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

RAISE Center of Excellence

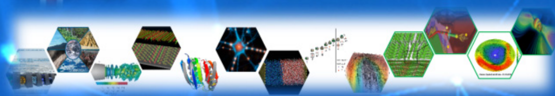
EURO

CoE RAISE Use Case Foundations & Lessons Learned from Fact Sheets

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

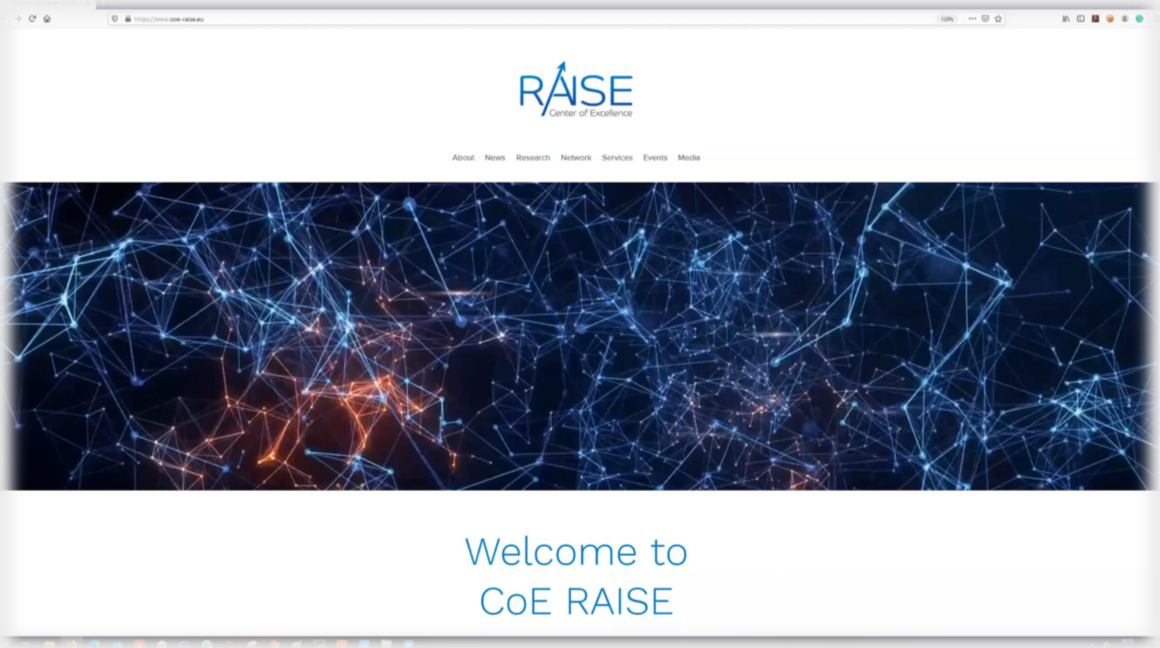
2021-04-08, RAISE CoE Seminar HPC Systems Engineering in the Interaction Room, Online

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



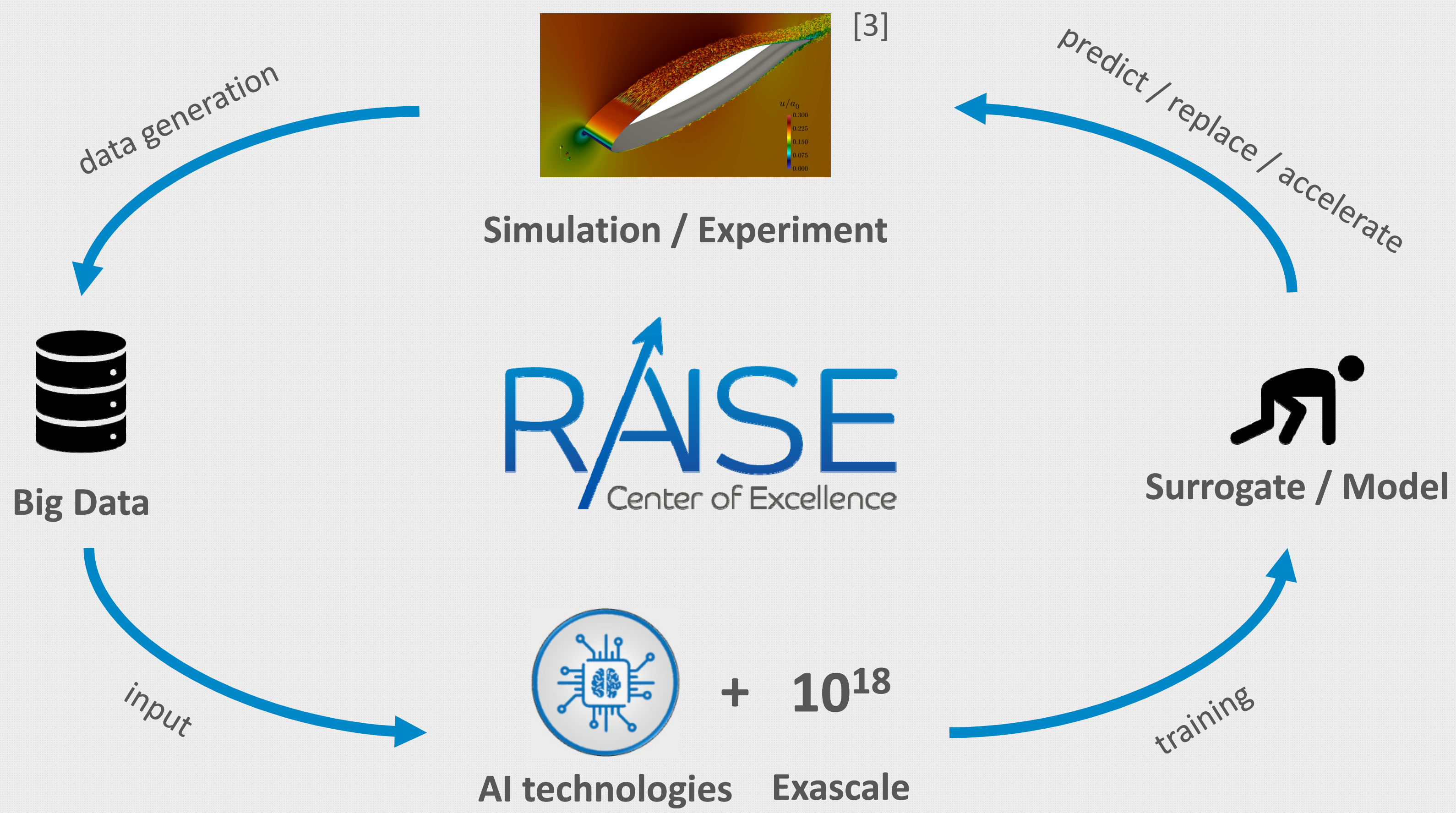
@ProfDrMorrisRiedel @Morris Riedel @MorrisRiedel @MorrisRiedel
https://www.youtube.com/channel/UCWC4VKHmL4NZgFfKoHtANKg
morris@hi.is

