UNIVERSITY OF ICELAND
SCHOOL OF ENGINEERING AND NATURAL SCIENCES
FACULTY OF INDUSTRIAL ENGINEERING, MECHANICAL ENGINEERING AND COMPUTER SCIENCE

RAISE Center of Excellence

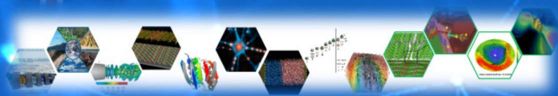
EURO

CoE RAISE – Need for Distributed Deep Learning

Prof. Dr. – Ing. Morris Riedel
School of Engineering & Natural Sciences, University of Iceland, Iceland
Juelich Supercomputing Centre, Forschungszentrum Juelich, Germany

2021-07-29, RAISE CoE Seminar Distributed Deep Learning, Online

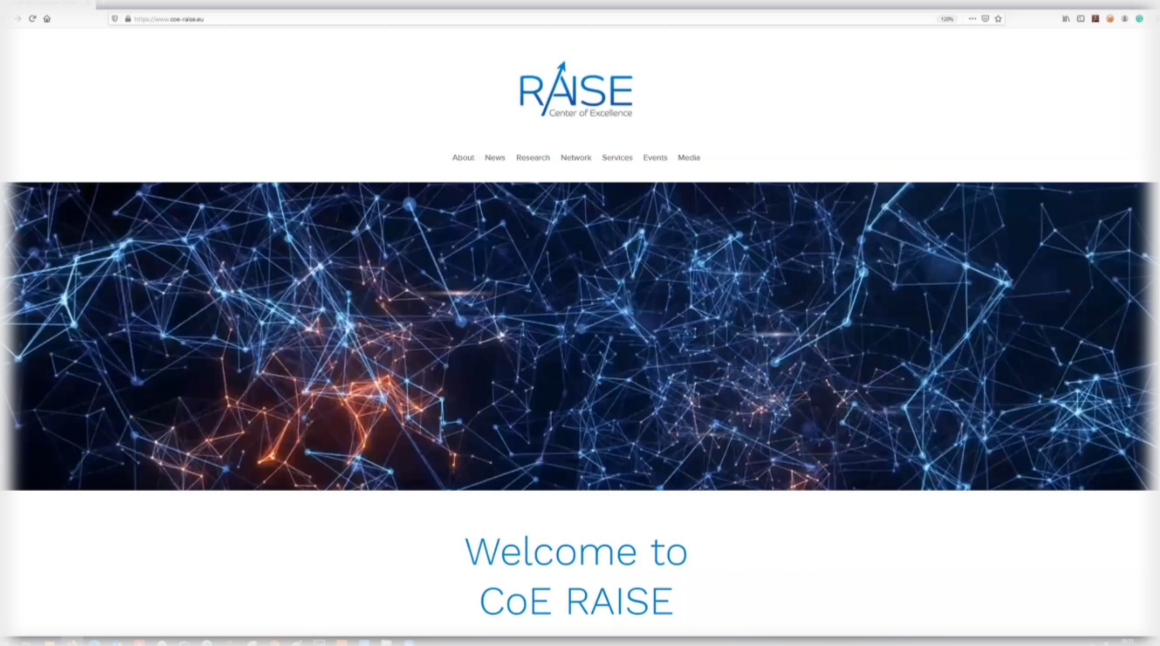
IHPC National Competence Center (NCC) for HPC & AI in Iceland



@ProfDrMorrisRiedel **@Morris Riedel** **@MorrisRiedel** **@MorrisRiedel**
https://www.youtube.com/channel/UCWC4VKHmL4NZgFfKoHtANKg **morris@hi.is**

