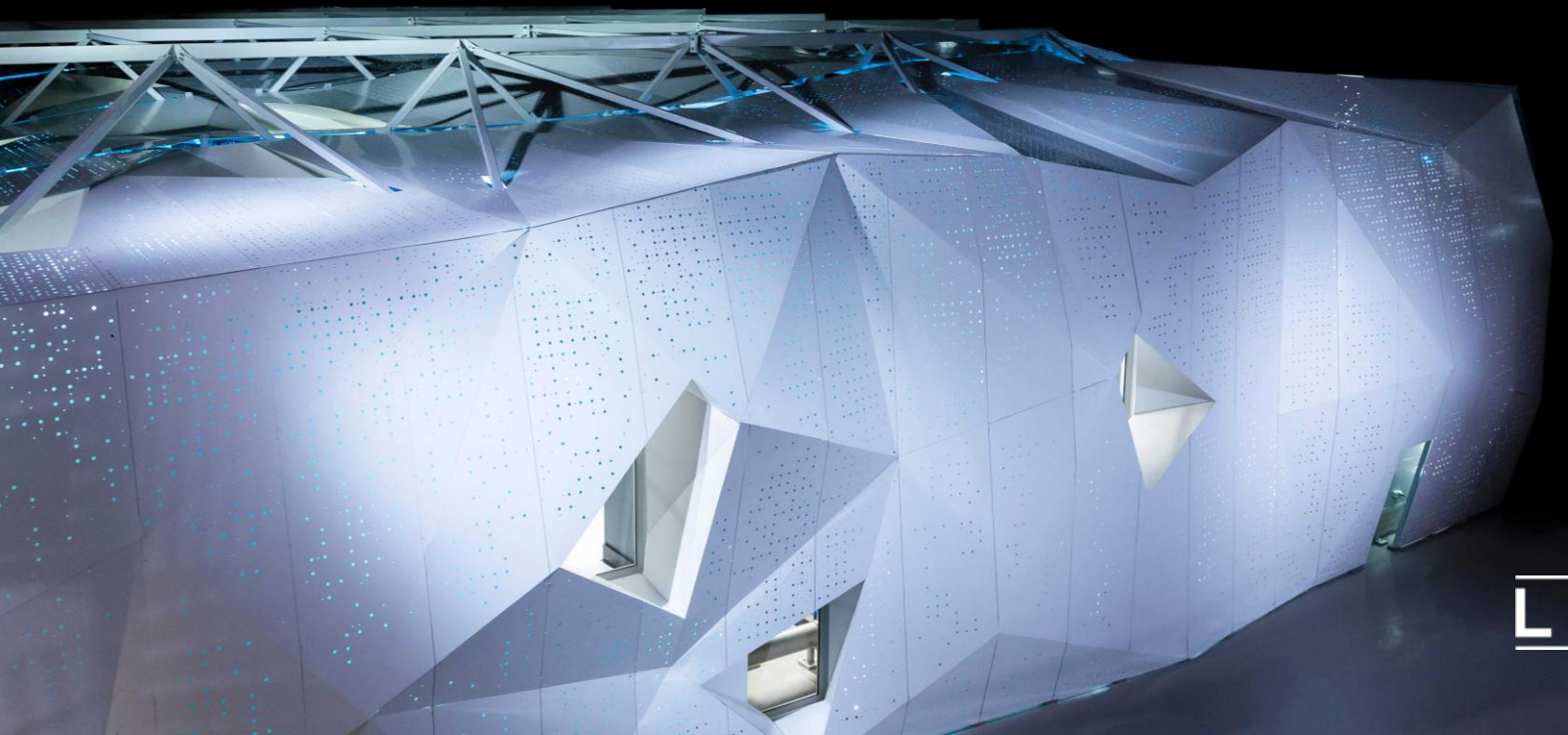
LUMI Supercomputer for geospatial companies

Katri Tegel, Kylli Ek, Jyrki Savolainen, CSC – IT Center for Science Ltd.

04.09.2025



Agenda

- LUMI supercomputer overview
- Spatial analysis LUMI use cases
- Working with LUMI
- Access to LUMI
- Next steps

Non-profit state organization with special tasks

Turn over in 2023

73,6 M€

csc

LUMI

Owned together with EuroHCP JU and unique consortium of 10 countries – Belgium, Czech Republic, Denmark, Estonia, Finland, Iceland, Norway, Poland, Sweden and Switzerland.