CoE RAISE Web Page & More Information



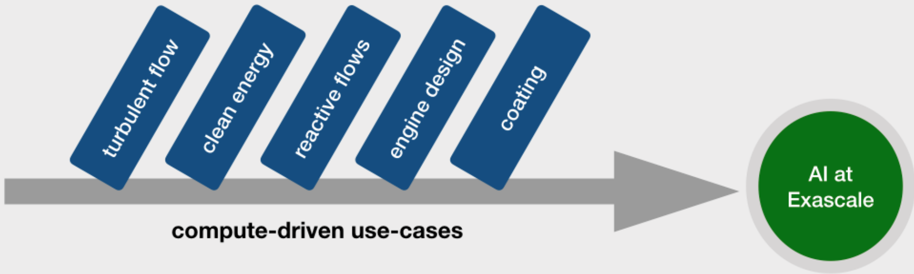
https://www.coe-raise.eu

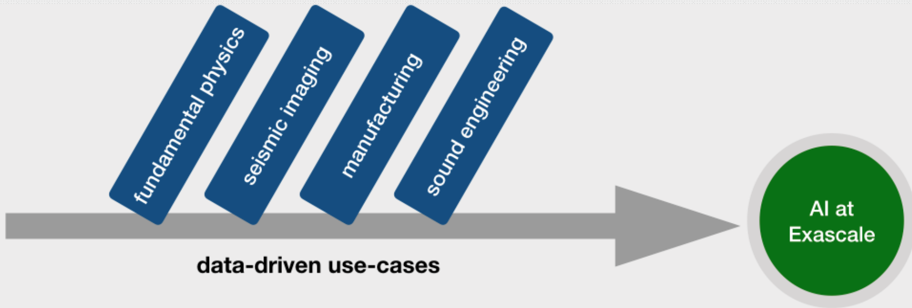
CoE RAISE – Motivation & Approach

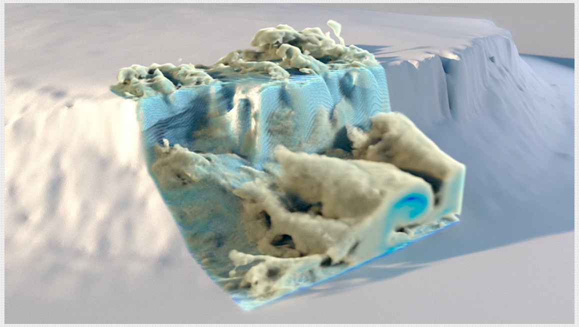


Use Cases in CoE RAISE

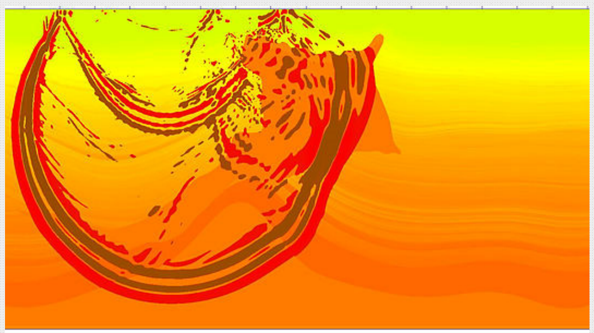
- Two kinds of use cases:





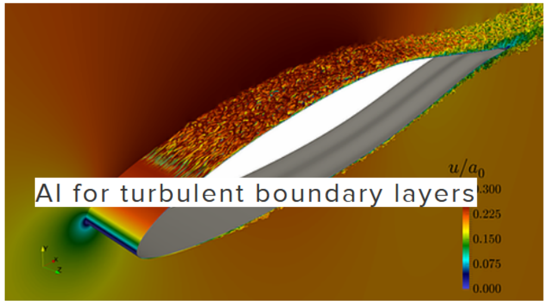


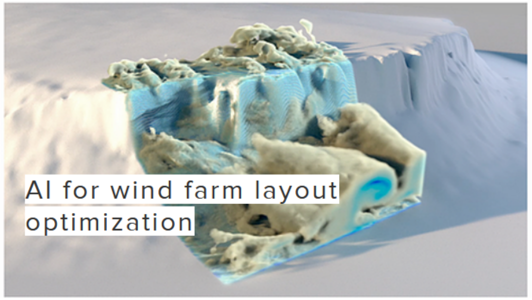
Example from use case "AI for wind farm layout": Turbulence generated by a cliff on Bolund Island, Denmark.

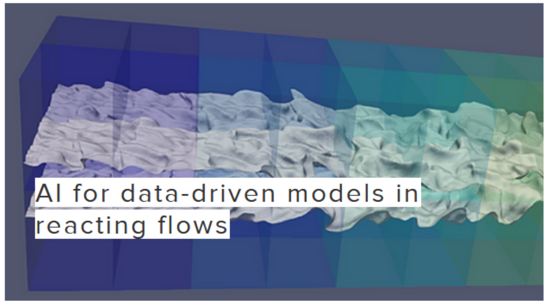


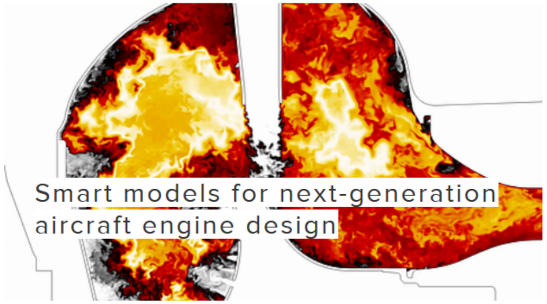
Example from use case " Seismic imaging with remote sensing - oil and gas exploration and well maintenance": Snapshot from a wavefield.