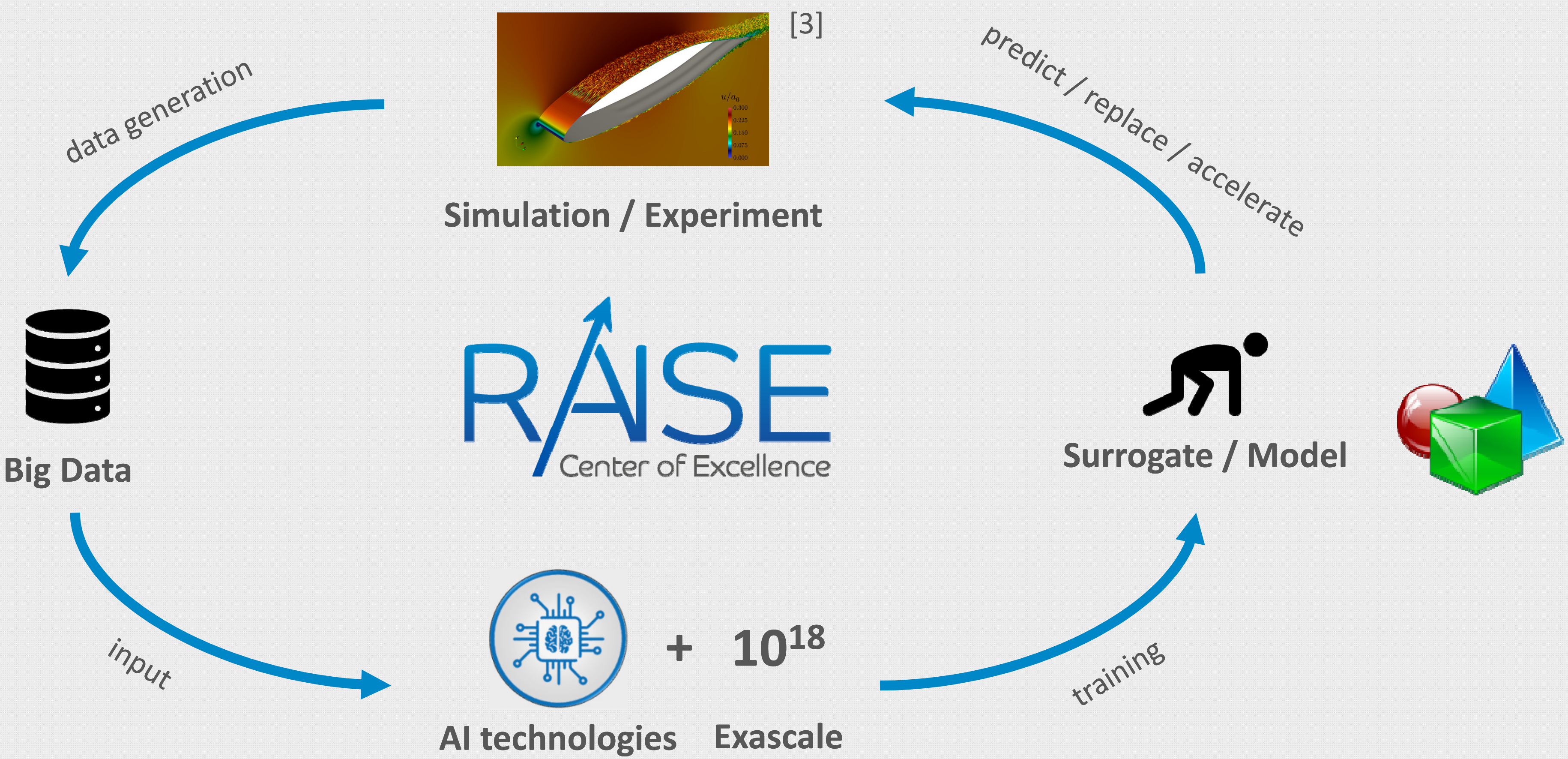
CoE RAISE Web Page & More Information

RAISE
Center of Excellence



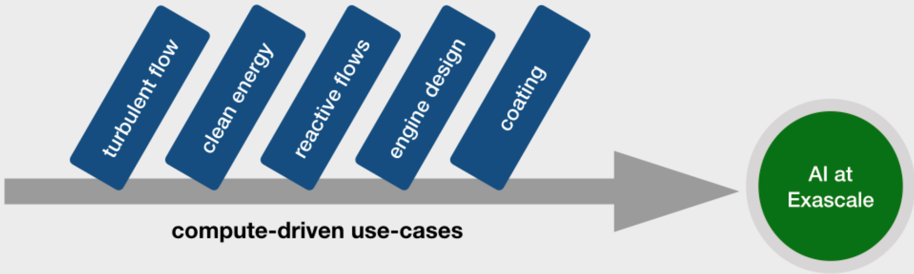
https://www.coe-raise.eu

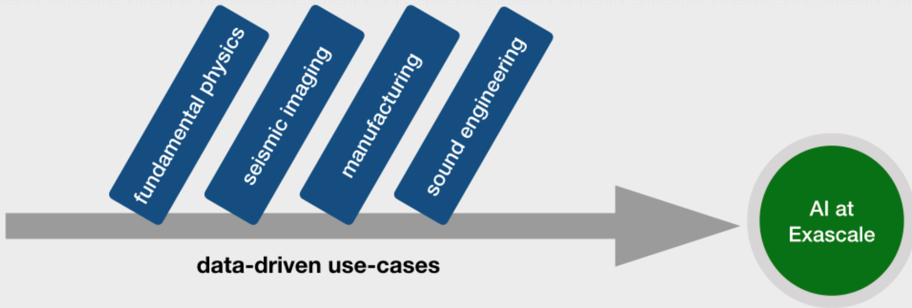
CoE RAISE – Motivation & Approach

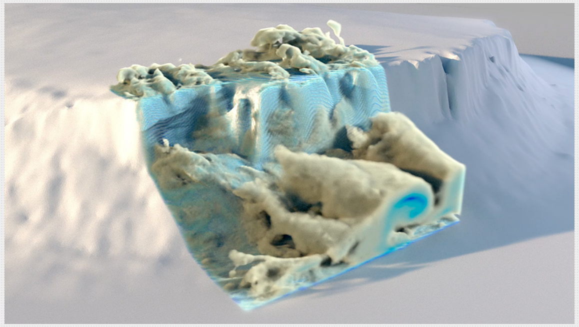


Use Cases in CoE RAISE

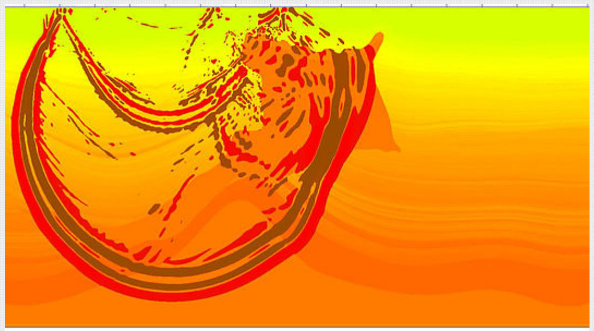
- Two kinds of use cases:





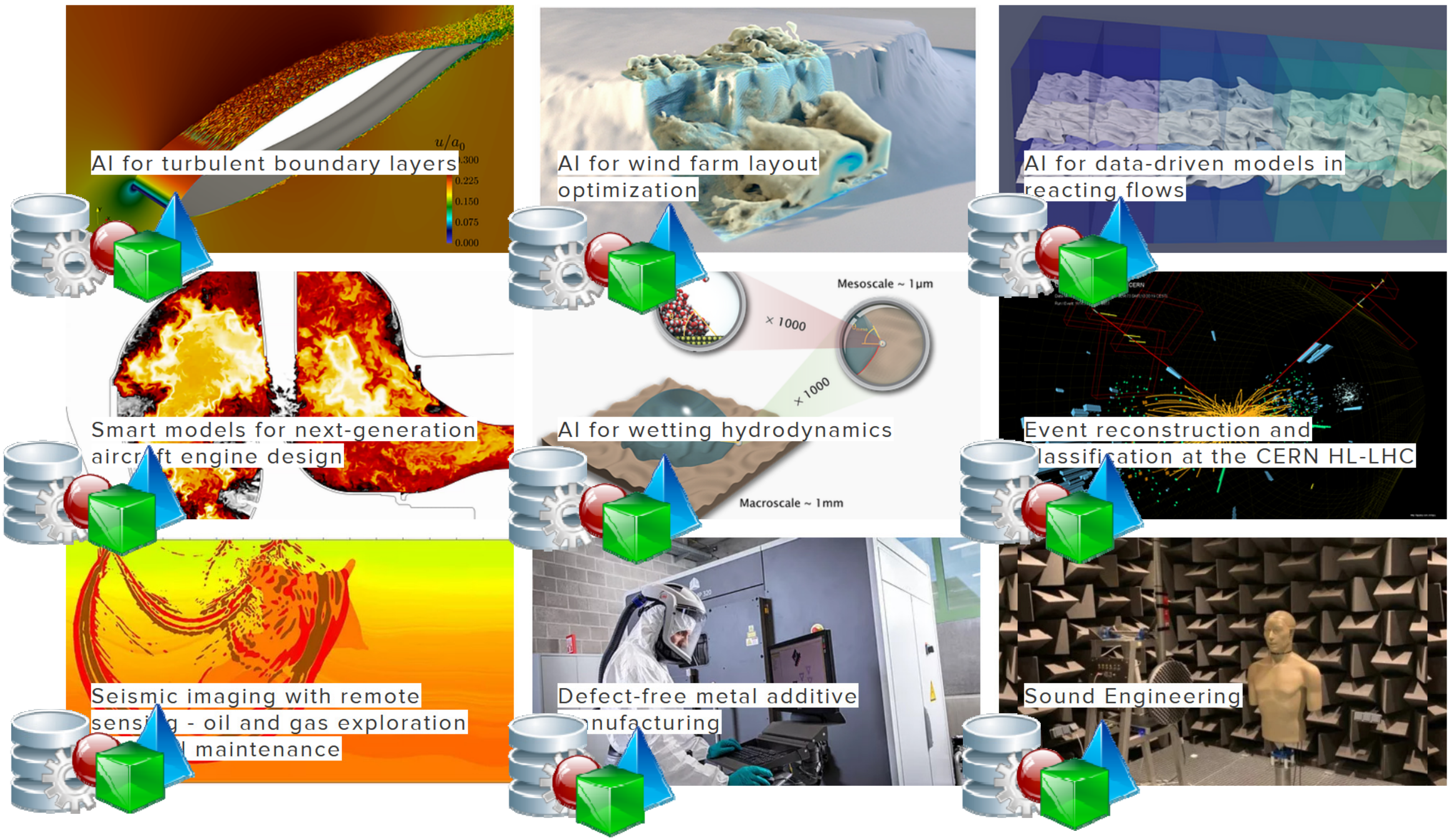


Example from use case "AI for wind farm layout": Turbulence generated by a cliff on Bolund Island, Denmark.



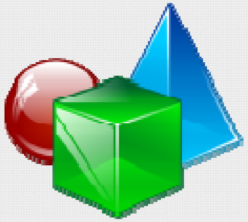
Example from use case " Seismic imaging with remote sensing - oil and gas exploration and well maintenance": Snapshot from a wavefield.

Compute- and Data-driven Use Cases – Use & Generate Data

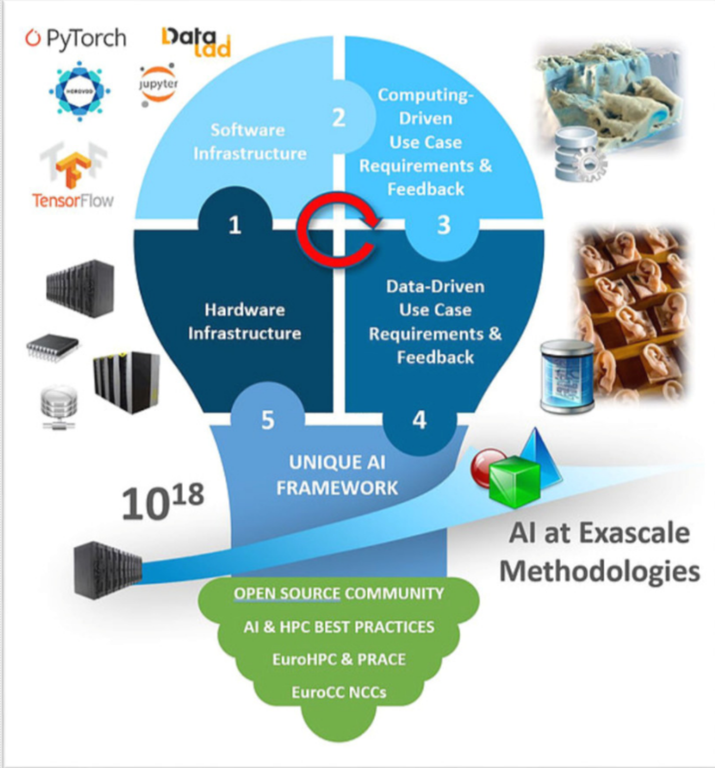


CoE RAISE's Objectives

- Development of AI methods towards Exascale along use-cases
- RAISE tightly connects
  - an exceptional hardware infrastructure,
  - an usable and versatile software infrastructure,
  - compute-driven use cases,
  - and data-driven use cases