Headquarters in Espoo, data center in Kajaani

Owned by the Finnish state **(70%)** and all Finnish higher education institutions **(30%)**

711

(09.2025)

Geoportti and Location Innovation Hub EDIH

LUMI

Geoportti – The Hub for Finnish Geospatial Research and Education Resources

- A shared service for using geospatial data and geocomputing tools. Geoportti RI helps the researchers in Finland to use, to refine, to preserve and to share their geospatial resources.
- Coordinated by National Land Survey of Finland (NLS)

https://WWW.geoportti.fi/

Location Innovation Hub (LIH)
European Digital Innovation Hub (EDIH)

- Large network that consists of companies, research organizations and universities, as well as the public sector.
- Goal: to expand the use of geospatial data in services, new business and technologies, and produce new benefits and innovations for society based on them.
- Coordinated by NLS

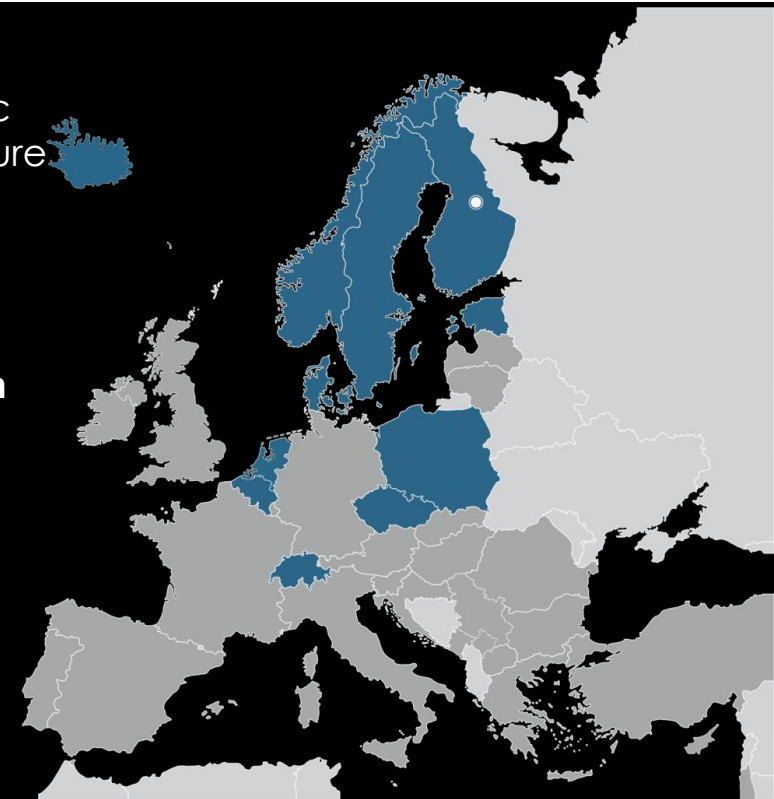
https://locationinnovationhub.eu/en/home/

LUMI supercomputer

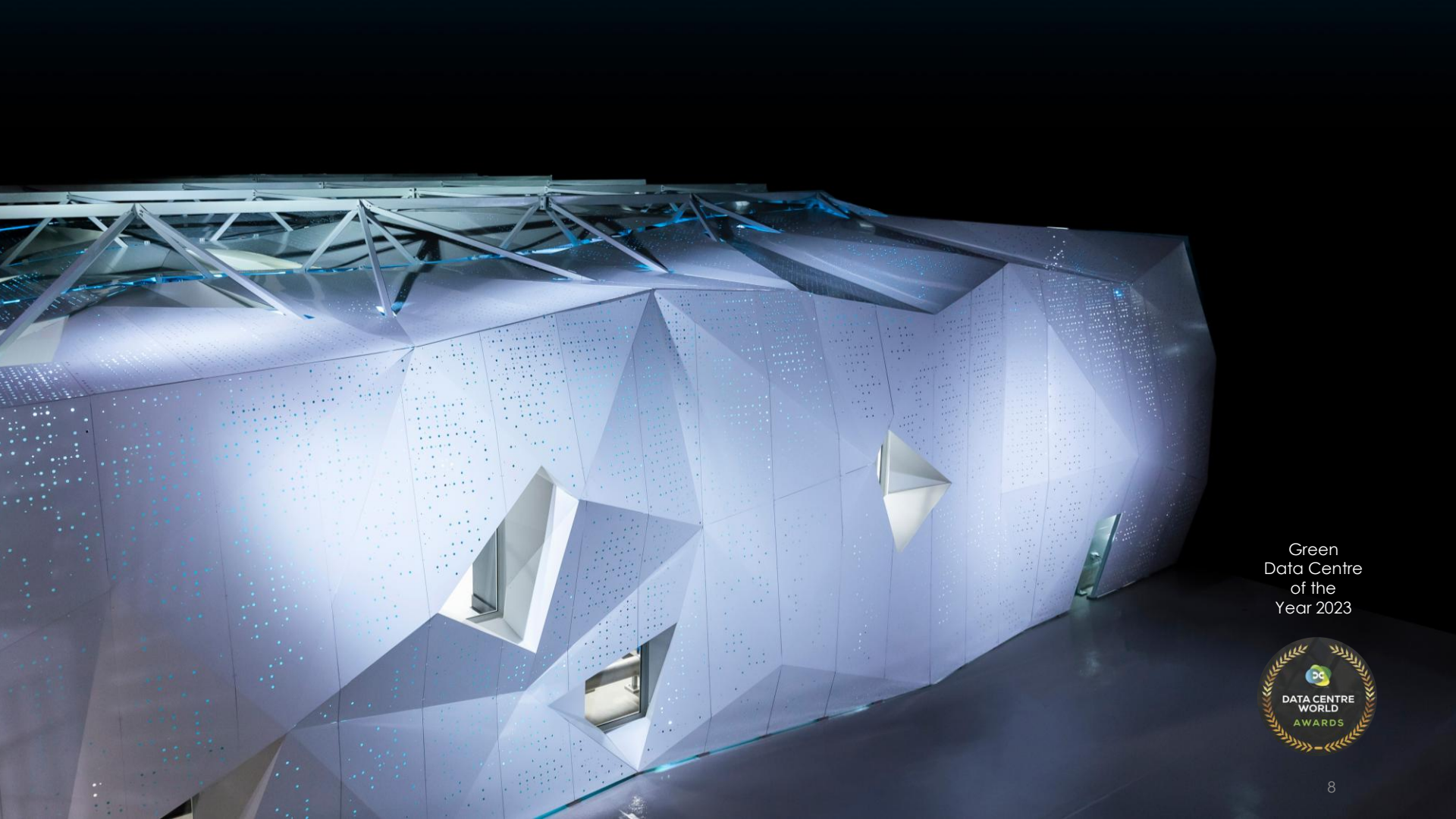
EuroHPC Joint Undertaking: Europe's world-class scientific computing and data management research infrastructure (high-performance computing, HPC)