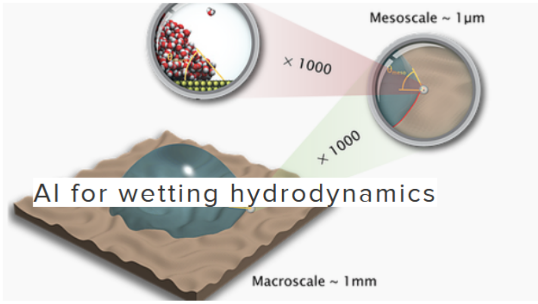
Compute- and Data-driven Use Cases – Overview

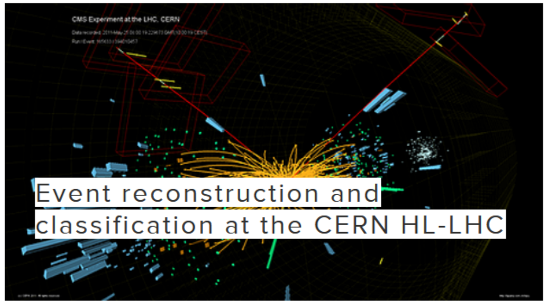


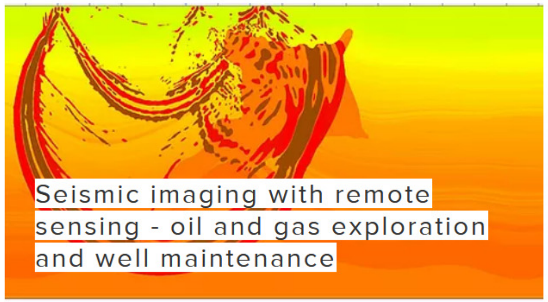


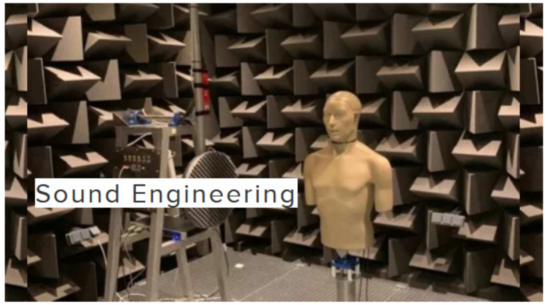












Partners in CoE RAISE

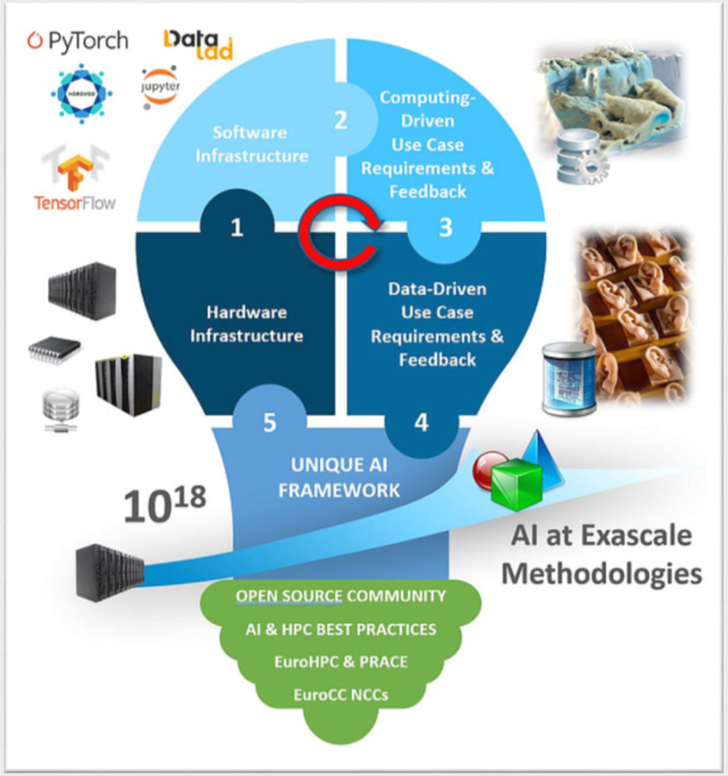
RAISE Center of Excellence

SAFRAN
JÜLICH Forschungszentrum
CERFACS
Atos
Delphi Consortium
1862 RĪGAS TEHNISKĀ UNIVERSITĀTE
CERN
UNIVERSITY OF ICELAND
FLANDERS MAKE
ParTec CLUSTER COMPETENCE CENTER
The Modular Supercomputing Company
BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación
RWTH AACHEN UNIVERSITY
The Cyprus Institute
RESEARCH·TECHNOLOGY·INNOVATION

CoE RAISE's Objectives (1)

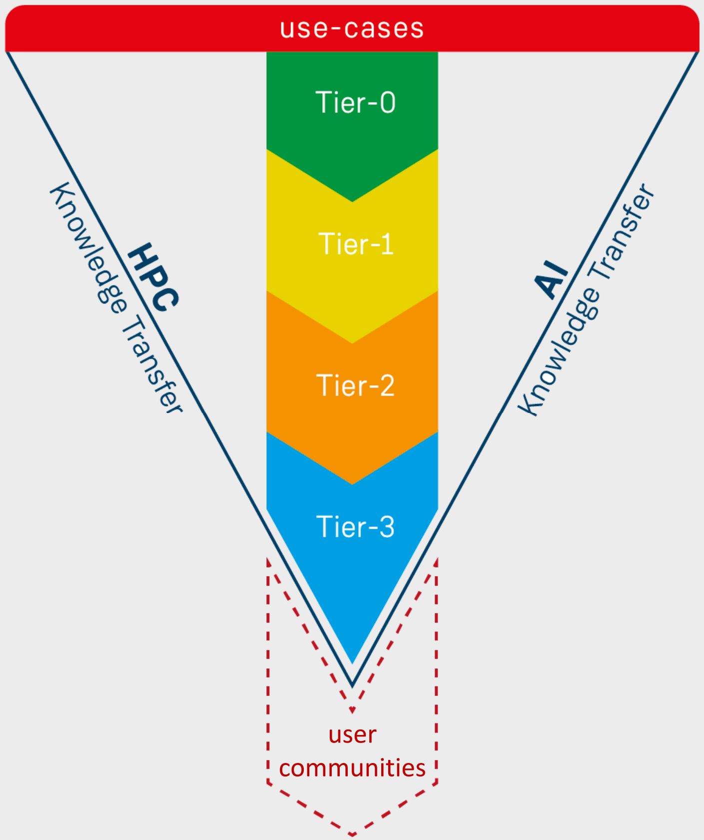
- Development of AI methods towards Exascale along use-cases
- RAISE tightly connects
  - an exceptional hardware infrastructure,
  - an usable and versatile software infrastructure,
  - compute-driven use cases,
  - and data-driven use cases

