

CoE RAISE AHM in Iceland

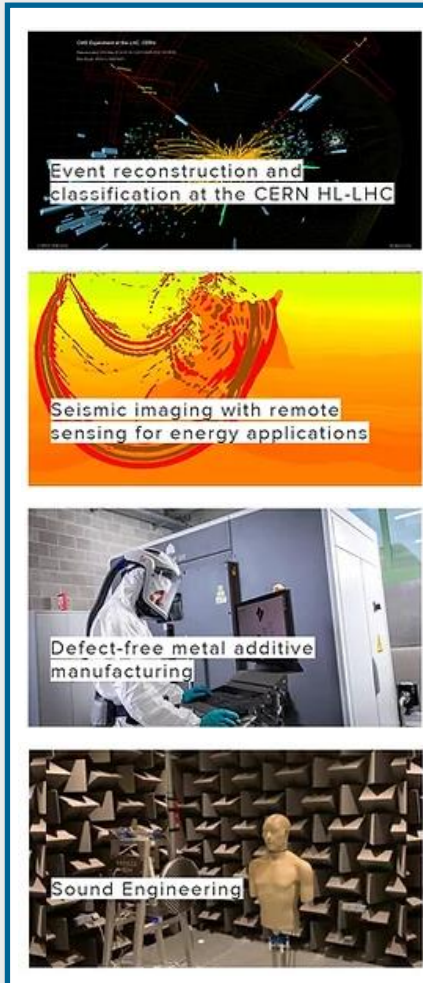
29 August 2023

WP4 "Data-Driven Use-Cases at Exascale": Status and Plans

Maria Girone, Eric Wulff, CERN

Objectives of WP4

- WP4 aims at integrating, developing, and scaling AI and HPC methods with a strong focus on data-driven workflows
 - Representative use-cases from academia and industry
 - Demonstrate the ability to analyze data-rich descriptions of physical phenomena
 - **Data challenge under way!**
- The use-cases have a wide application range from science and industry with specific targeted goals
 - Generate essential information of the underlying dynamics in multidisciplinary systems
 - Find new energy sources and improve the pipeline in storage and supply
 - Identify unknown risks and the associated weakness in automation systems
 - Obtain higher productivity, robust quality control, and cost-efficiency in manufacturing

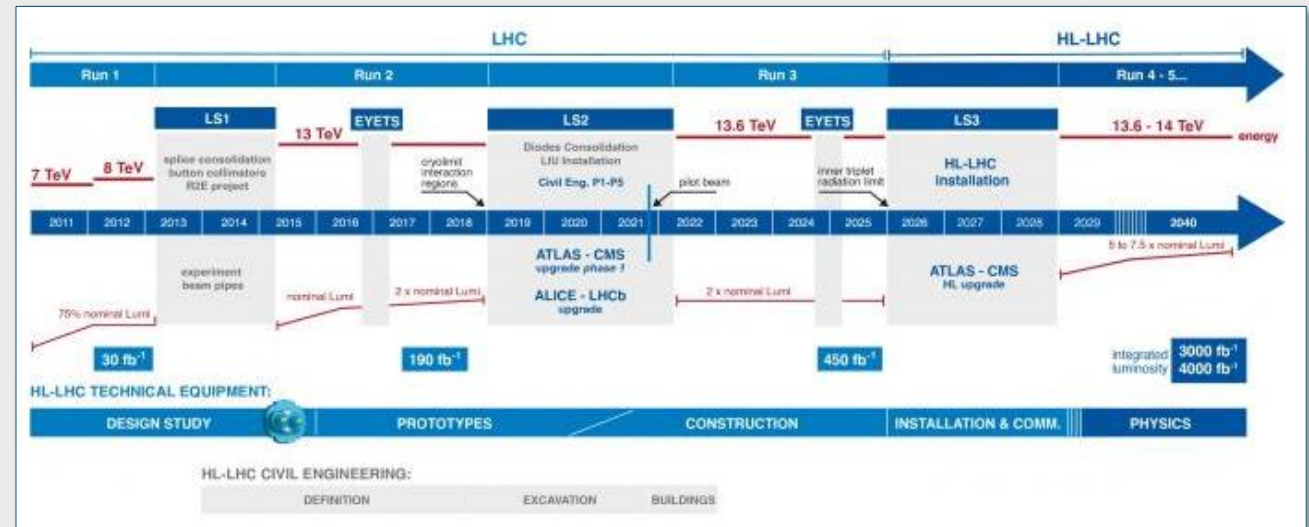
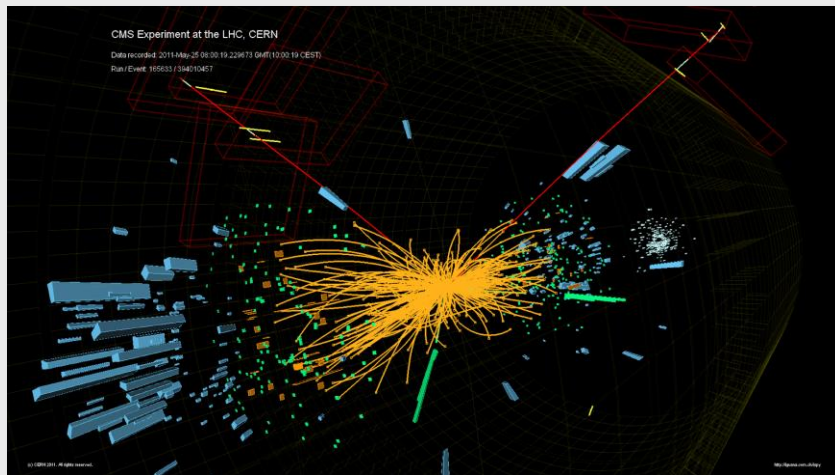


