KNMI & JSC & UNIVERSITY OF ICELAND COLLABORATION – WELCOME

PROF. DR. – ING. MORRIS RIEDEL, JUELICH SUPERCOMPUTING CENTRE (JSC) / UNIVERSITY OF ICELAND
HEAD OF HIGH PRODUCTIVITY DATA PROCESSING & CROSS-SECTIONAL TEAM DEEP LEARNING
3TH DECEMBER JUELICH SUPERCOMPUTING CENTRE, FORSCHUNGSZENTRUM JUELICH, GERMANY

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and Water Management

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

DEEP Projects

RDA
RESEARCH DATA ALLIANCE

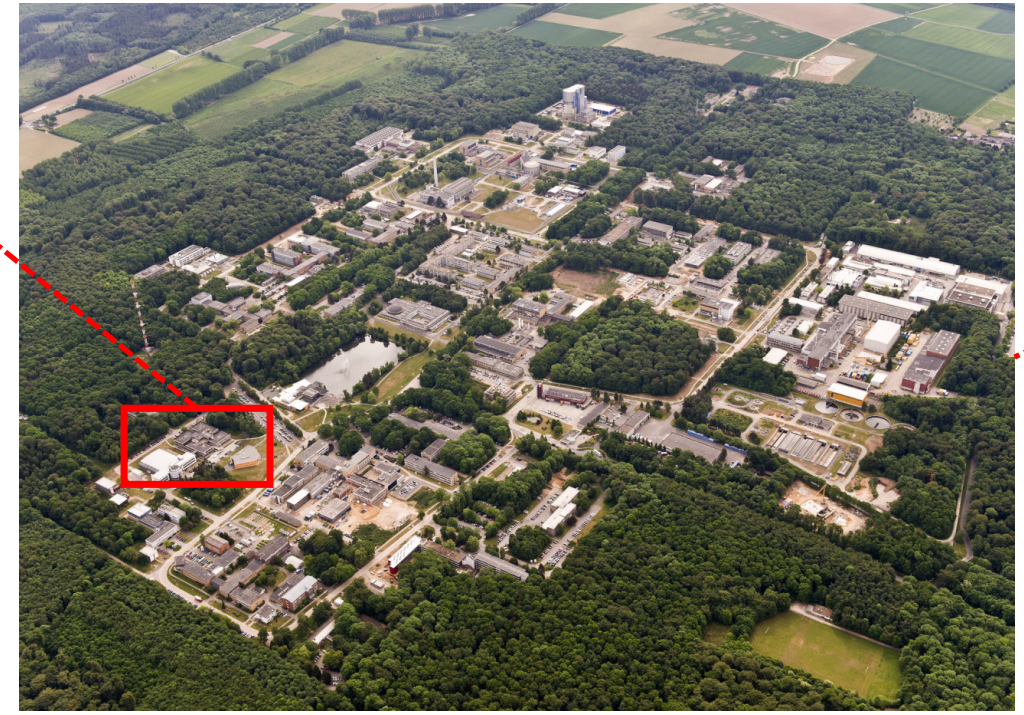
UNIVERSITY OF ICELAND
SCHOOL OF ENGINEERING AND NATURAL SCIENCES
FACULTY OF INDUSTRIAL ENGINEERING, MECHANICAL ENGINEERING AND COMPUTER SCIENCE