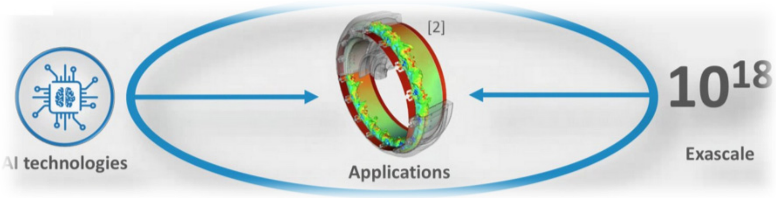


to contribute to a Unique AI framework that will be provided to academic and industrial communities (RAISE AI-Exascale library)

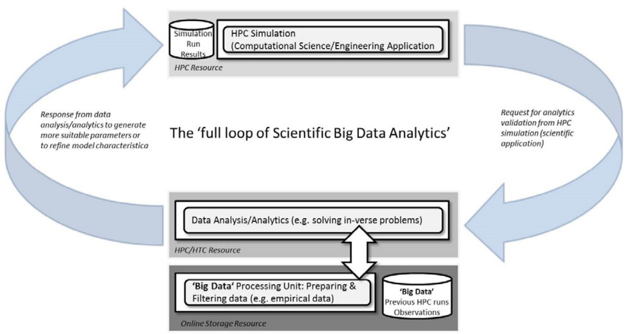


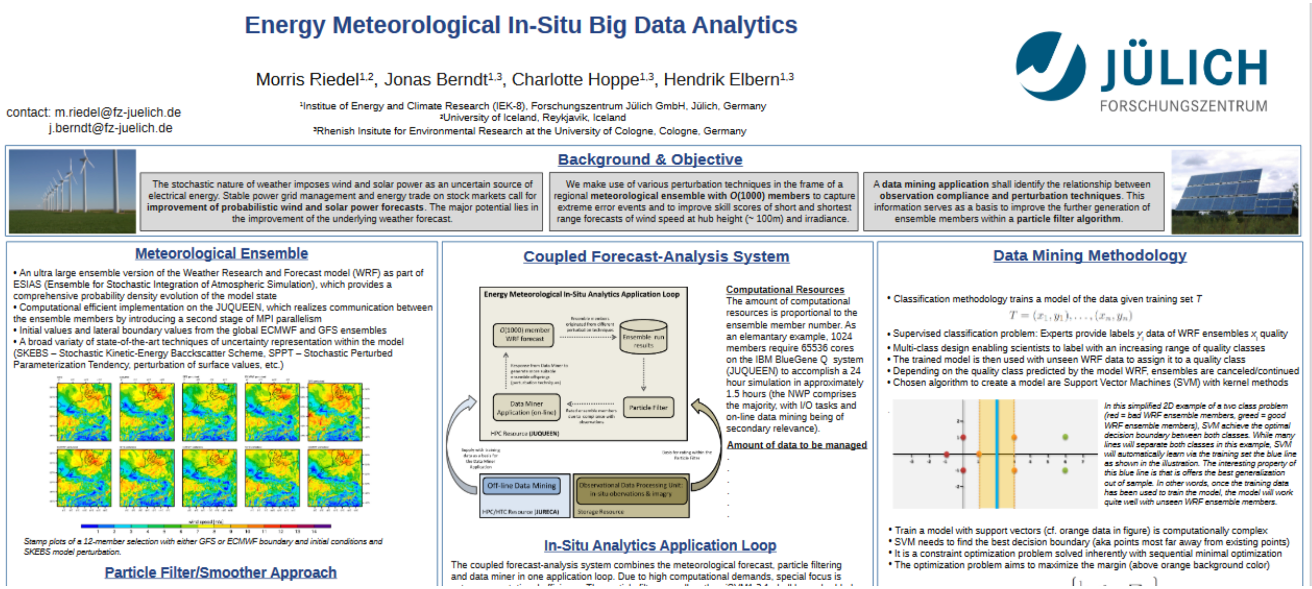
Vision – Intertwined HPC Simulations & AI – ‘full loop‘ ?

- What means AI & HPC Cross Methods at Exascale?



Today rather high performance data analytics (HPDA)





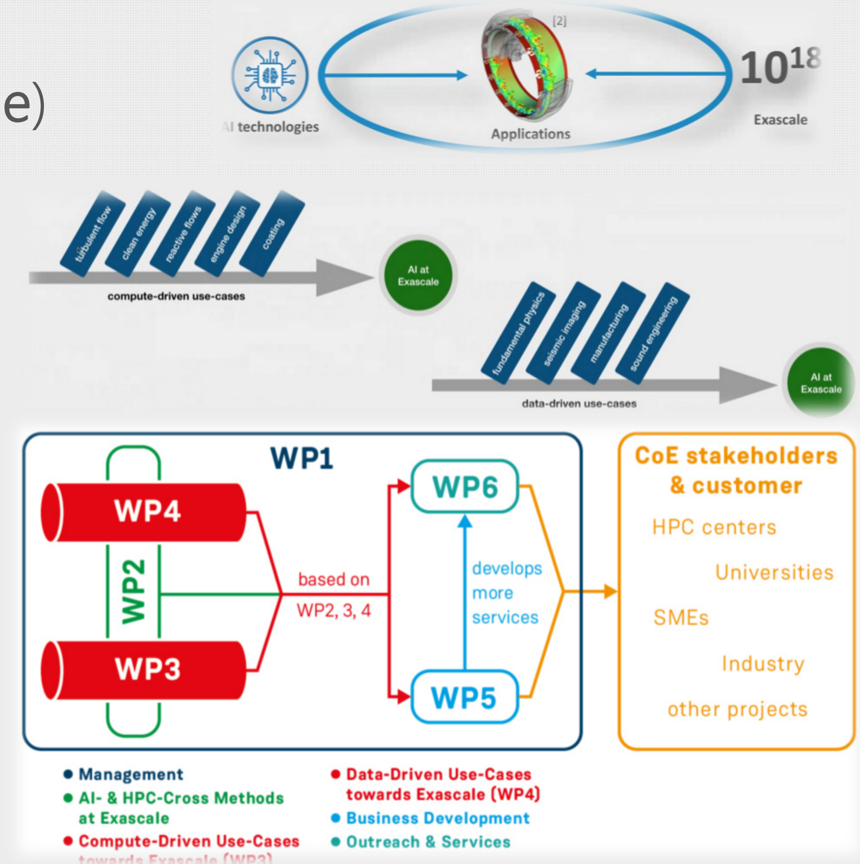
Lippert, T., Mallmann, D., **Riedel, M.**: Scientific Big Data Analytics by HPC, in Symposium proceedings of NIC Symposium 2016 – publication Series of the John von Neumann Institute for Computing (NIC), NIC Series 48 (417), ISBN 978-3-95806-109-5, February 11-12, 2016, Juelich, Germany

Riedel, M., Berndt, J., Hoppe, C., Elbern, H., Energy Meteorological In-Situ Big Data Analytics, Helmholtz Program Meeting, Karlsruhe Institute of Technology (KIT), July 1, 2016, Karlsruhe, Germany, [PDF (~ 4,08 MB)]

WP2 – AI- & HPC-Cross Methods at Exascale in a nutshell

RAISE Center of Excellence

- WP3 (Compute-Driven Use-Cases towards Exascale)
- WP4 (Data-Driven Use-Cases towards Exascale
- Developments in these WPs will be supported by the cross-linking activities of WP2
  - E.g. scaling machine & deep learning codes with frameworks like Horovod/Deepspeed
  - E.g. introduction to new AI methods such as Long-Short Term Memory (Time series)
  - E.g. data augmentation approaches
  - E.g. benchmarking HPC machines and offer also pre-trained AI algorithms (i.e., transfer learning)
  - E.g. offer neural architecture search methods for hyperparameter – tuning in semi-automatic way



Towards AI & HPC at Exascale with CoE RAISE Results

RAISE Center of Excellence

Hardware Infrastructure

Prepare & Document available production systems at partners' HPC centers

Examples: JUWELS (JUELICH), LUMI (UoICELAND), DEEP Modular Prototypes, JUNIQ (JUELICH), etc.