EuroHPC funding: EU and participating countries

- **The supercomputer hosted by the LUMI consortium** (11 countries) **in Kajaani**
- Developing competence in high-performance computing
  - **CSC operates EuroCC Finland, National comptence center in Finland**



Renforsin Ranta Business Park, Kajaani


Green
Data Centre
of the
Year 2023
DATA CENTRE
WORLD
AWARDS

LUMI is an HPE Cray EX Supercomputer

LUMI

Hewlett Packard Enterprise

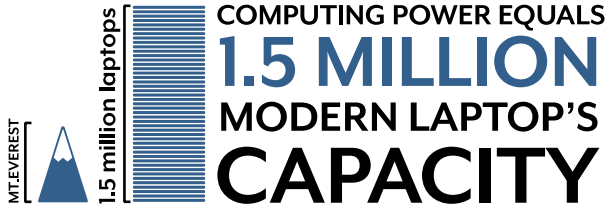
Its powerful processors can perform millions of calculations simultaneously, so it can operate millions of times faster than conventional computers and help us solve highly complex problems.

9th fastest supercomputer globally

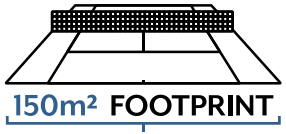
LUMI

LASKENTATEHO

375 PETAFLOP/S



2 x



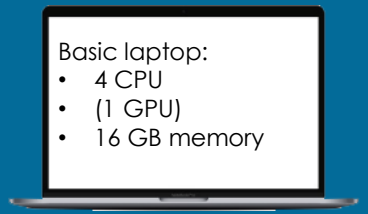
Suurteho-laskennalle | **AI** | Data-analytiikalle

200 000 000+ €

Computing resources per project/job

- CPU: max 65 536 cores
  = 128 cores x 512 nodes
- **GPU: max 8 192 GPUs**
  = 8 GPU x 1024 nodes
- Memory: max 1 Tb
- HPC local storage: max 500 Tb

+ LUMI-O S3 storage
150+ TB



Why use a supercomputer?

LUMI

- Computing resource needs
  - time, memory, storage, GPU
- Prebuilt environments
- Run many experiments at the same time
- **LUMI / CSC specialist support**

Enabling work that would either be impossible or slower on smaller systems.

LUMI usage

LUMI

- **DEEP LEARNING MODEL TRAINING**
- DEEP LEARNING INFERENCE
- Big data analysis
- Advanced modelling
- Not for web services nor time-critical/real-time applications

LUMI AI Factory

LUMI AI Factory

Empowering the European AI ecosystem

LUMI AI Factory

The three pillars of LUMI AI Factory

- AI-optimised supercomputer LUMI AI
- AI-factory service center
- Experimental quantum-computing platform LUMI-Q

CSC (Finland) coordinates consortium with participation from Czechia, Denmark, Estonia, Norway and Poland

- Other Finnish partners are FCAI (Aalto University, Univ of Helsinki) and AI Finland (Technology industries)

Total budget over 612 million euros

- EU 306.4M€, FI 250M€, CZ 11M€, DK 10M€, EE 5M€, NO 20.4M€, PL 10M€

- Largest public computing ecosystem investment in Finland, among the largest in Europe
- Largest EuroHPC AI Factory investment

Significant investment in talent and competence development

https://www.lumi-supercomputer.eu/lumi-ai-factory/

Timeline for the LUMI AI Factory

LUMI AI Factory

Supercomputing use cases

ICEYE

- Finnish space technology company
- Microsatellites for global real-time radar (SAR)
- Several use case about floods and wildfires.
- Deep learning models training and optimization