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



CoE RAISE's Objectives

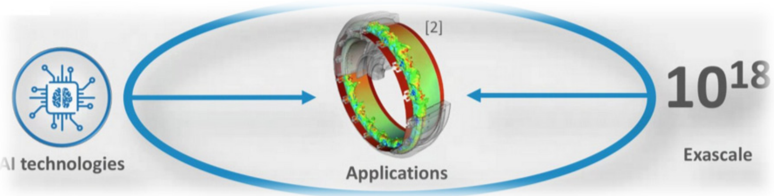
- Knowledge and technology transfer to science and industry
  - Service development and education
- Creation of a European network to support less developed entities
  - Equalize knowledge and technology across borders
- Business development to ensure sustainability
  - Market analysis and finding a market niche for the CoE
- Connect to others
  - Connect to other CoEs, Pre-Exascale projects, and other existing projects



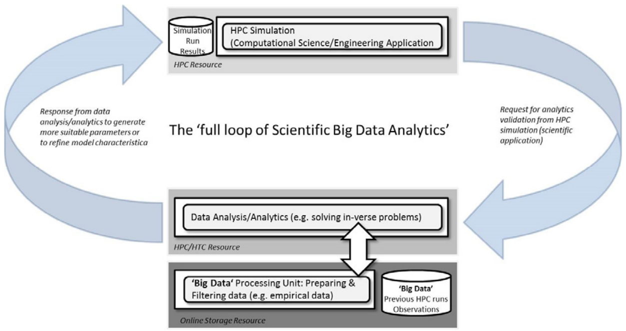
Vision – Intertwined HPC Simulations & AI – 'full loop' ?

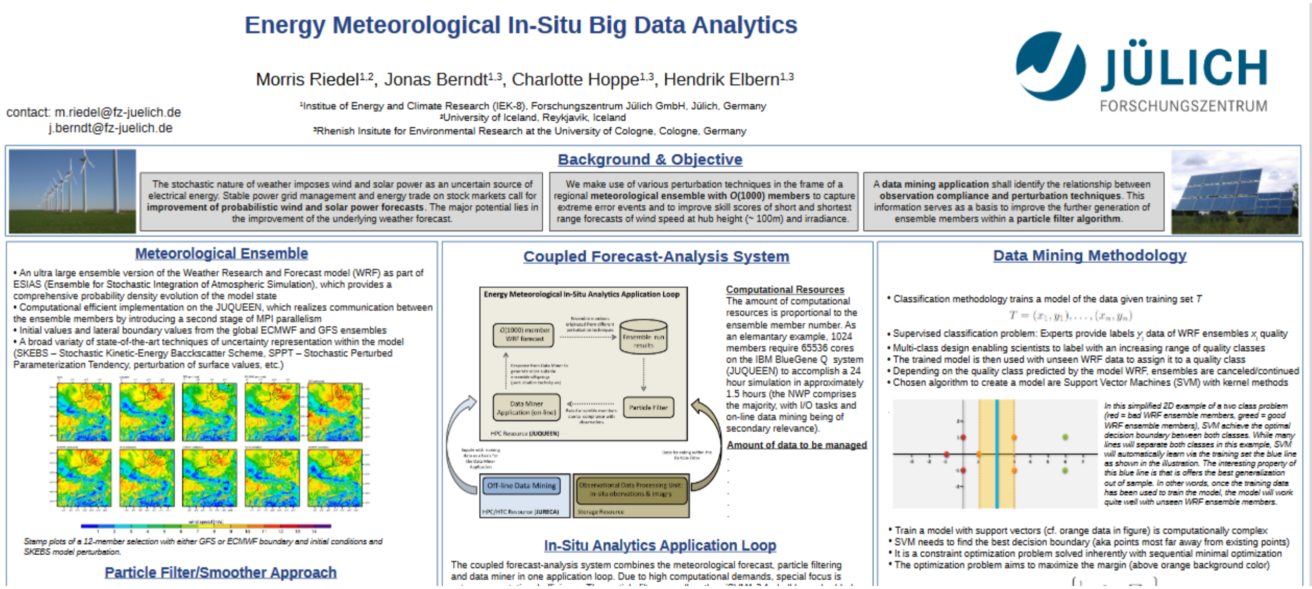
RAISE Center of Excellence

- What means AI & HPC Cross Methods at Exascale?



Today rather high performance data analytics (HPDA)





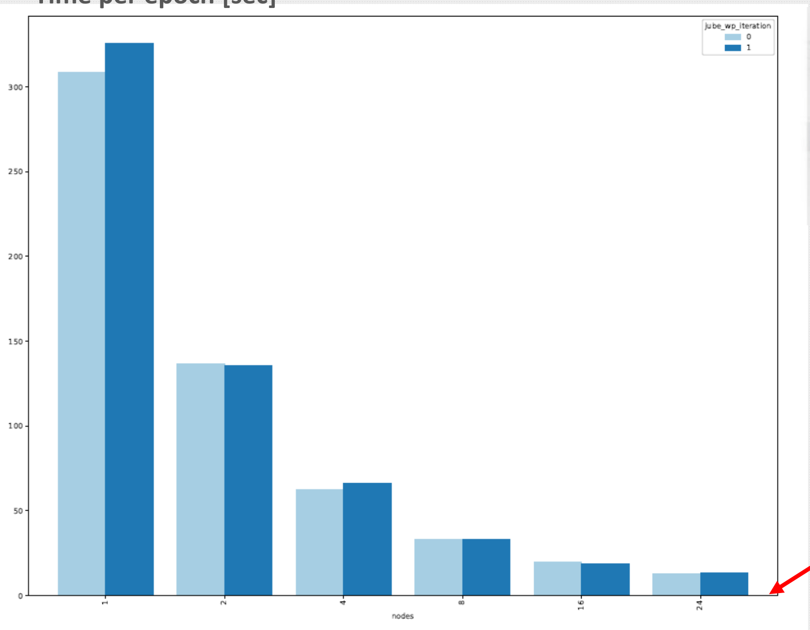
Lippert, T., Mallmann, D., **Riedel, M.**: Scientific Big Data Analytics by HPC, in Symposium proceedings of NIC Symposium 2016 – publication Series of the John von Neumann Institute for Computing (NIC), NIC Series 48 (417), ISBN 978-3-95806-109-5, February 11-12, 2016, Juelich, Germany

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)]

Can AI do Exascale & use Disruptive Technologies?