Software Infrastructure

Prepare & Document available open source tools & libraries for HPC & AI useful for implementing use cases

Examples: DeepSpeed and/or Horovod for interconnecting N GPUs for a scalable deep learning jobs

Computing-driven Use Cases Requirements & Feedback

Use cases with emphasize on computing bring in co-design information about AI framework & hardware

Examples: Use feedback that TensorFlow does not work nicely, so WP2 works with use cases on pyTorch

Data-driven Use Cases Requirements & Feedback

Use cases with emphasize on data bring in co-design information about AI framework & hardware

Examples: Deployment blueprint by using AI training on cluster module & inference/testing on booster

→ UNIQUE AI FRAMEWORK

Living design document & software framework blueprint for using HPC & AI offering also pretrained AI models

Example HeAT → CoE RAISE Seminar June: on YouTube soon

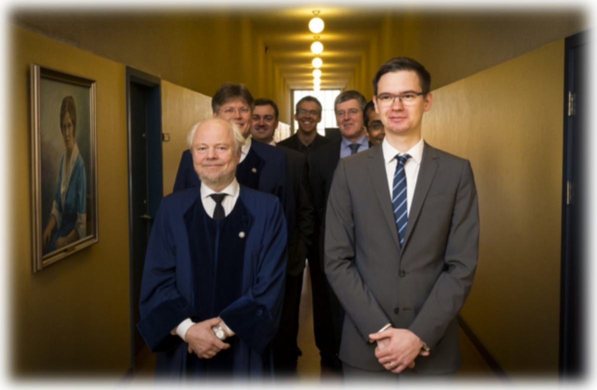
Package	*Multi CPU*	*Single GPU*	*Multi GPU*	*NumPy API*	AD	Ref.
PyTorch	✓	✓	✓	✓	✓[a]	[4]
Legate	✓	✓	✓	✓		[14]
Dask	✓			✓		[5]
Intel DAAL	✓					[16]
TensorFlow	✓	✓	✓	✓	✓[a]	[3]
MXNet	✓	✓	✓	✓		[8]
DeepSpeed	✓	✓	✓	✓		[17]
DistArray	✓					[18]
Bohrium	✓	✓				[11]
Grumpy	✓		✓			[12]
JAX	✓	✓	✓	✓		[19]
Weld	✓		✓			[13]
NumPywren	✓					[20]
Arkouda	✓					[21]
GAiN	✓					[22]
Spartan	✓					[10]
Phylanx	✓					[23]
Ray	✓	✓	✓	✓		[15]
HeAT	✓	✓	✓	✓	✓	

[a]Based on RPC, no MPI support.

HeAT – a Distributed and GPU-accelerated Tensor Framework for Data Analytics

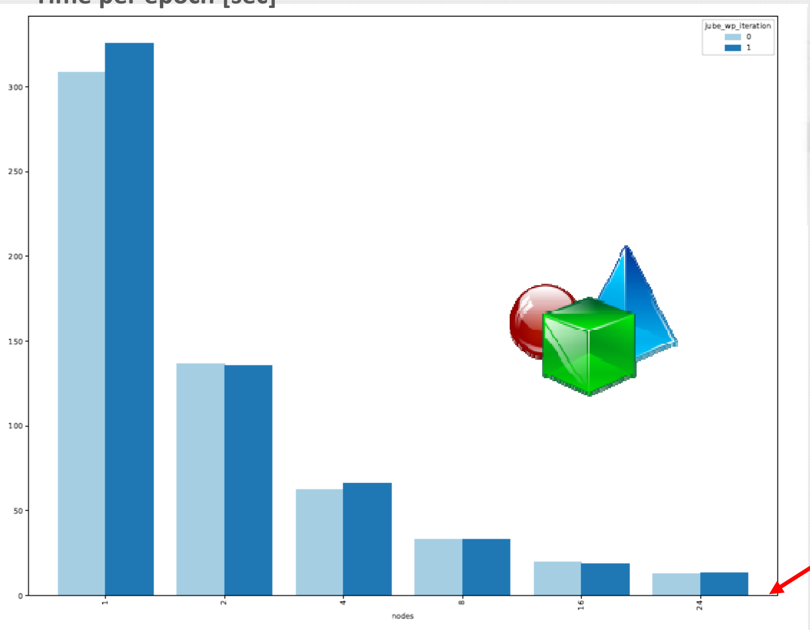
Markus Götz[‡], Daniel Coquelin[†‡], Charlotte Debus[*], Kai Krajsek[†], Claudia Comito[†], Philipp Knechtges[*], Björn Hagemeier[†], Michael Tarnawa[†], Simon Hanselmann[‡], Martin Siggel[*], Achim Basermann[*] and Achim Streit[‡]
[*]*Institute for Software Technology (SC)*
German Aerospace Center (DLR)
Cologne, Germany
{charlotte.debus, philipp.knechtges, martin.siggel, achim.basermann}@dlr.de
[†]*Jülich Supercomputing Centre (JSC)*
Forschungszentrum Jülich (FZJ)
Jülich, Germany
{k.krajsek, c.comito, b.hagemeier, m.tarnawa}@fz-juelich.de
[‡]*Steinbuch Centre for Computing (SCC)*
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
{markus.goetz, daniel.coquelin, simon.hanselmann, achim.streit}@kit.edu

Götz M, Debus C, Coquelin D, Krajsek K, Comito C, Knechtges P, Hagemeier B, Tarnawa M, Hanselmann S, Siggel M, Basermann A. HeAT–a Distributed and GPU-accelerated Tensor Framework for Data Analytics. In2020 IEEE International Conference on Big Data (Big Data) 2020 Dec 10 (pp. 276-287). IEEE.



Can AI do Exascale & use Disruptive Technologies?

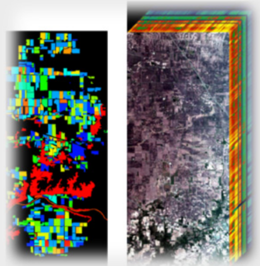
Time per epoch [sec]

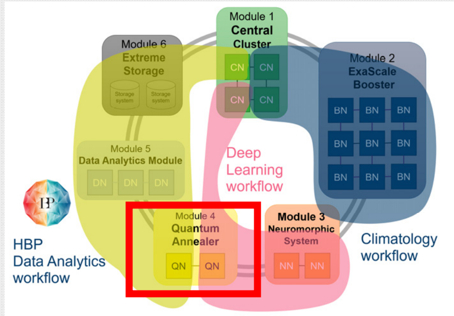


Europe #1 Supercomputer (11/2020)

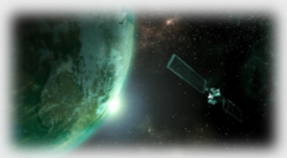
Example from 2019:
Using partition of the JUWELS system has 56 compute nodes, each with 4 NVIDIA V100 GPUs (equipped with 16 GB of memory)

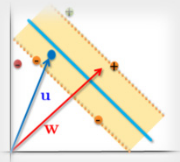
24 nodes x 4 GPUs = 96 GPUs






```
In [ ]: from quantum_SVM import *

        # Hyperparameters
        B=[2,3,5,10]
        K=[2,3]
        xi=[0,1,5]
        gamma=[-1,0.125,0.25,0.5,1,2,4,8]
        n_experiments=len(B)*len(K)*len(xi)*len(gamma)

        hyperparameters=np.zeros([n_experiments,4], dtype=float)

        path_data_key='input_datasets/calibration/'+id_dataset+'/'
        data_key = id_dataset+'calibtrain'
        path_out='outputs/calibration/'+id_dataset+'/'

        trainacc=np.zeros([fold], dtype=float)
        trainauroc=np.zeros([fold], dtype=float)
        trainauprc=np.zeros([fold], dtype=float)
```





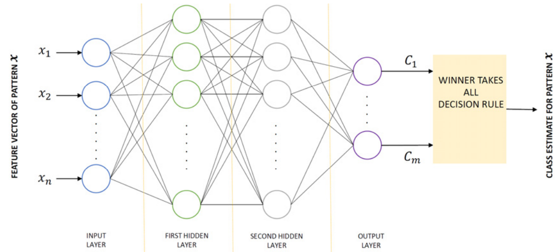
Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., **Riedel, M.**, Book, M.: SCALING UP A MULTISPECTRAL RESNET-50 TO 128 GPUS, in conference proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020), September 26 – October 2nd, 2020, Virtual Conference, Hawai, USA

Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., **Riedel, M.**, Benediktsson, J.A.: Remote Sensing Big Data Classification with High Performance Distributed Deep Learning, Journal of Remote Sensing, Multidisciplinary Digital Publishing Institute (MDPI), Special Issue on Analysis of Big Data in Remote Sensing, 2019

Cavallaro, G., Willsch, D., Willsch, M., Michielsen, K., **Riedel, M.**: APPROACHING REMOTE SENSING IMAGE CLASSIFICATION WITH ENSEMBLES OF SUPPORT VECTOR MACHINES ON THE D-WAVE QUANTUM ANNEALER, in conference proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020), September 26 – October 2nd, 2020, Virtual Conference, Hawai, USA

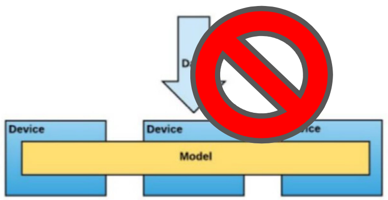
Horovod Example of Distributed Training Tool

