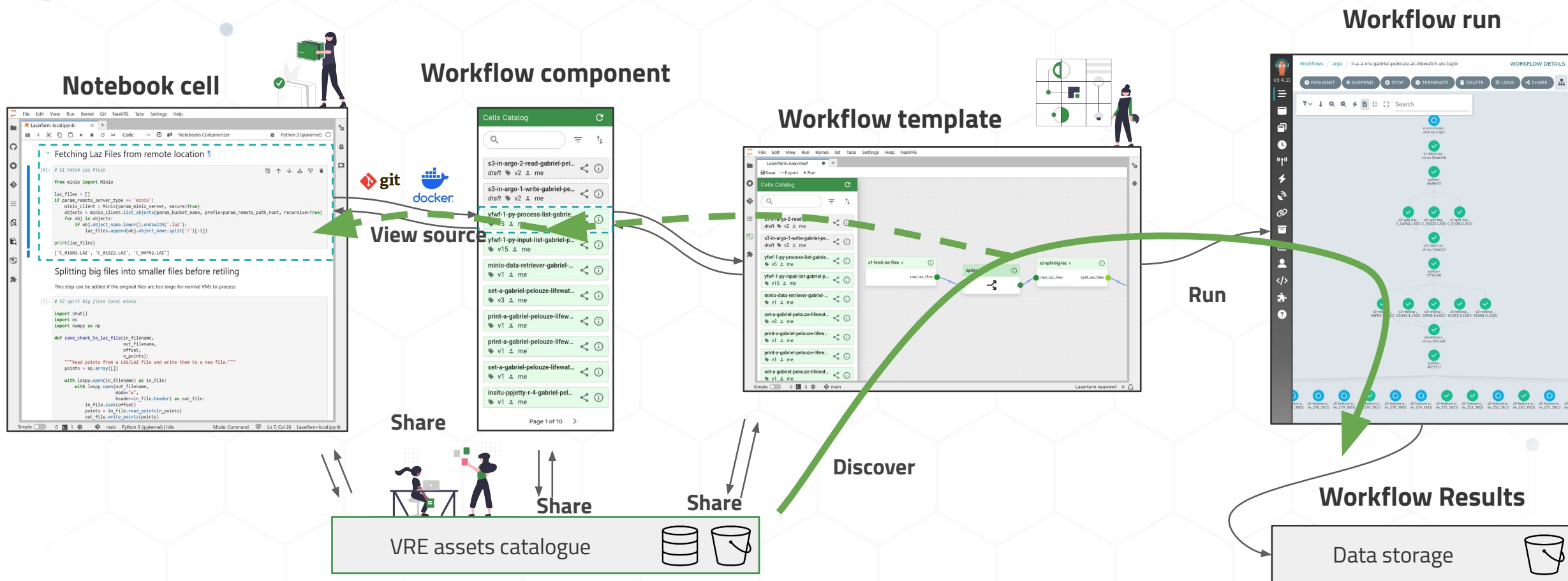
Rasterization

Data products

High-resolution LiDAR metrics
across space and time



Other researchers can run the workflow on their own data and modify the components.





The LifeWatch ERIC Biodiversity
& Ecosystem eScience Conference
PLOVDIV, BULGARIA 7-10 JULY 2026

THANK YOU!

Questions?

vlc@lifewatch.eu