RAISE Center of Excellence

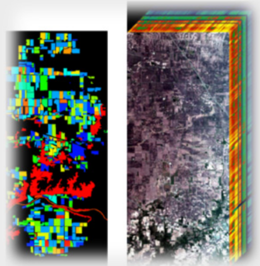
Time per epoch [sec]

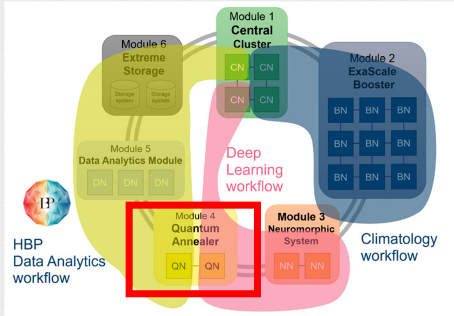


Europe #1 Supercomputer (11/2020)

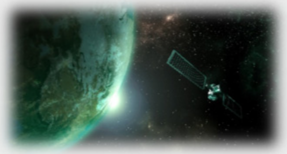
Example from 2019:
Using partition of the JUWELS system has 56 compute nodes, each with 4 NVIDIA V100 GPUs (equipped with 16 GB of memory)

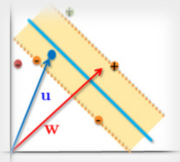
24 nodes x 4 GPUs = 96 GPUs






```
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)
```



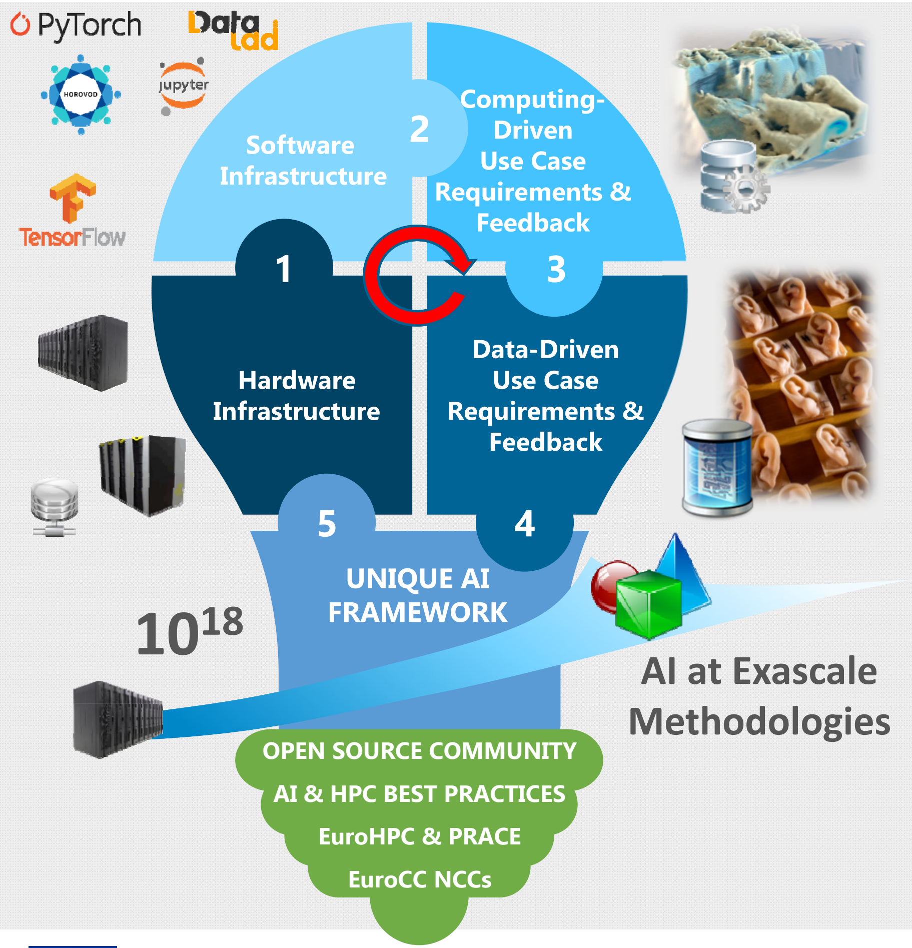


Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., **Riedel, M.**, Book, M.: SCALING UP A MULTISPECTRAL RESNET-50 TO 128 GPUS, in conference proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020), September 26 – October 2nd, 2020, Virtual Conference, Hawai, USA

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

Towards AI & HPC at Exascale with CoE RAISE Results



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

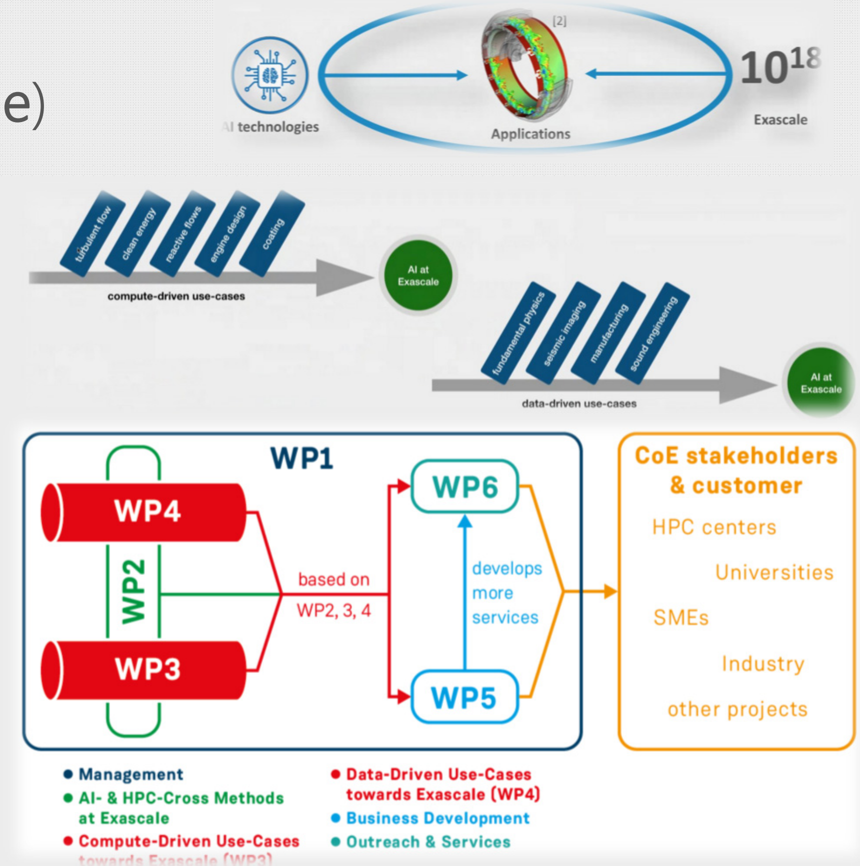
Examples: Deployment blueprint by using AI training on cluster module & inference/testing on booster