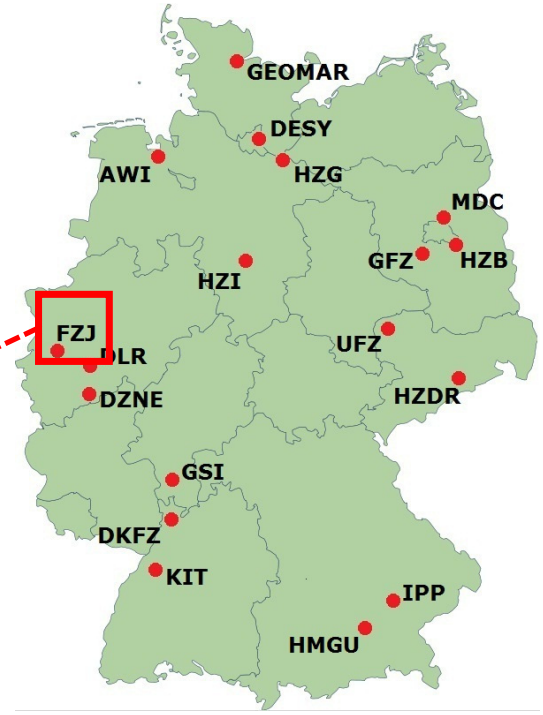
JÜLICH
Forschungszentrum
JÜLICH SUPERCOMPUTING CENTRE

FORSCHUNGSZENTRUM JUELICH (FZJ)

Multi-Disciplinary Research Centre of the Helmholtz Association in Germany

(Juelich Supercomputing Centre known as JSC)





- Selected Facts
  - One of EU largest inter-disciplinary research centres (~5000 employees)
  - Special expertise in physics, materials science, nanotechnology, neuroscience and medicine & information technology (HPC & Data)

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

[1] Holmholtz Association Web Page

EUROPEAN UNION & COMMISSION PLANS

Supporting Artificial Intelligence & Supercomputers – Objectives are In-line with EU Strategic Plans

"By supporting strategic projects in frontline areas such as artificial intelligence, supercomputers, cybersecurity or industrial digitisation, and investing in digital skills, the new programme will help to complete the Digital Single Market, a key priority of the Union."

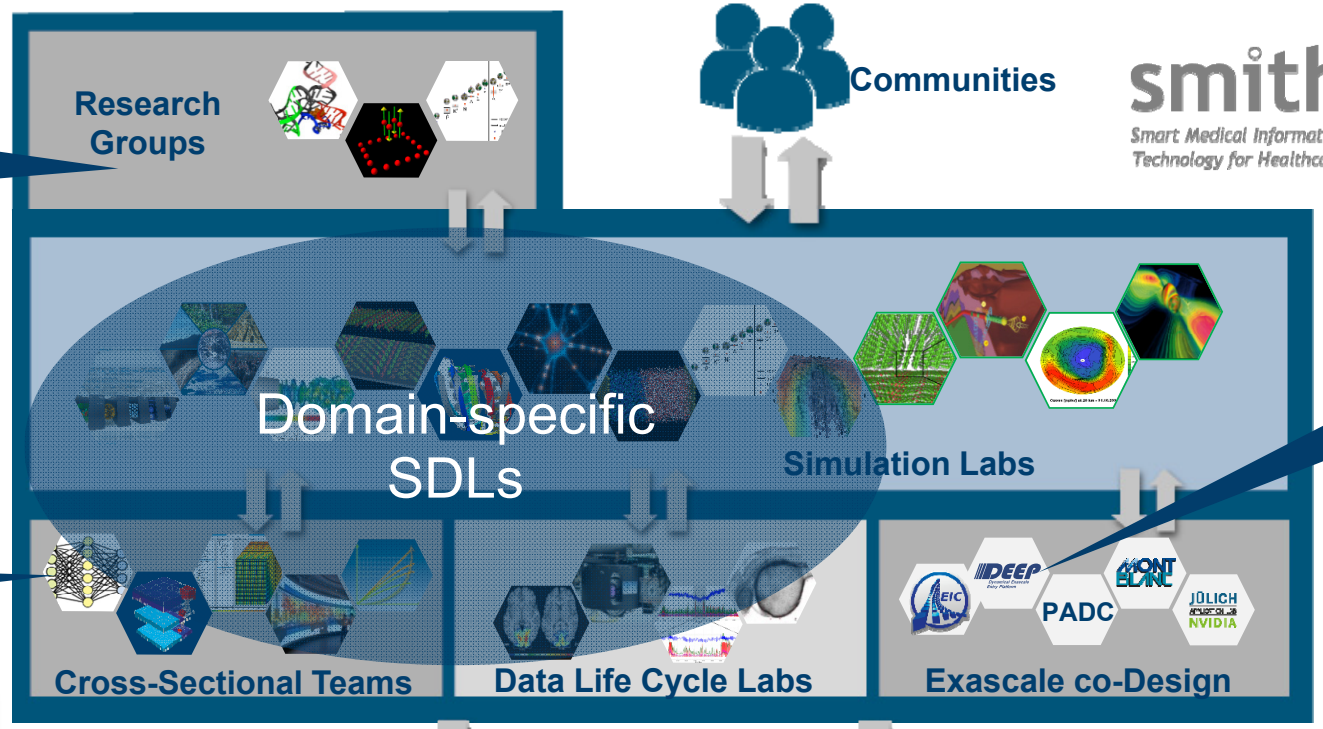
[11] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, EC, 2018, 2nd May 2018