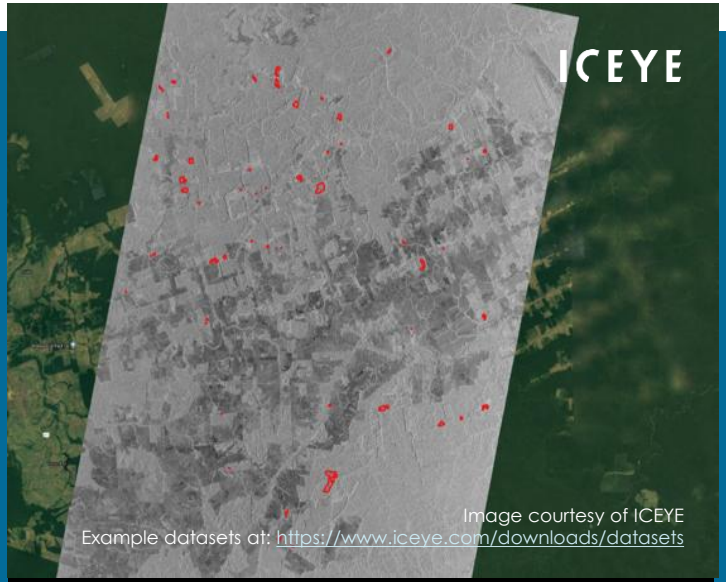


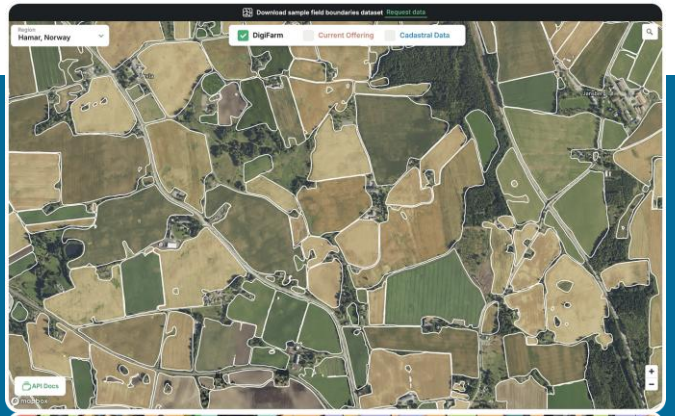
Image courtesy of ICEYE
Example datasets at: https://www.iceye.com/downloads/datasets

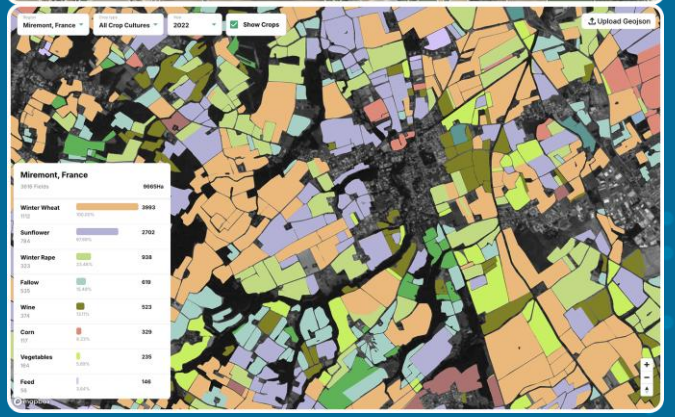
READ ARTICLE:

Space Technology Company ICEYE praises LUMI user support: "A CSC expert guided us step by step."
https://www.linkedin.com/pulse/space-technology-company-iceye-praises-lumi-user-support-bcanf/

DigiFarm

- Norwegian AI start-up
- Deep learning models training and optimization
- Precision farming services
  - Automatic Field Boundary Delineation
  - Crop Suitability Mapping
  - Multi-Year Zoning Analysis
  - Monitor crop health and growth dynamic
  - In-season Crop Identification





READ ARTICLE:

A EuroHPC Success Story: HPC on the Farm
https://www.eurohpc-ju.europa.eu/eurohpc-success-story-hpc-farm_en

Images from DigiFarm:
https://digifarm.io/blog/empowering-customers-seed-production

BioDT / University of Jyväskylä

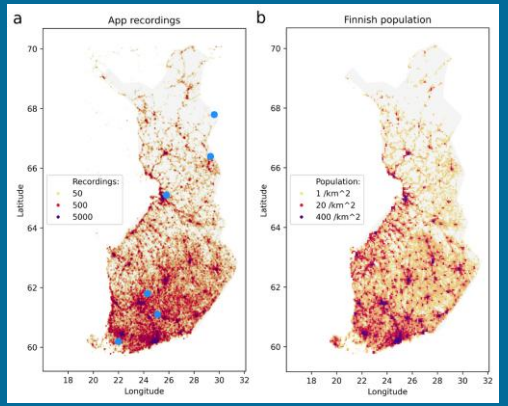
LUMI

- Bird sound recognition CNN model training and optimization
- Joint species distribution modelling with Hmsc R-package and Tensorflow