The way we work

- Regular WP4 meetings with minutes ever 2 weeks
- Very productive cooperation with other WPs, especially WP2 and WP6
- (Still) very enthusiastic and highly competent team. Thank you!!!!
- All tasks will present their status shortly

Task 4.1 – Event Reconstruction and classification at the CERN HL-LHC

- At the upgraded HL-LHC programme, each experiment will produce exabytes yearly, which will result in an unprecedented computing challenge



- Develop AI and exascale solutions for event reconstruction and classification
 - AI-based particle-flow reconstruction algorithms
 - Developing and generating datasets for supervised learning
 - Development, training and HPO of deep learning models
 - Development of algorithms for hybrid quantum-classical hyperparameter optimization
 - Development and optimization of a GPU-accelerated clustering algorithm

Task 4.2 – Seismic imaging with remote sensing for energy applications

Remote sensing

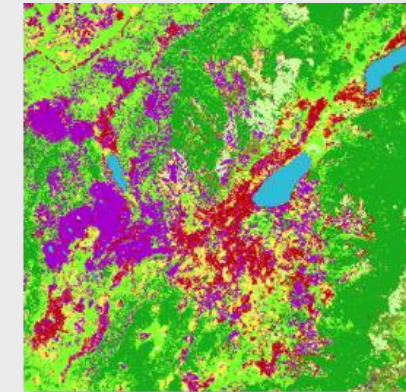
- for land-cover classification and prediction
- develop AI methods to automatically identify land use from satellite data [3]



Data is freely available (Sentinel-2 data)



True-color composition of a Sentinel-2 image (2019)

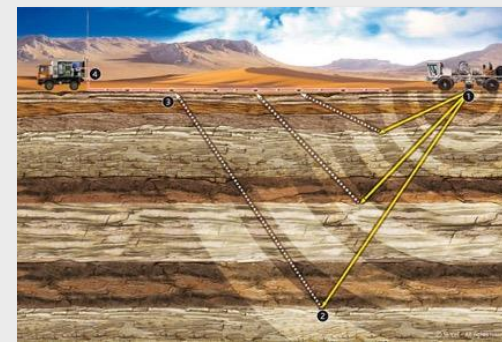


LSTM land cover classification map

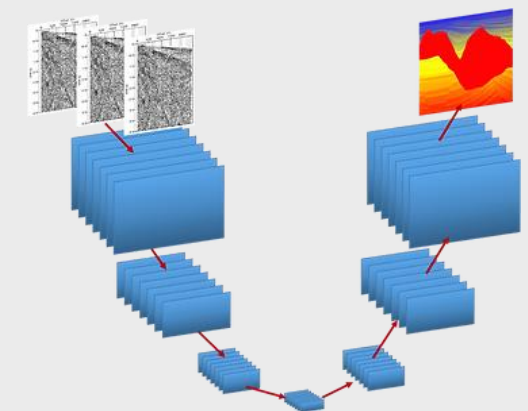


Seismic imaging

- seismic imaging of subsurface structures involves solving complex large-scale inverse problems
- replace computationally expensive parts of seismic imaging with a ML based approach