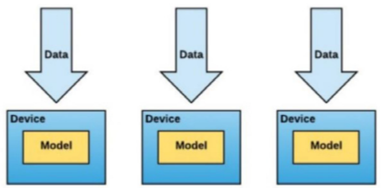
- Free open-source
  AI tool: Horovod
  - https://github.com/horovod/horovod



- Distributed Training of Deep Learning Models
  - Used on HPC systems to speed-up model training
  - Significant experience at the University of Iceland and Juelich Supercomputing Centre (both partners in CoE RAISE)
  - Used in Science & Engineering, e.g. remote sensing image analysis

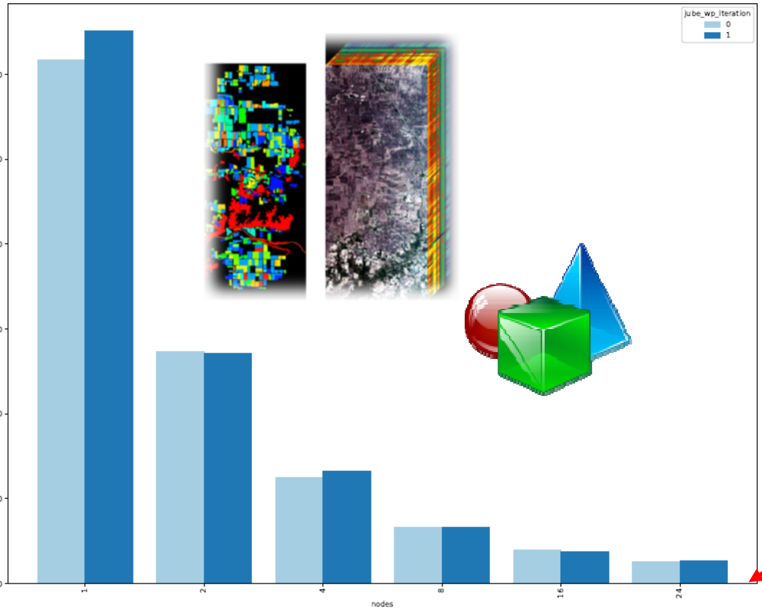
Horovod
- **Data parallel**, each GPU has a copy of the model and a chunk of the data
- Efficient **decentralized framework**,
based on MPI and NCCL libraries, where actors exchange parameters without the need of a parameter server
- Works on top of Keras, TensorFlow, PyTorch and Apache MXNet





Horovod Example: Challenges & Benefits

Time per epoch [sec]



Europe #1 Supercomputer (11/2020)

Example from 2019:
Using partition of the JUWELS system has 56 compute nodes, each with 4 NVIDIA V100 GPUs (equipped with 16 GB of memory)

24 nodes x 4 GPUs = 96 GPUs

Distributed training challenges w.r.t. batch sizes & accuracy

batch size	n. GPUs	training time [s]
512	8	49,400
8,000	128	3,400
16,000	128	2,800
32,000	128	2,500

batch size	n. GPUs	warm-up	initial LR	F1
512	8	5	0.2	0.78
8,000	128	5	3.2	0.74
16,000	128	5	6.4	0.64 (diverge)
32,000	128	5	12.8	0.43 (diverge)

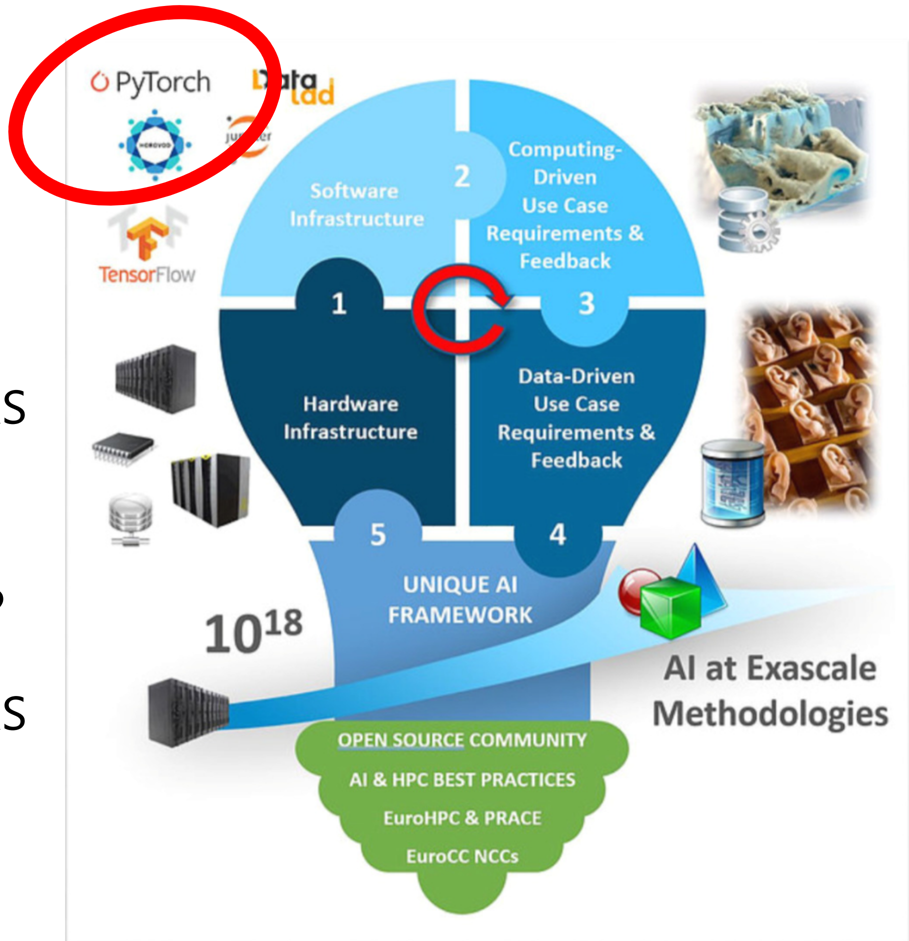
Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., **Riedel, M.**, Book, M.: SCALING UP A MULTISPECTRAL RESNET-50 TO 128 GPUS, in conference proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020), September 26 – October 2nd, 2020, Virtual Conference, Hawai, USA

Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., **Riedel, M.**, Benediktsson, J.A.: Remote Sensing Big Data Classification with High Performance Distributed Deep Learning, Journal of Remote Sensing, Multidisciplinary Digital Publishing Institute (MDPI), Special Issue on Analysis of Big Data in Remote Sensing, 2019

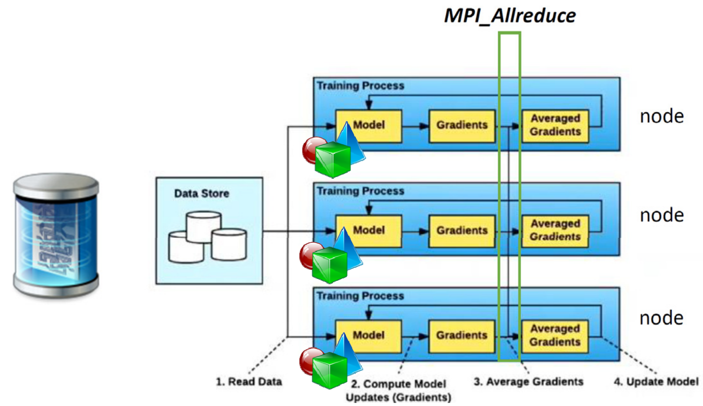
- Accuracy stable up to batch size = 8k
- For batch size > 8k training diverges
- Horovod enabled to significantly cut training time
- Scaling slightly less than linear (possibly due to data loading issues)

CoE RAISE: Distributed training influences the framework

- CoE RAISE AI models?
- Distributed Training?
- How it works with known toolks like TensorFlow?
- How it works with batch job scripts?



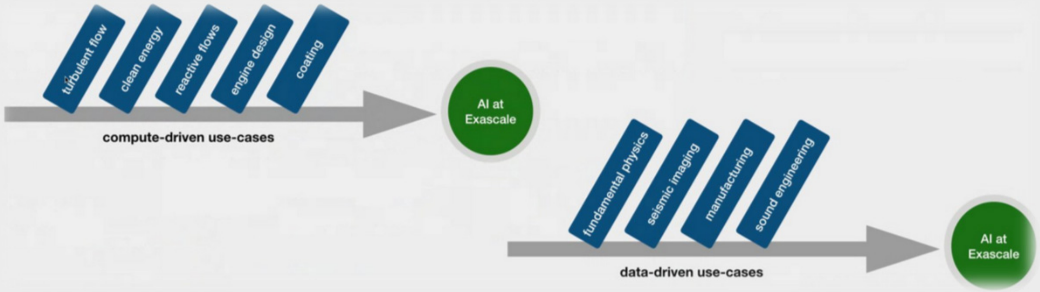
- The gradients for different batches of data are calculated separately on each node
- But averaged across nodes to apply consistent updates to the model copy in each node



→ Is that the preferred solution that scales really well?!