READ ARTICLES:

A Mobile Application–Based Citizen Science Product to Compile Bird Observations
https://doi.org/10.5334/cstp.710

Accelerating joint species distribution modelling with Hmsc-HPC by GPU porting
https://doi.org/10.1371/journal.pcbi.1011914

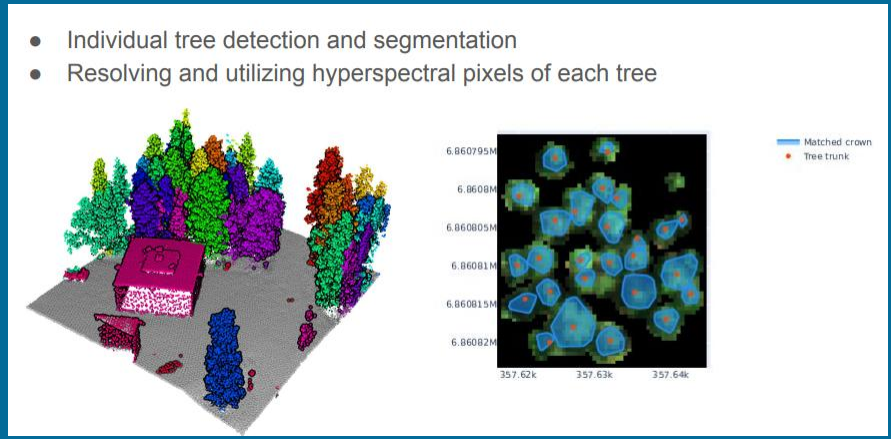


Images from the first article

Aalto University / Department of Computer Science

LUMI

- Superresolution of hyperspectral images
- Modeling and prediction of forestry variables from hyperspectral images
- Individual tree detection and segmentation from LiDAR
- Tree species classification from tree segments and hyperspectral data
- Modeling of forest diversity and naturalness measures



READ ARTICLES:

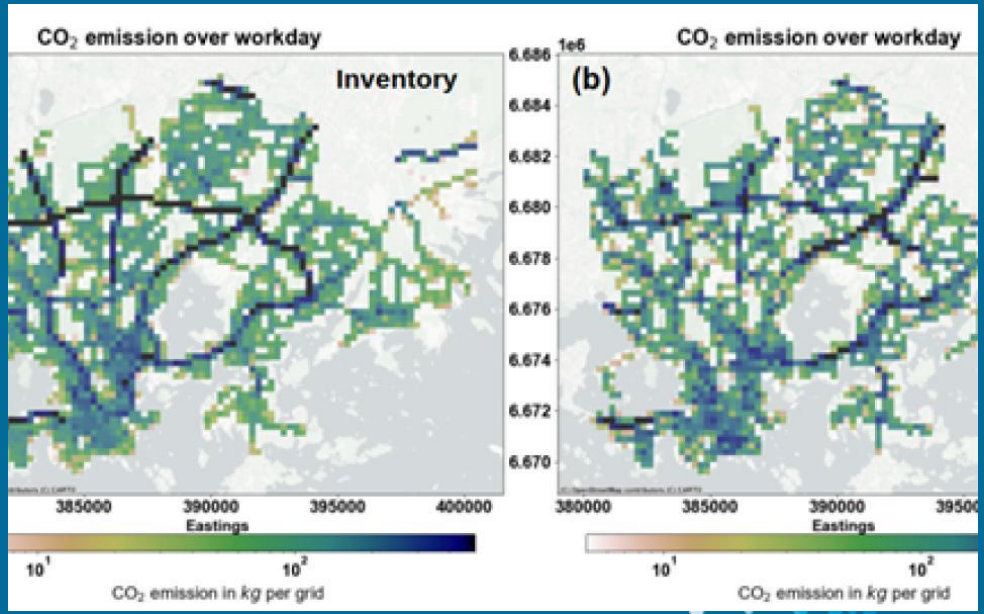
A Comparison of Hyperspectral Super-Resolution Techniques for Boreal Forest Imagery
https://doi.org/10.1109/IGARSS53475.2024.10640724

Mesh Surface And Morphological Hierarchies For Individual Tree Detection And Segmentation From LiDAR Data
https://doi.org/10.1109/IGARSS53475.2024.10640418

Images from the first article

Helsinki University / Spatiotemporal data analysis group

- Dynamic on-road emission estimation
- City-wide emission upscaling for deep reinforcement learning methods
- Design of dynamic routing strategies
- Design of reward functions for Electric Vehicle drivers



READ ARTICLES:

Artificial Intelligence for Urban Low-Emission Autonomous Traffic (AIForLEssAuto)
https://www.helsinki.fi/en/researchgroups/spatiotemporal-data-analysis/research/aiforlessauto

Images from the first article

GROKE

Groke Technologies Oy is a Finnish company that focuses on developing intelligent methods for autonomous navigation to improve maritime safety.