Land Seismic Survey



Schematic of a DL NN approach to assist seismic inversion

Working on the convergence of the two sciences

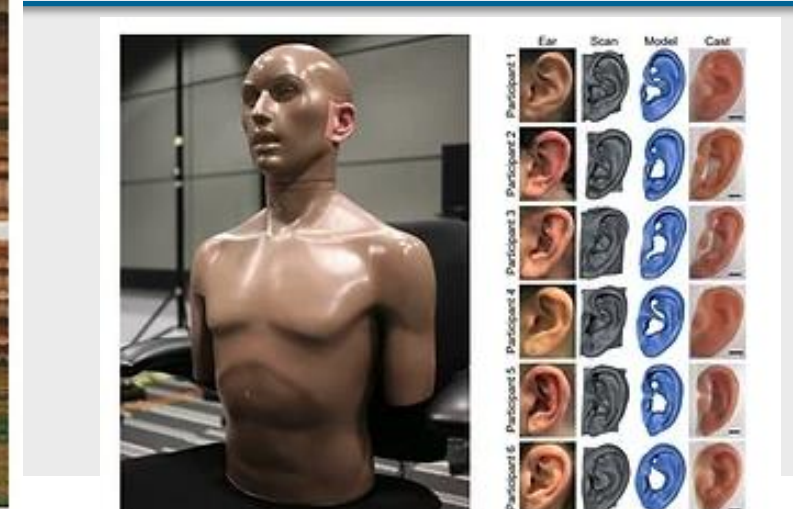
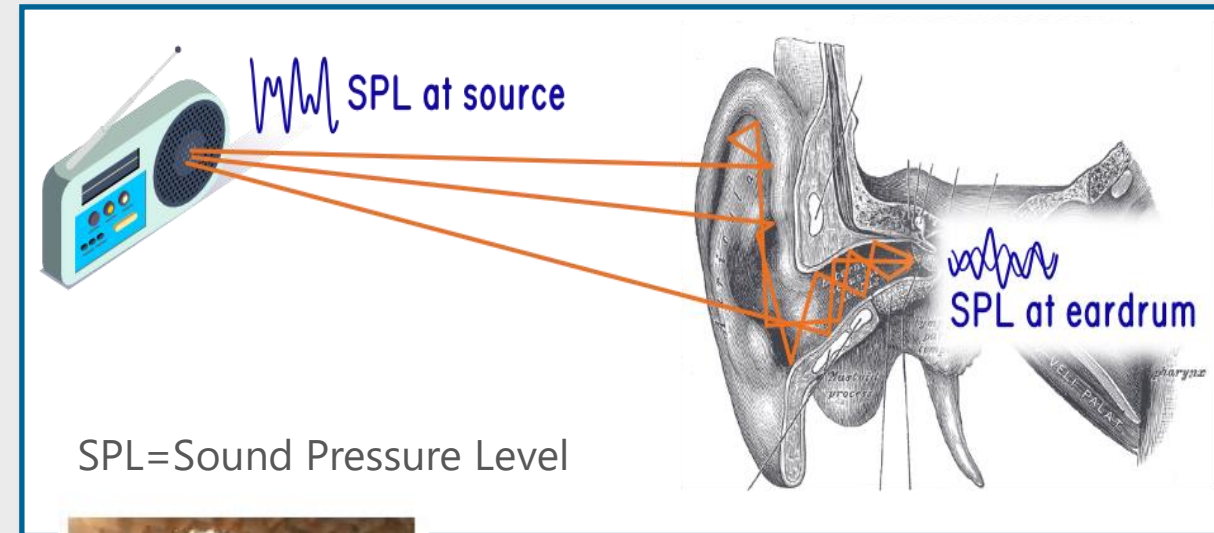
Task 4.3 – Defect-free metal additive manufacturing

- Laser powder bed fusion (LPBF) is the main industrial 3D printing process for metal parts
- Major bottleneck: quality control
- Process monitoring with high-speed video
- Train computer vision models
 - Reconstruct laser params
 - Direct porosity prediction