Selected Techniques to Identify Cross-Methods for HPC & AI

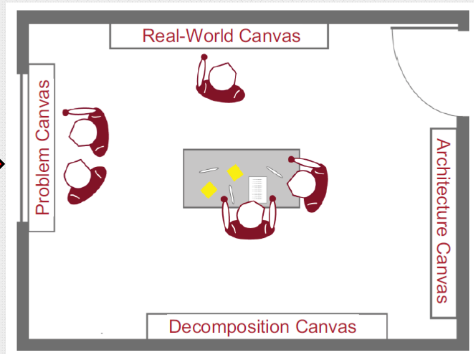
- Fact Sheets
  - Foster initial understanding
  - Living document & each Fact Sheet per WP3/WP4 Use Case
  - *(Experience from many other EU projects)*



- Selected Contents
  - Short Application Introduction
  - Clarify Primary Contacts
  - Codes/Libraries/Executables
  - HPC System Usage Details
  - Specific Platforms & 'where is what data'?
  - Machine/Deep Learning Approaches of Interest

initial steps

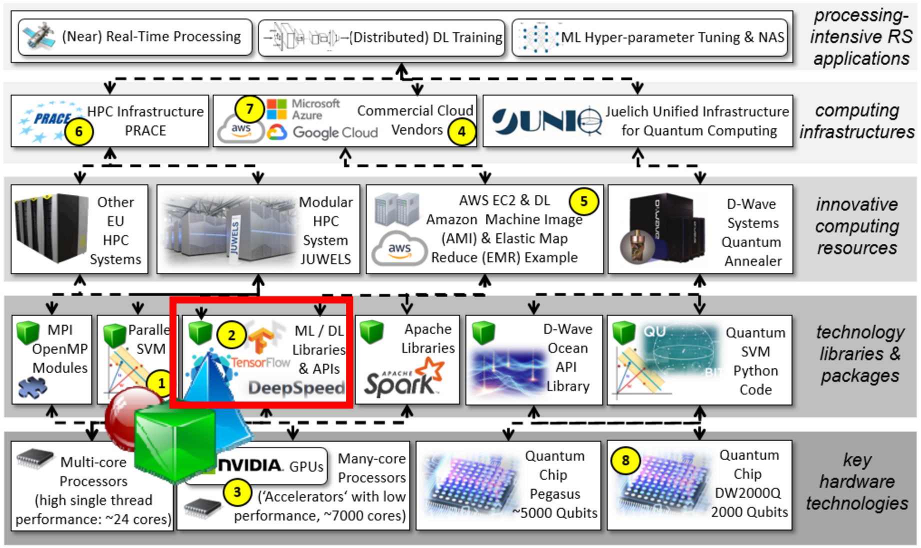
Driven by
Prof. Matthias Book &
Prof. Helmut Neukirchen



interaction room process

Fact Sheet Process of CoE RAISE & Early Co-Design Examples

RAISE
Center of Excellence



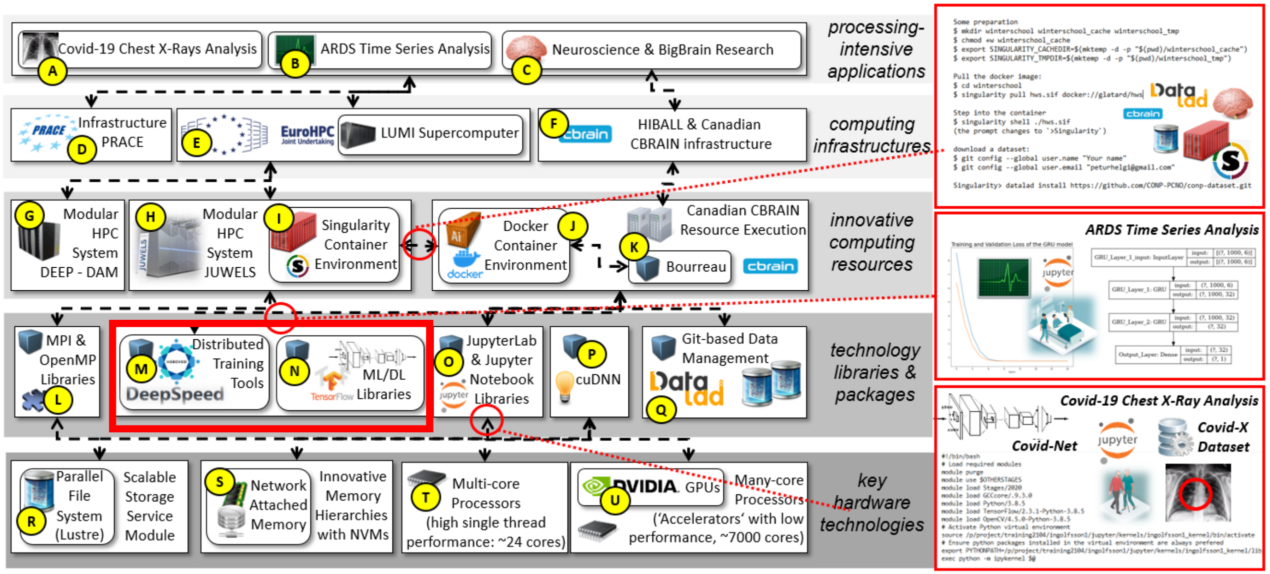
1 Parallel ML implementations still rare (MPI/OpenMP)
2 Open source tools good, but all need to fit in versions
3 Using very many GPUs beyond NVlink could be tricky
4 Look & feel of CC vendor ML services differ significantly
5 Costs of GPUs of CC vendors (e.g., EC2) tough, 24$/hour
6 GPU hours are free, but requires time grant proposal
7 Free GPUs in Google Colab vary in the available types
8 Works not yet with multi-class problems & large data

Legend:
N Highlighted Challenges & Experiences

Riedel, M., Cavallaro, G., Benediktsson, J.A.: Practice and Experience in using Parallel and Scalable Machine learning in Remote Sensing from HPC over Cloud to Quantum Computing, in conference proceedings of the IEEE IGARSS Conference, Brussels, Belgium, 2021, Physical and Online event, to appear
https://igarss2021.com/