CHALLENGE

How do you create an awareness system that monitors the environment and detects and recognizes objects around a vessel, for example from the side of another vessel to a far-away sea buoy? How can machine learning models be programmed for object detection with an optimal trade-off between accuracy and performance.

SOLUTION AND IMPACT

By building scalable data processing pipelines for training, evaluating and optimizing machine learning algorithms, and running large-scale training on a GPU-equipped HPC system, better training speed and better target detection were achieved leading to lower unit costs.

LUMI

Working with LUMI

Working with LUMI

- Main work is done via **scripts**
- **Graphical tools** available for:
  - Code development (Visual Studio Code, Jupyter)
  - Input or output review (QGIS)
- Scripts are sent for calculation as **batch jobs**, predefining the needed computational resources
  - SLURM Workload Manager schedules the jobs
- Important to use **parallel computing**

Pre-installed tools in LUMI

- Deep learning:
  - **PyTorch**
  - Keras/TensorFlow
- GIS:
  - QGIS
  - GRASS-GIS
  - SagaGIS
  - GDAL
  - PDAL
  - **Geoconda** – Python package for spatial data analysis
- https://docs.lumi-supercomputer.eu/software/
- https://docs.lumi-supercomputer.eu/software/local/csc/

GIS tools suitable for a supercomputer
(available on CSC national Puhti supercomputer)

LUMI

- **R GIS packages**
- ArcGIS Python API (not arcpy)
- CloudCompare
- FORCE
- LasTools
- OpenDroneMap
- Orfeo Toolbox
- PCL
- SNAP, sen2cor, sen2mosaic
- WhiteboxTools