Task 4.4 – Sound Engineering

- Spatial audio systems are algorithms designed to stimulate sound localization
 - applications in VR, audiology, safety equipment
- The operation of spatial audio systems relies on the listener's Head-Related Transfer Function (HRTF)
- The human brain learns to use HRTF to determine localization
- The goal is to use DL on HPC to produce individualized HRTF estimates



Partners and Tasks of WP4

Partner	FZJ	UOI	CYI	CERN	RTU	FM
PM	10	21	18	29	6	22

Task	Title	Lead	Duration	Status
4.1	Event reconstruction and classification at the CERN HL-LHC	CERN	M1 – M36	Ongoing
4.2	Seismic imaging with remote sensing for energy applications	CYI	M1 – M36	Ongoing
4.3	Defect-free metal additive manufacturing	FM	M1 – M36	Ongoing
4.4	Sound engineering	UOI	M1 – M36	Ongoing

Deliverables and Milestones of WP4

ID	Title	Due	Lead	Status
D4.1	Report on outcomes of WP4 use-cases	M12	CERN	Submitted
D4.2	Report on outcomes of WP4 use-cases	M24	CERN	Submitted
D4.3	Report on outcomes of WP4 use-cases	M36	CERN	Not started

ID	Title	Due	Status
MS1	Project kick-off	M1	Achieved
MS3	Use-cases/technical developments	M24	Achieved
MS5	All final reports	M36	Not yet achieved

➤ Together with WP2, the WP4 use cases have 6 scientific papers in preparation/published.

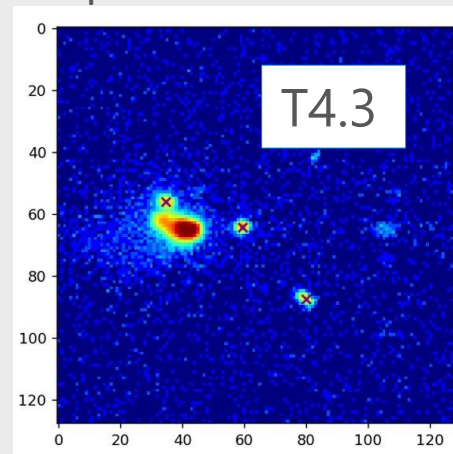
Status of WP4 in 2023 – Reports and Deliverables (1/2)

- All quarterly reports delivered on time
- WP4 actively participates in WP2 and WP6
 - News articles
 - AHM at CERN (<https://www.coe-raise.eu/news-2023-01-ahm>)
 - Eric and David presented work from WP4 at the CERN openlab Technical Workshop (<https://www.coe-raise.eu/news-2023-03>)
 - Success story (https://www.coe-raise.eu/files/ugd/248388_dd20ddf9eee1414ea186164736597147.pdf)
 - In collab with WP2
 - T4.1: Poster at ISC, presented by Marcel Aach (<https://juser.fz-juelich.de/record/1007703>)
 - T4.1: Abstract submitted to the Quantum Technologies in Machine Learning conference (QTML), taking place at CERN

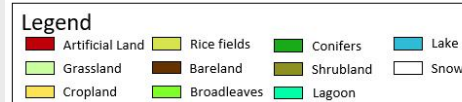
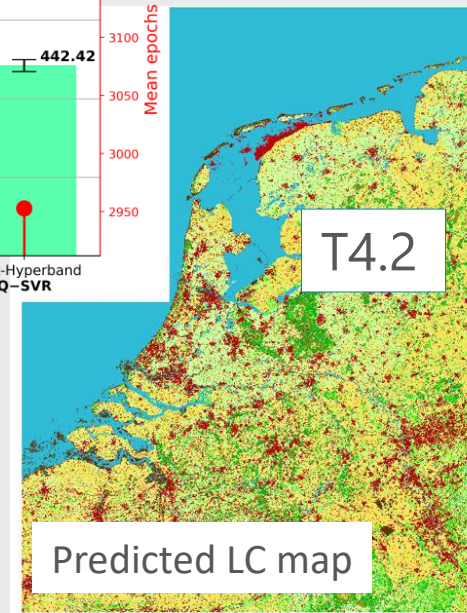
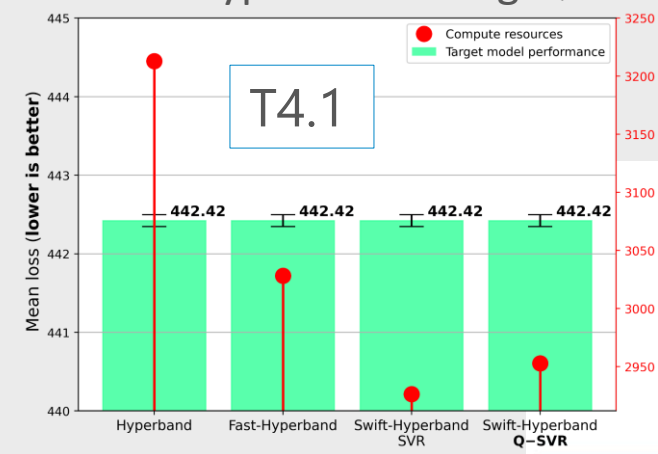
Status of WP4 in 2023 – Reports and Deliverables (TOBEUPDATED)