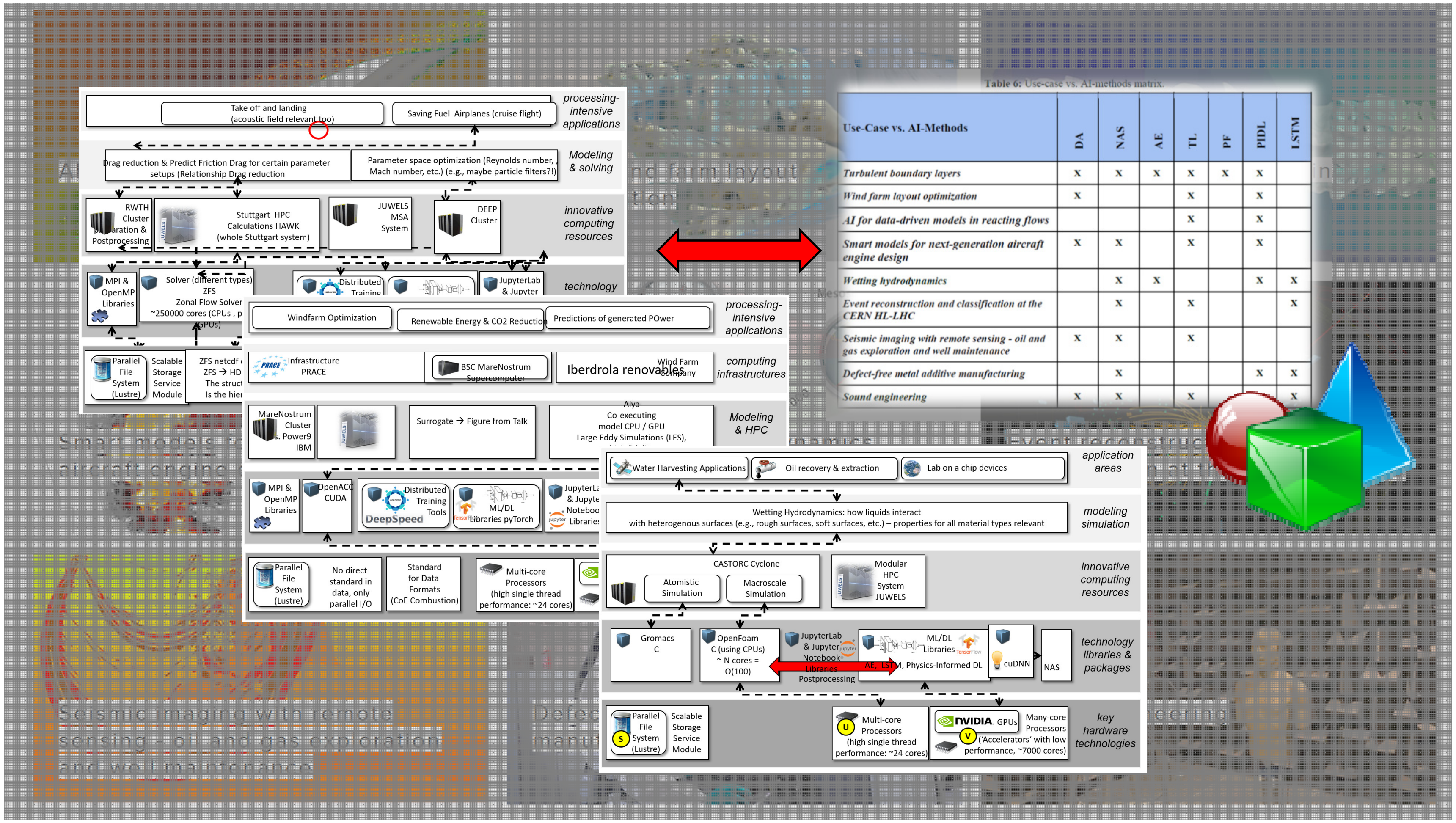
Riedel, M., Sedona, R., Barakat, C., Einarsson, P., Hassanian, R., Cavallaro, G., Book, M., Neukirchen, H., Lintermann, A.: Practice and Experience in using Parallel and Scalable Machine learning with Heterogenous Modular Supercomputing Architectures, in conference proceedings of the IEEE IDPDS Conference, Heterogenous Computing Workshop (HCW), Portland, USA, 2021, Online, to appear
https://www.ipdps.org/

IPDPS 2021 Portland



Compute- and Data-driven Use Cases Fact Sheets – Drafts(!)

RAISE Center of Excellence

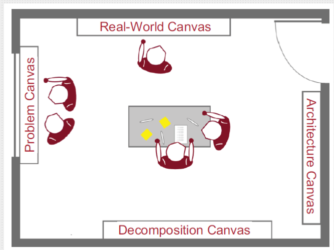


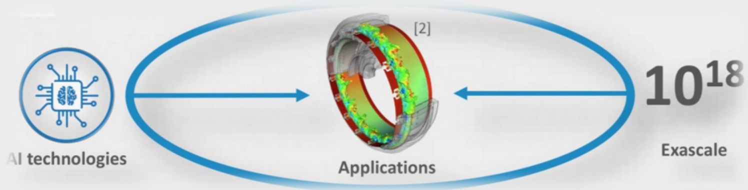
Use-Case vs. AI-Methods	DA	NAS	AE	TL	PF	PIDL	LSTM
Turbulent boundary layers	x	x	x	x	x	x	
Wind farm layout optimization	x			x		x	
AI for data-driven models in reacting flows				x		x	
Smart models for next-generation aircraft engine design	x	x		x		x	
Wetting hydrodynamics		x	x			x	x
Event reconstruction and classification at the CERN HL-LHC		x		x			x
Seismic imaging with remote sensing - oil and gas exploration and well maintenance	x	x		x			
Defect-free metal additive manufacturing		x				x	x
Sound engineering	x	x		x			x

WORK IN PROGRESS

HPC Systems Engineering in the Interaction Room Seminar

- CoR RAISE Interaction Room Process as Next Step
  - Supports the proper software engineering design of the unique AI framework blueprint
  - Expecting to work with WP3 & WP4 experts in an open minded way
  - Process will be guided by **Prof. Dr. Matthias Book** (University of Iceland)
  - Supported by Software Engineering & testing expert **Prof. Dr. Helmut Neukirchen** (University of Iceland)
  - CoE RAISE @ YouTube: https://www.youtube.com/channel/UCAdIZ-v6cWwGdapwYxdN7dg
  - Methology as one CoE RAISE outcome

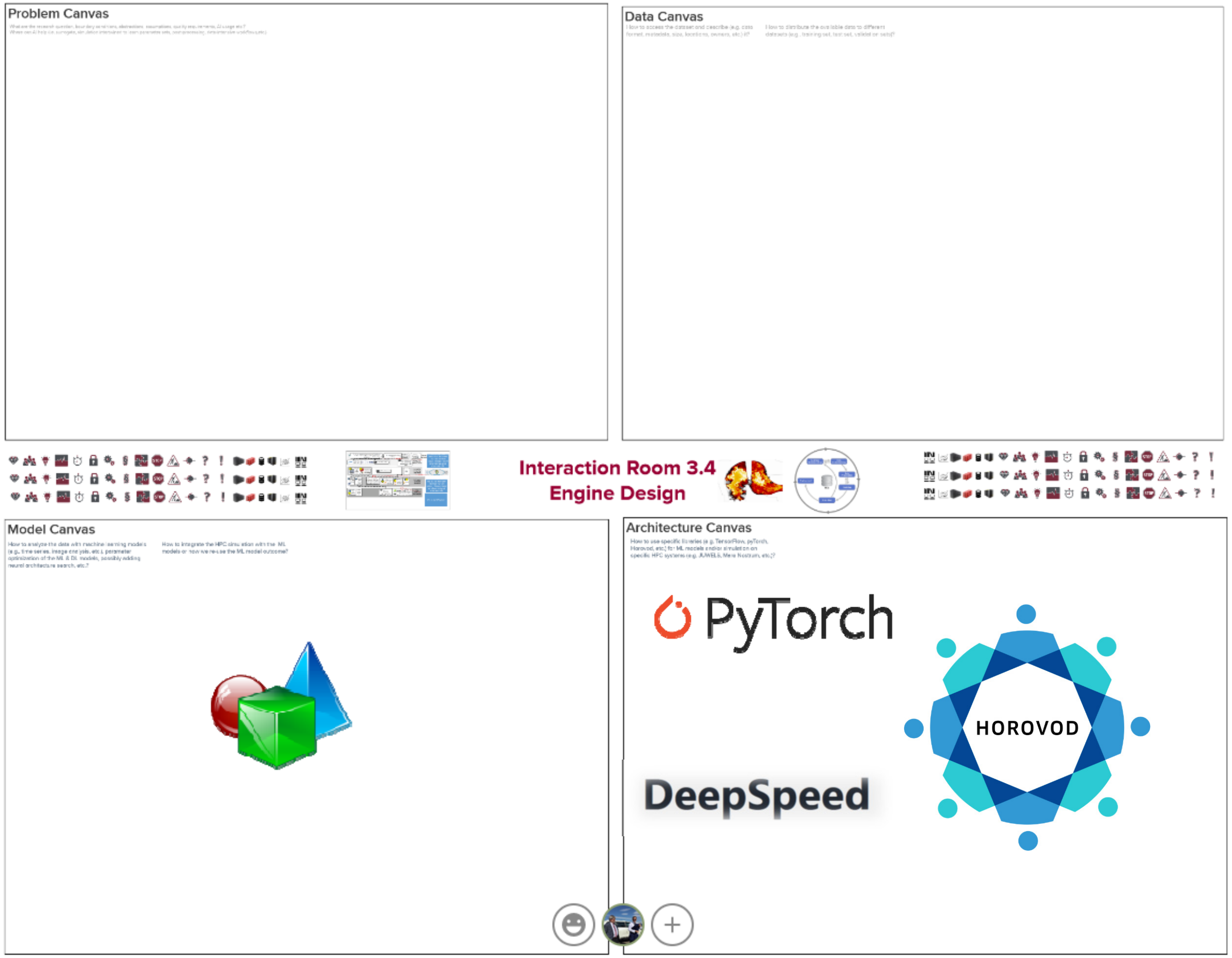




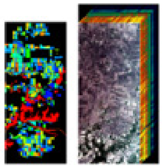
Book, M., **Riedel, M.**, **Neukirchen, H.**, Goetz, M.: Facilitating Collaboration in High-Performance Computing Projects with an Interaction Room, in conference proceedings of the 4th ACM SIGPLAN International Workshop on Software Engineering for Parallel Systems (SEPS 2017), October 22-27, 2017, Vancouver, Canada