→ UNIQUE AI FRAMEWORK

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

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

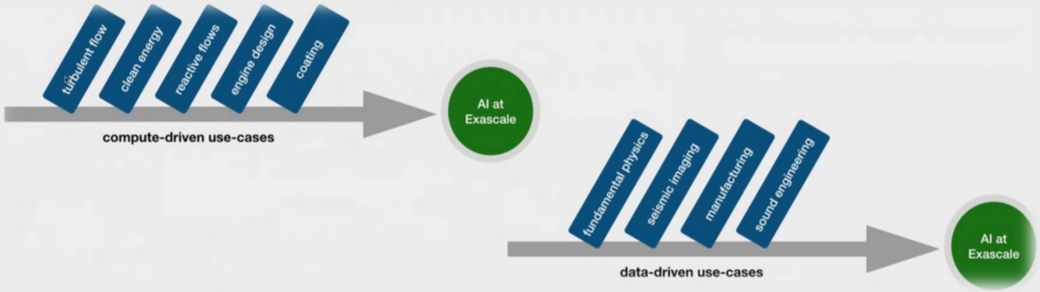
- WP3 (Compute-Driven Use-Cases towards Exascale)
- WP4 (Data-Driven Use-Cases towards Exascale
- Developments in these WPs will be supported by the cross-linking activities of WP2
  - E.g. scaling machine & deep learning codes with frameworks like Horovod/Deepspeed
  - E.g. introduction to new AI methods 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