Open source tools available for **Linux**

Running on: CPU or AMD GPU

GIS tools NOT suitable for supercomputers

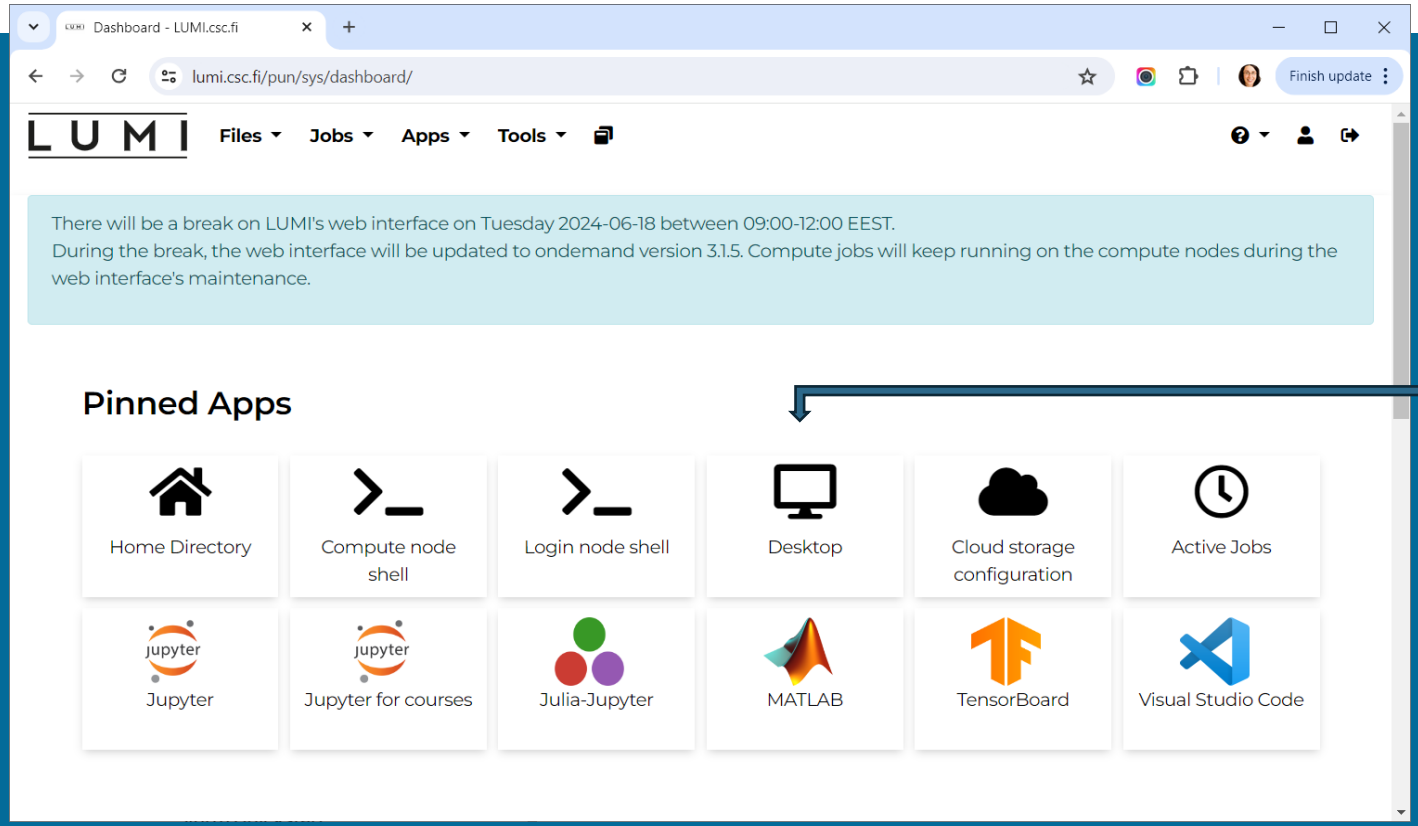
LUMI

- **Windows only tools**
  - ArcGIS, TerraScan
- **APIs**
  - GeoServer, MapServer
- **Databases**
  - PostGIS, MongoDB
- **Web-map front-end**
  - OpenLayers, Leaflet

Own installations

- Tools for installation:
  - **LUMI container wrapper**
    - **Conda, pip or Docker based installations**
  - Apptainer containers
  - EasyBuild
  - SPACK
- Add Python packages to existing installations
  - https://docs.lumi-supercomputer.eu/software/installing/python/

LUMI web interface



QGIS, GRASS, SagaGIS

Steps of a typical use case

1. Get access and user accounts
2. Log-in to LUMI
3. Move data and scripts to LUMI
4. Check your tools
   - Are they already pre-installed?
   - Or install them yourself
5. Create batch job file
6. Run your analysis as batch job
7. Test, debug and analyse the performance of your application
8. Iterate previous steps, if needed

Parallel computing

Python libraries

- Built-in parallel: **xarray**, xarray-spatial, dask-geopandas
- **dask**, multiprocessing, joblib
- https://docs.csc.fi/support/tutorials/dask-python/

R libraries

- Built-in parallel: **terra**, **lidR**, gdalcubes
- **future**, (snow, foreach, Rmpi, …)
- https://docs.csc.fi/support/tutorials/parallel-r/

Parallelization tools

- **GNU parallel** for command line software
- Workflow tools: **Snakemake, NewtFlow…**
- https://docs.csc.fi/computing/running/throughput/

Data in LUMI

- Besides supercomputer disk, LUMI-O provides S3 storage
- For moving data, use the usual Linux data moving tools
  - scp, rsync, S3 tools
  - Small datasets via LUMI web interface
- For Finnish raster data use Paituli STAC
  - https://paituli.csc.fi/stac.html
  - https://radiantearth.github.io/stac-browser/#/external/paituli.csc.fi/geoserver/ogc/stac/v1?.language=en
  - Many of these datasets are stored in the same Kajaani data center

What skills are needed?

LUMI

- Scripting: Python, R, bash …
- Basic Linux
- Using remote service
- Parallelisation (speed-up)
- Resource management

Working with LUMI – Limitations to pay attention

LUMI service promise is not ”24/7”

- LUMI is a highly advanced, multi-purpose ”scientific instrument” that is mainly purposed to support research & development.
  - LUMI may not be able to fulfil the most demanding service quality requirements for the continuous production runs.

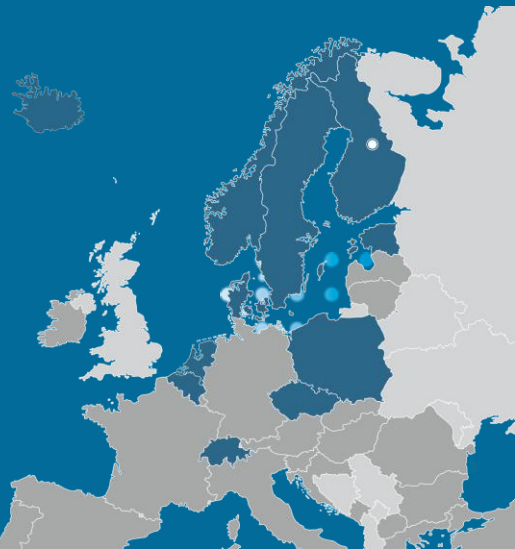
LUMI does not support processing of sensitive personal data

- At the moment LUMI have not the necessary security solutions in place to support processing of sensitive personal data (according to GDPR).

How to start using LUMI in a COMPANY?