JUELICH SUPERCOMPUTING CENTRE (JSC) OF FZJ

Simulation & Data Labs (SDL) using High Performance Computing (HPC)

Smart Data Innovation Lab

Research Group High Productivity Data Processing



smith
Smart Medical Information Technology for Healthcare

SOCCERWATCH BETA

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

JÜLICH Forschungszentrum | JÜLICH SUPERCOMPUTING CENTRE

DEEP-EST EU PROJECT

Cross-Sectional Team Deep Learning

DEEP Projects

HPC Systems JURECA

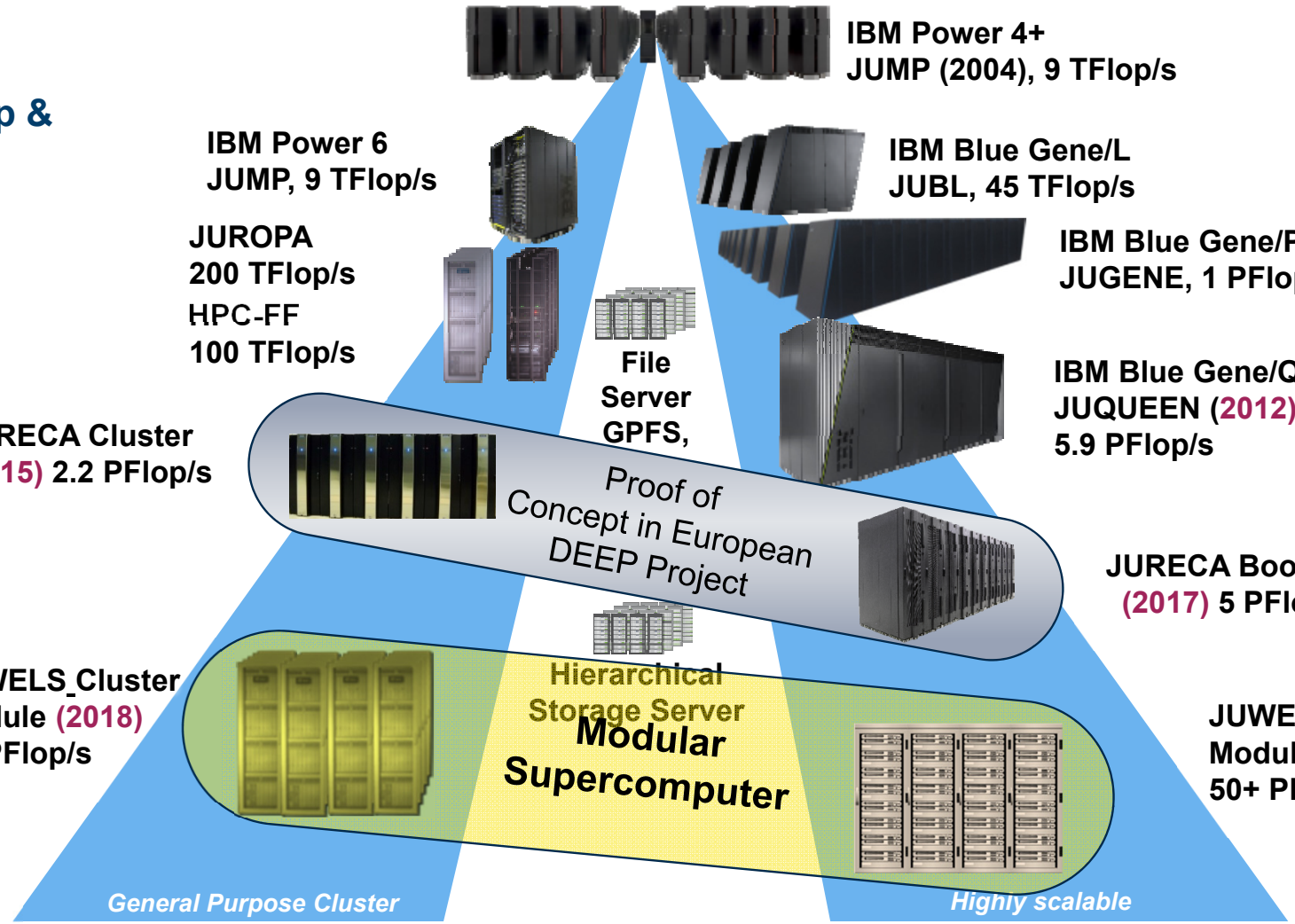
Facilities

Modular Supercomputer JUWELS

Industry Relations Team

JSC

HPC Roadmap & Key Vendors

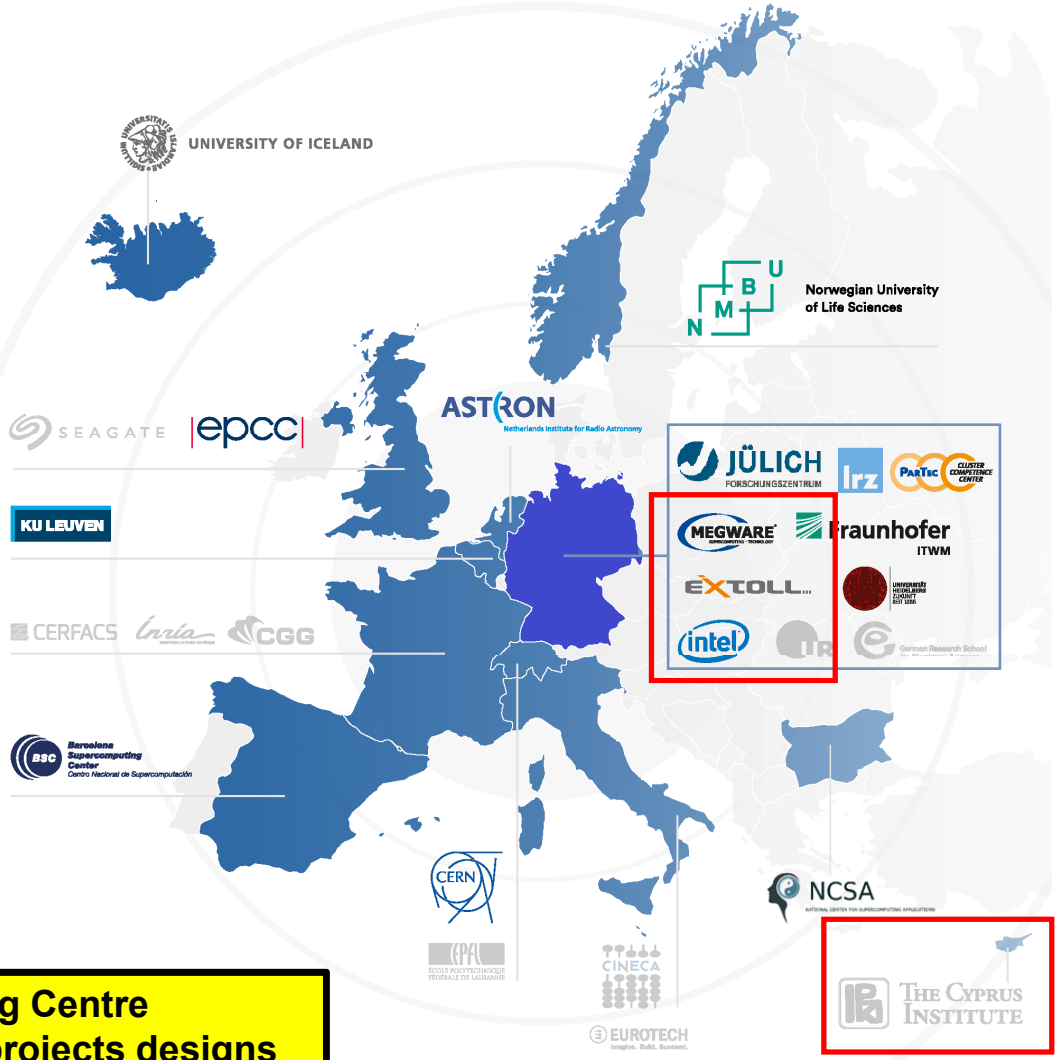


DEEP SERIES OF PROJECTS

EU Projects Driven by Co-Design of HPC Applications

- 3 EU Exascale projects
  - DEEP
  - DEEP-ER
  - DEEP-EST
- 27 partners
  - Coordinated by JSC
- EU-funding: 30 M€
  - JSC-part > 5,3 M€
- Nov 2011 – Jun 2020