Selected Techniques to Identify Cross-Methods for HPC & AI

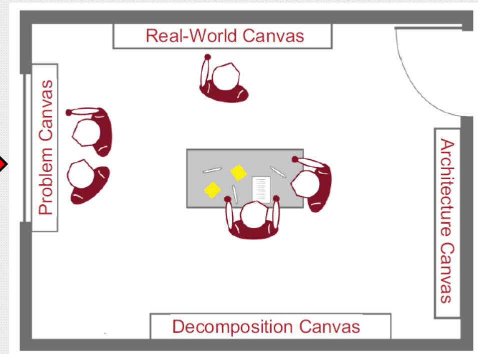
RAISE Center of Excellence

- 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

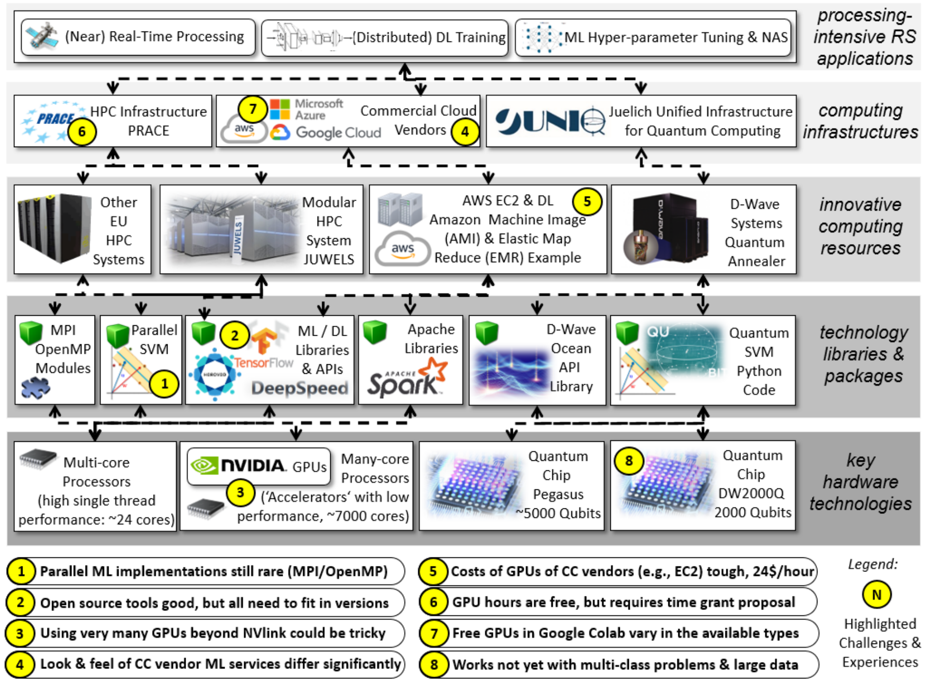
initial steps



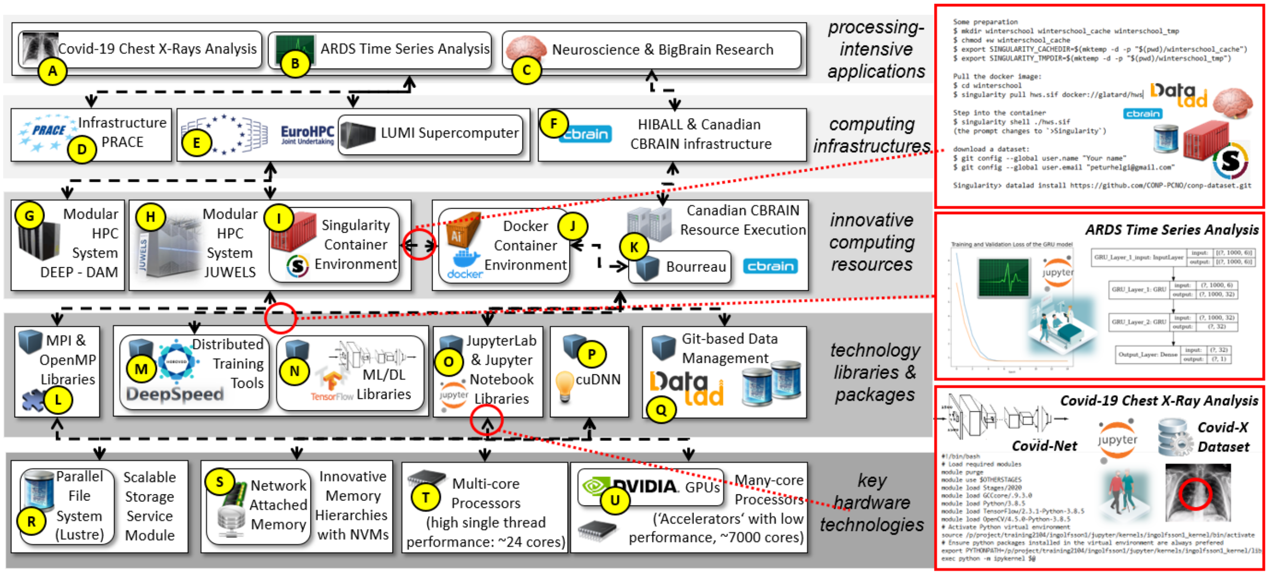
Presentations by Prof. Matthias Book & Prof. Helmut Neukirchen

interaction room process

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



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/



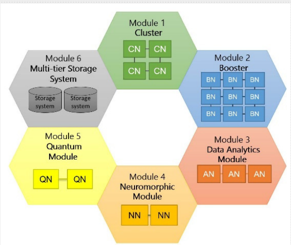
Riedel, M., Cavallaro, G., Benediktsson, J.A.: Practice and Experience in using Parallel and Scalable Machine learning in Remote Sensing from HPC over Cloud to Quantum Computing, in conference proceedings of the IEEE IGARSS Conference, Brussels, Belgium, 2021, Physical and Online event, to appear
https://igarss2021.com/

Fact Sheet Process of CoE RAISE & Paper Example

RAISE Center of Excellence

- 'not a waste of time': Fact Sheets can be re-used for publications, project presentations & dissemination

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/



Practice and Experience in using Parallel and Scalable Machine Learning with Heterogenous Modular Supercomputing Architectures

Morris Riedel, Department of Computer Science, University of Iceland, Reykjavik, Iceland, morris@hi.is
Rocco Sedona, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, r.sedona@fz-juelich.de
Chadi Barakat, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, c.barakat@fz-juelich.de
Petur Einarsson, Department of Computer Science, University of Iceland, Reykjavik, Iceland, peturhelgi@gmail.com
Reza Hassanian, Department of Computer Science, University of Iceland, Reykjavik, Iceland, seh38@hi.is
Gabriele Cavallaro, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, g.cavallaro@fz-juelich.de
Matthias Book, Department of Computer Science, University of Iceland, Reykjavik, Iceland, book@hi.is
Helmut Neukirchen, Department of Computer Science, University of Iceland, Reykjavik, Iceland, helmut@hi.is
Andreas Lintermann, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, a.lintermann@fz-juelich.de

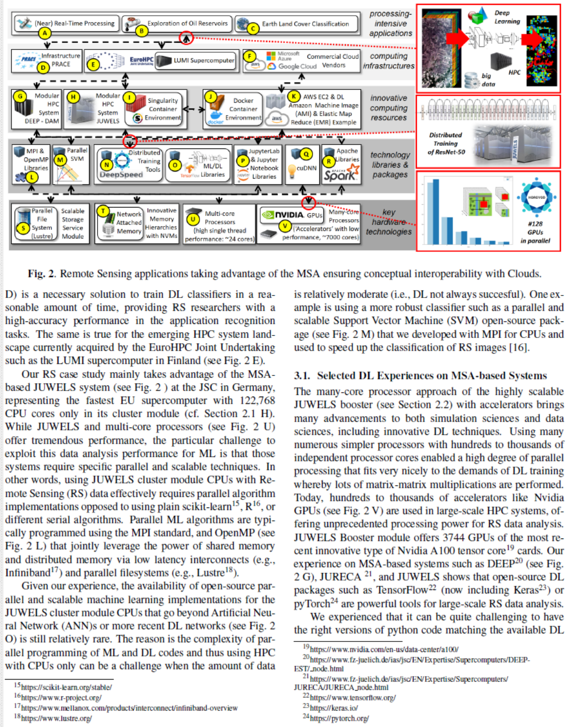


Fig. 2. Remote Sensing applications taking advantage of the MSA ensuring conceptual interoperability with Clouds.

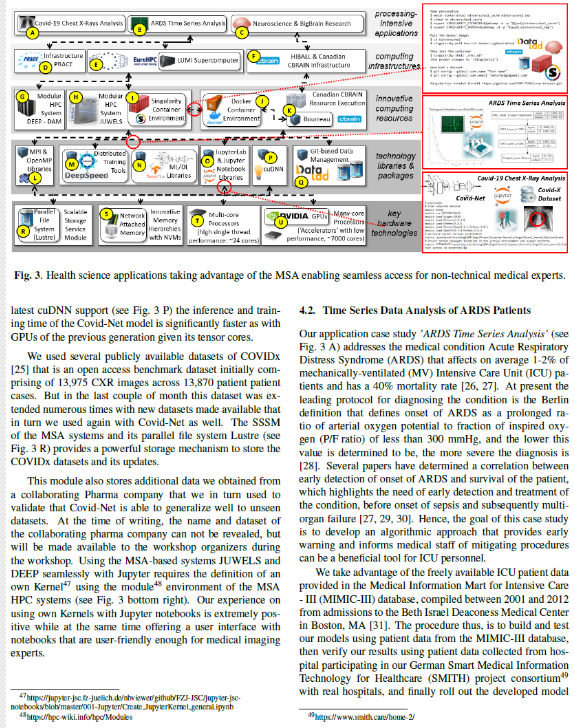


Fig. 3. Health science applications taking advantage of the MSA enabling seamless access for non-technical medical experts.