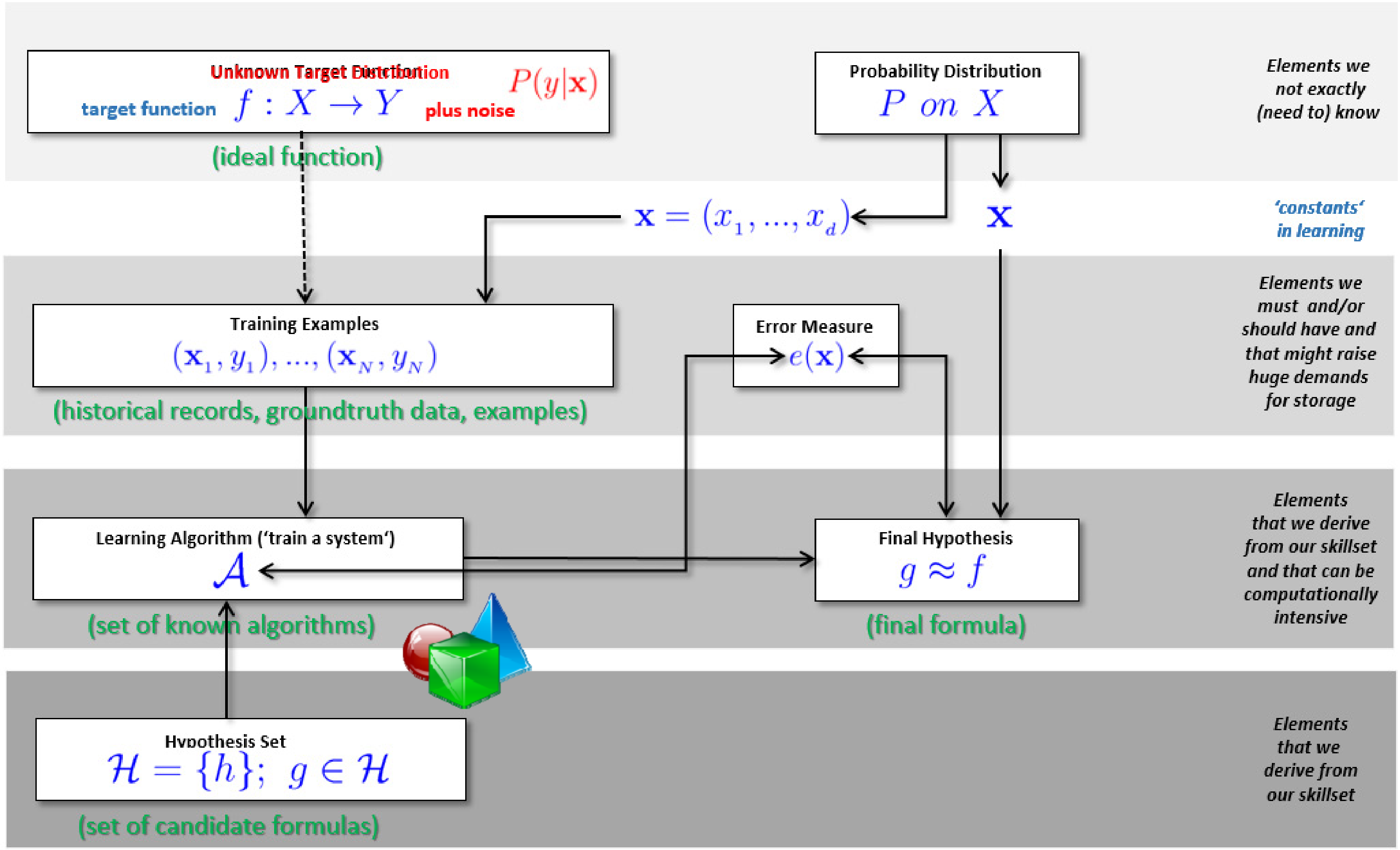
Interaction Rooms via MURAL Boards: Distributed Training?!

RAISE Center of Excellence



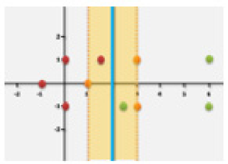
Distributed Training: 1st Impact in Model Development





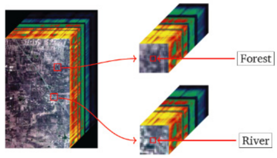
PyTorch

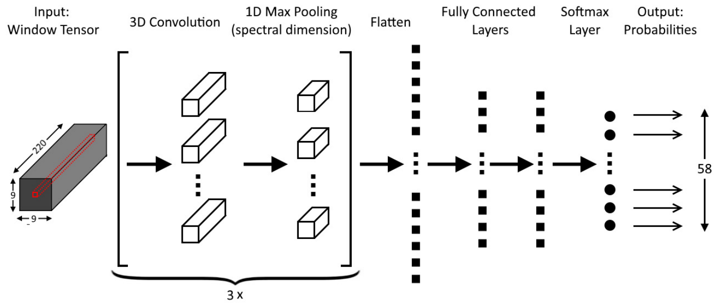
DeepSpeed



Distributed Training: 2nd Impact in Model Fine-Tuning

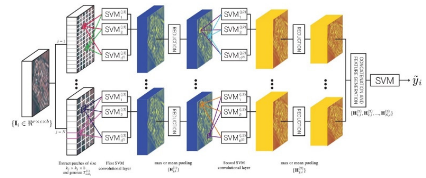
- **Using Convolutional Neural Networks (CNNs) with hyperspectral remote sensing image data**





Feature	Representation / Value
Conv. Layer Filters	48, 32, 32
Conv. Layer Filter size	(3, 3, 5), (3, 3, 5), (3, 3, 5)
Dense Layer Neurons	128, 128
Optimizer	SGD
Loss Function	mean squared error
Activation Functions	ReLU
Training Epochs	600
Batch Size	50
Learning Rate	1
Learning Rate Decay	5×10^{-6}

- **Find Hyperparameters & joint 'new-old' modeling & transfer learning given rare labeled/annotated data in science (e.g. 36,000 vs. 14,197,122 images ImageNet)**

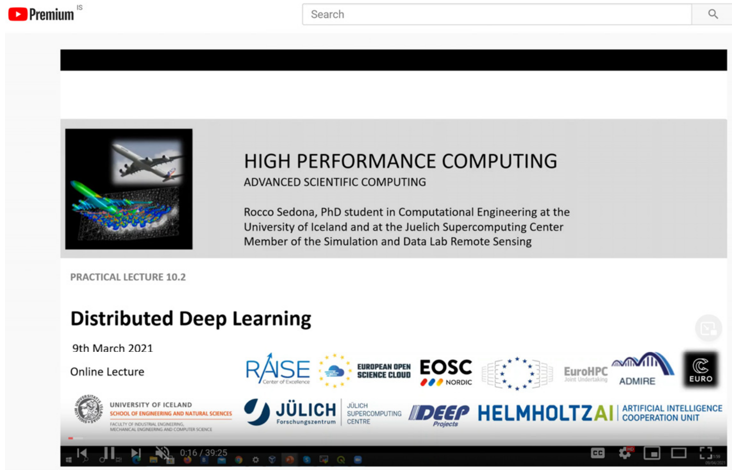


J. Lange, G. Cavallaro, M. Goetz, E. Erlingsson, M. Riedel, 'The Influence of Sampling Methods on Pixel-Wise Hyperspectral Image Classification with 3D Convolutional Neural Networks', Proceedings of the IGARSS 2018 Conference, Online: https://www.researchgate.net/publication/328991957_The_Influence_of_Sampling_Methods_on_Pixel-Wise_Hyperspectral_Image_Classification_with_3D_Convolutional_Neural_Networks

G. Cavallaro, Y. Bazi, F. Melgani, M. Riedel, 'Multi-Scale Convolutional SVM Networks for Multi-Class Classification Problems of Remote Sensing Images', Proceedings of the IGARSS 2019 Conference, Online: https://www.researchgate.net/publication/337439088_Multi-Scale_Convolutional_SVM_Networks_for_Multi-Class_Classification_Problems_of_Remote_Sensing_Images

Distributed Training with Horovod: More Information

- High-Peformance Computing Course
  - University of Iceland
  - In collaboration with FZJ



  - YouTube Channel: https://www.youtube.com/channel/UCWC4VKHmL4NZgFfKoHtANKg
  - Practical Lecture 10.2: Distributed Deep Learning (by Rocco Sedona)
  - https://www.youtube.com/watch?v=8dtg0IDnQO0&list=PLmJwSK7qduwVnlrIPjrfSn7QRcv3wIQj5&index=32
  - 2 x 40 minutes

drive. enable. innovate.

RAISE Center of Excellence

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