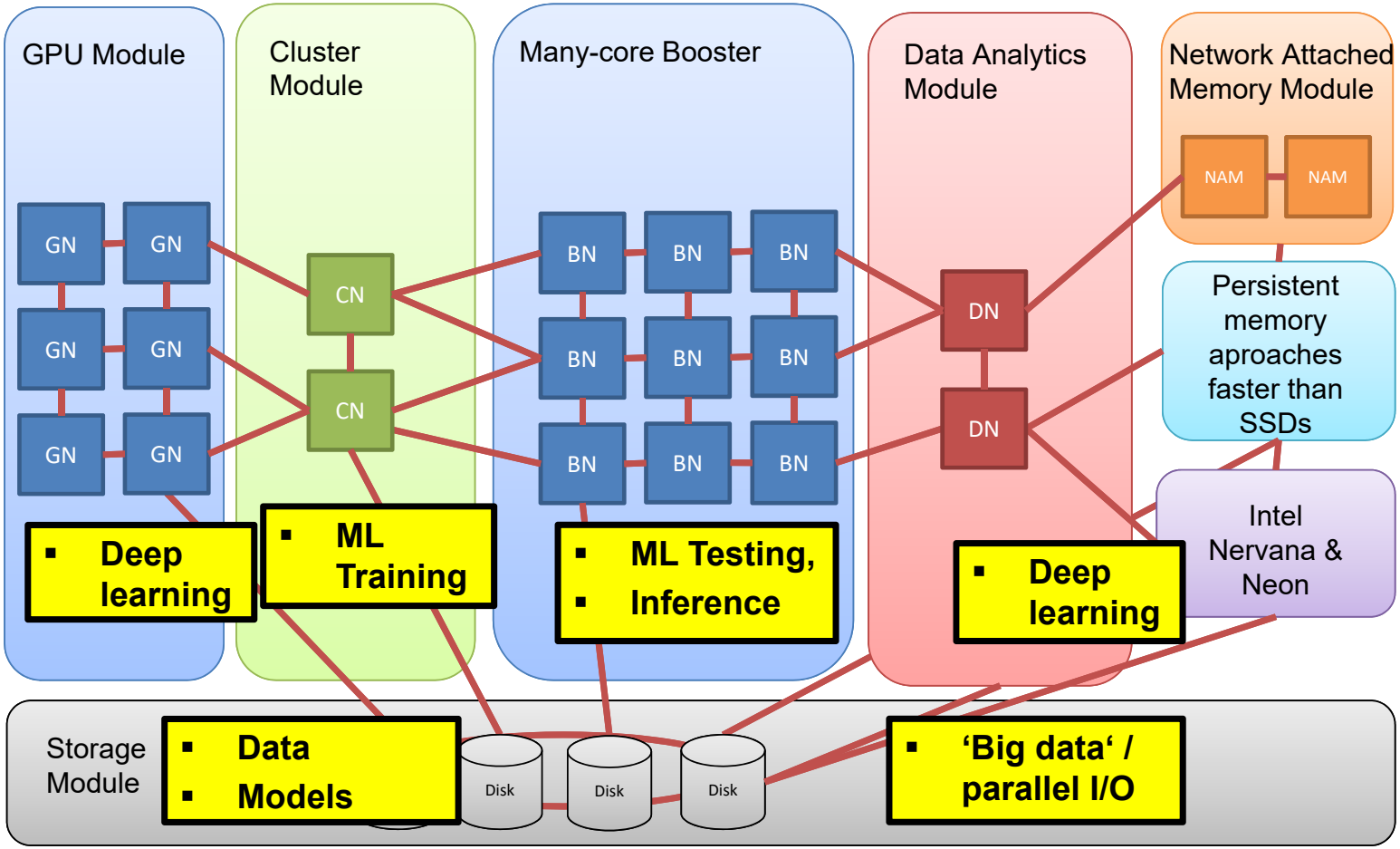
- **Strong collaboration with our industry partners Intel, Extoll & Megware**
- **Innovative HPC hardware like Intel Nervana Neon and persisten RAMs**
- **Juelich Supercomputing Centre implements the DEEP projects designs in its HPC production infrastructure**