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

RAISE Center of Excellence

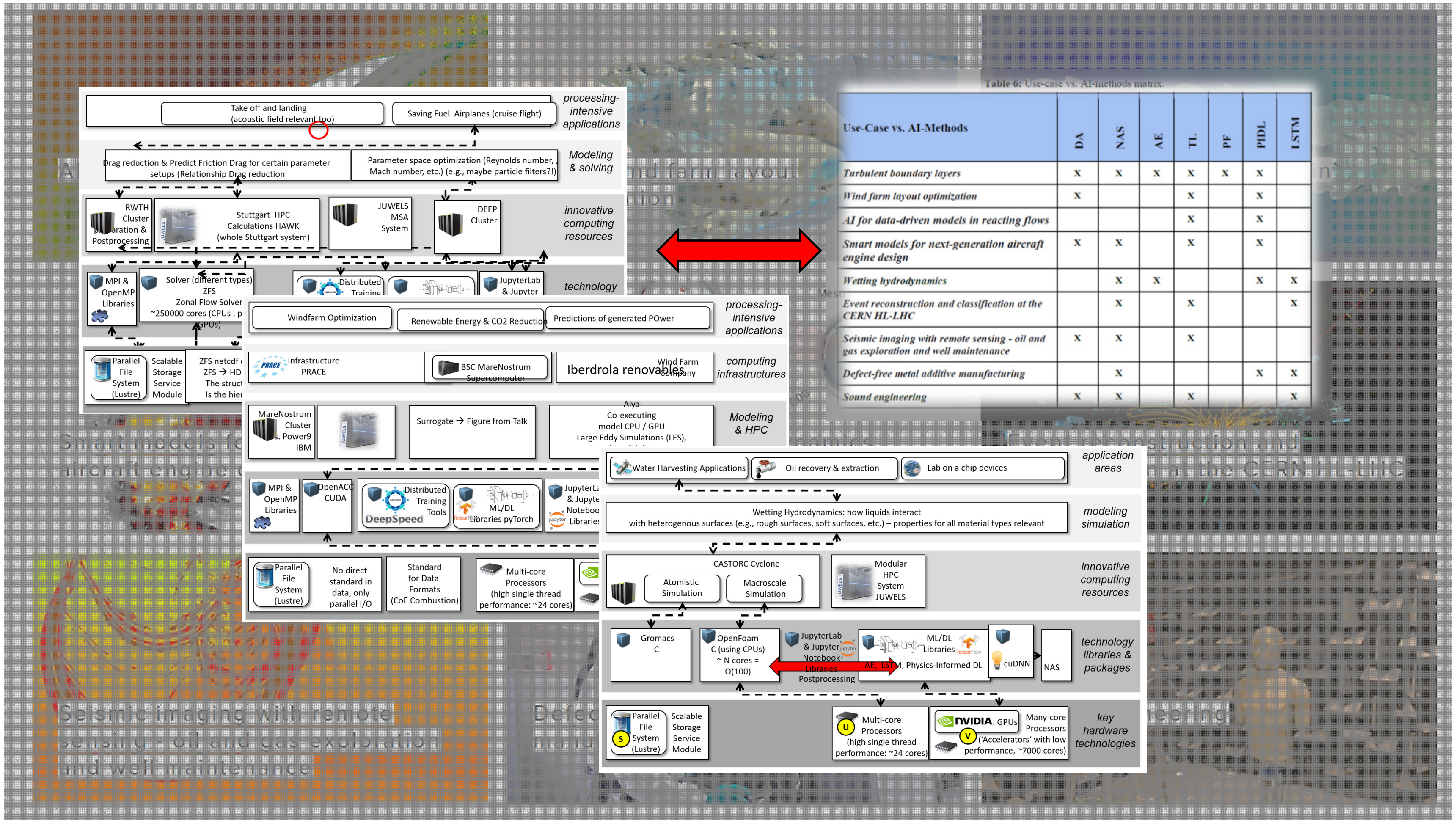


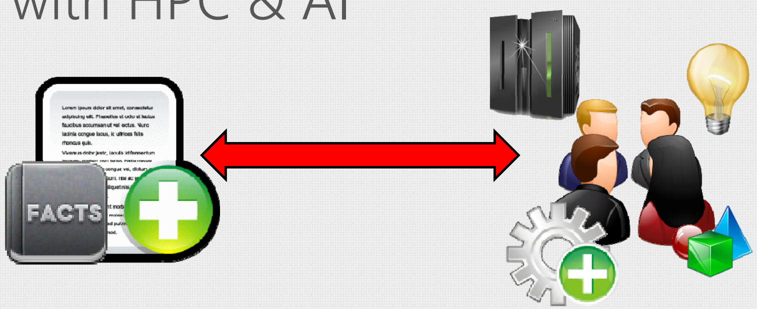
Table 6: Use-case vs. AI-methods matrix.

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

Lessons Learned from CoE RAISE Draft Factsheets

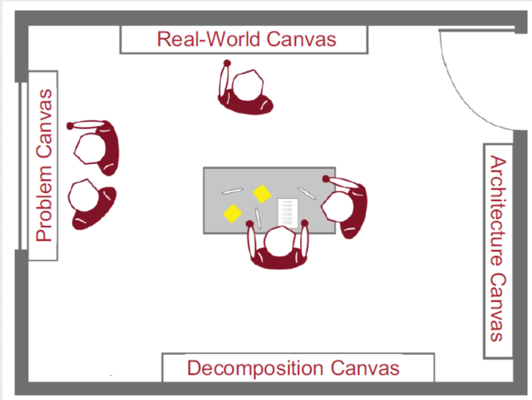
- Fact Sheet process
  - Participants have been very positive understanding use cases & each other much better
  - Enables really to understand where what components of use cases are running & why needed
  - Fosters understanding what is confidential, e.g., what could be goals to for industry for a patent
- Massive complexity observed moving towards Exascale with HPC & AI
  - Software engineering expertise required
  - AI expertise required
  - HPC expertise required
  - Application domain-science know-how required

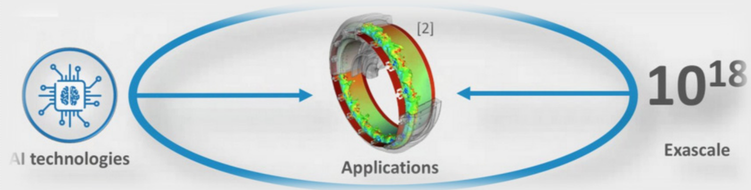


- Understanding & Communication
  - Essential and lots of expertise area-specific terminology and misunderstandings
  - Need for a systematic method to succeed in the nine use cases of CoE RAISE & external use cases

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)
  - 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 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