- Common objectives
 - Advance the SOTA of AI/ML in the different fields
 - Enable more efficient usage of HPC infrastructure on the path to Exascale
 - Enable more use-cases to leverage heterogeneous resources
- Synergies
 - Data is central to all use-cases
 - Common need for using heterogeneous hardware
 - All use-cases benefit from advanced hyperparameter optimization
 - All use-cases aim to increase speed and/or accuracy of various tasks using data-driven AI-based approaches

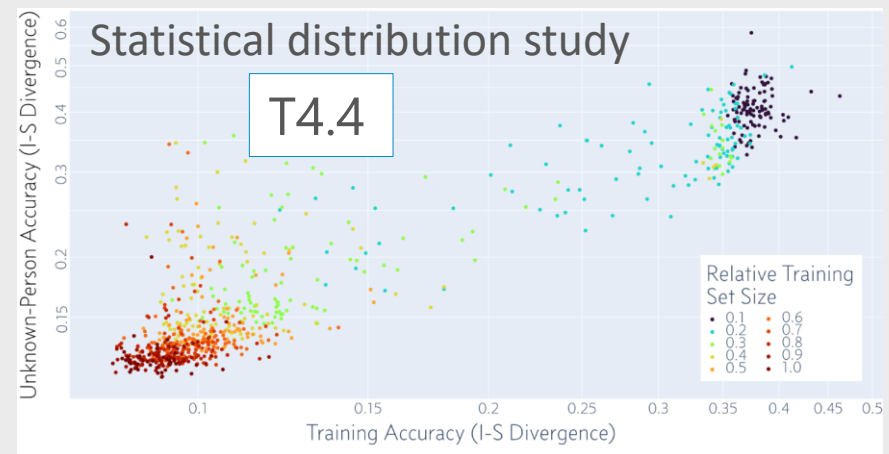
Spatter detection



Swift-Hyperband using QSVR



Statistical distribution study



Data Challenge in a Nutshell

As we move to larger datasets, transferring data in a predictable, practical and reliable manner is still something that's a challenge. We can provide this expertise from RAISE for future AI workflows:

Verify open-source transfer tools

- UFTP transfers from RTU, FZJ, GÉANT. ---PREPARING TOOLING

Verify high-speed site transfers

- CERN – FZJ 200G ---CONNECTIONS ESTABLISHED, IN PROGRESS

Summer student (CERN) working on data transfer testing with FZJ, to be completed soon!

Many thanks to all experts for supporting these tests!

- WP4 is invested in exploring synergies and sharing knowledge within all RAISE WPs
 - AI/HPC technologies (HPO, GNNs, autoencoders, FNO, GANs, Transformers, CNNs, (Q)SVR)
 - use of HPC infrastructures (i.e., JURECA, JUWELS, DEEP, JUPSI, ..)
- In our last year of CoE RAISE, WP4 is working to enhance synergies and common solutions with close communities/use-cases, on the path to Exascale.
 - Exploring synergies with other EC-funded projects (interTwin, SPECTRUM)
 - Data challenges

drive. enable. innovate.



The CoE RAISE project has received funding from the European Union's Horizon 2020 – Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733

Follow us:



R^G