Access to LUMI

- **EuroHPC JU** allocates 50% based on project applications
  - Free of charge, results must be public
  - Acedemia, companies, public authorities
- **Consortium counties** allocate 50% based on own rules
  - Finland: Free of charge projects for universities and research institutes
- Project applies for CPU, GPU and data storage resources
- Counted in CPU-core-hours, GPU-hours, TB-hours
- **https://www.lumi-supercomputer.eu/get-started**



Up to 20%

of LUMIs capacity is reserved for European industry and SMEs

LUMI

For companies LUMI offer

- A world-class **supercomputing capacity with cost-efficient pricing**
- **Superfast product development** and **new business opportunities** in areas such as Data analytics and AI
- **Top technical expert and training support**
- A way to initiate or strengthen **cooperation with universities and research institutes**
- **Data security** based on ISO/IEC 27001 standard

LUMI

Private–Public engagement	Pay per use model	Business Finland funding
• Project in cooperation with Finnish university or research organization (academic partner) • Project lead (PI) assigned from academic partner • Free of charge if results are published	• **National LUMI capacity** Company pays market price to CSC • **EuroHPC JU capacity** Company pays market price to JU. Funding possible through PRACE SHAPE and EuroHPC JU programs for SMEs	• **Start-up and SMEs** can request HPC grant at a value of 20,000-80,000 €. Can be added to an already running project. • **Large and mid cap** companies can include computing capacity into their R&D project budget, 40 % of costs covered • Capacity is valued at market price

Test Before Use (“Try&Buy”)

Do you want to develop your company's business faster and more efficiently? Why don´t you try high-performance computing, data analytics or artificial intelligence to support research and development work? The LUMI supercomputer can now be used by your company.

FREE-OF-CHARGE ”TEST BEFORE USE”-PROJECT

In the project companies will receive

- CPU-, GPU, and data storage resources
- our expertise support

If interested to take LUMI in use, please contact our experts!

Price list – LUMI Computing Services

Service	€, VAT 0%
LUMI computing project base package	1000 €
LUMI-C – computing nodes with CPU-processors (AMD Milan) • 1 CPU-node hour equals to 128 CPU-core hours	0,57 € / CPU-node-hour ~0,445 cents / CPU-core-hour
LUMI-G – GPU graphics processing units (AMD MI250)	0,535 € / GPU-hour
LUMI-P – Lustre parallel file system	0,005 € / TiB/h
LUMI-F – Flash memory storage	0,05 € / TiB/h
LUMI-O – LUMI-O – CEPH object storage	0,0025 € / TiB/h

Working with LUMI – Principles

LUMI

LUMI Projects and User Accounts

- The use of LUMI is based on LUMI projects and personal user accounts.
- Each project has a Principal Investigator (PI) who is the lead contact and the responsible person for the project.
- A commercial LUMI project and user accounts for commercial company persons are requested via CSC ServiceDesk servicedesk@csc.fi

Running your jobs with LUMI Resources

- LUMI projects have CPU / GPU / data storage resources allocated according to your contract made with CSC.
- Running your application, i.e. a job, is based on batch jobs.
  - Your job is put to a queue to wait for requested resources, and run when they are available.
  - Note: **You pay for the used resources only**, not for service in general, like in monthly based cloud services.

Next steps

How to get help / Further info

- CSC ServiceDesk: **servicedesk@csc.fi**
- CSC LUMI User Support https://lumi-supercomputer.eu/user-support/need-help/
- CSC LUMI documentation https://docs.lumi-supercomputer.eu/
- CSC training: https://www.csc.fi/en/training#training-calendar
- LUMI training: https://www.lumi-supercomputer.eu/events/
- CSC User guides and tutorials https://docs.csc.fi/
- Geocomputing with supercomputer: https://research.csc.fi/geocomputing
- EuroCC-Finland LinkedIn: www.linkedin.com/company/eurocc-finland
- HPC GIS/remote sensing use cases: https://docs.csc.fi/support/training-material/geocomputing-seminars/

Extra:

- How does accessing a supercomputer look like?
    - https://www.youtube.com/watch?v=HB9JUH0mPqI (by EuroCC Sweden)

Free training and examples

- 8-9.10.2025 Moving your AI training jobs to LUMI: A Hands-On Workshop (Stockholm / online)
  - https://www.lumi-supercomputer.eu/events/moving-your-ai-training-jobs-to-lumi-a-hands-on-workshop/
- **8.–9.10.2025 Geocomputing course (Espoo / online)**
  - https://csc.fi/koulutuskalenteri/geocomputing-on-the-supercomputer-2/
- 20.-21.10.2025, LUMI Intro Course (Tallinn / online)
  - https://www.lumi-supercomputer.eu/events/lumi-intro-course-tallinn/
- **10-12.12.2025 Practical machine learning with spatial data (Espoo / online)**
  - https://csc.fi/koulutuskalenteri/practical-machine-learning-with-spatial-data-2/

- Geocomputing examples: https://github.com/csc-training/geocomputing
- Previous training materials: https://research.csc.fi/gis-learning-materials

Summary

- LUMI is huge supercomputer, with a lot of GPU resources
- LULMI enables data analysis that would otherwise be impossible or very time consuming
- Supercomputers excel in Deep Learning tasks
- LUMI provides cost-effective access to GPUs

It takes a bit of practice,
but we are here to help:

firstname.lastname@csc.fi

Katri Tegel

Kylli Ek

Jyrki Savolainen