[2] DEEP Projects Web Page

MODULAR SUPERCOMPUTING ARCHITECTURE

JSC Roadmap



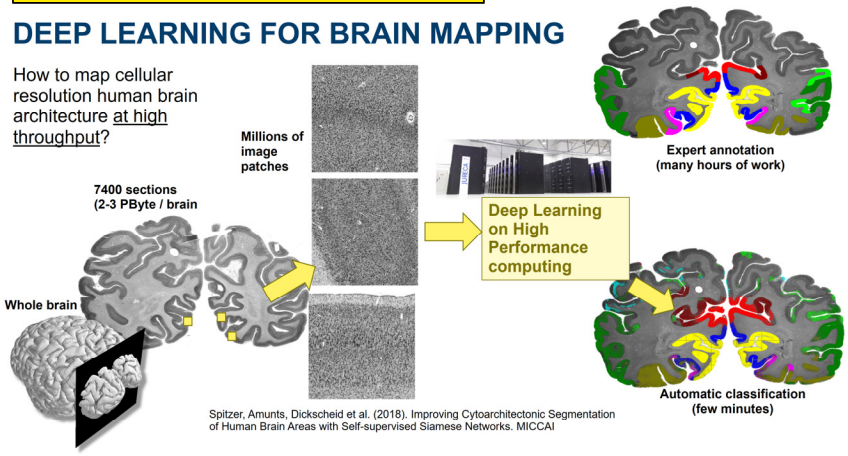
- **Innovative Ideas, e.g. trained models in memory**
- **Innovative memory, e.g. persistent RAM**
- **Innovative chips, e.g. use of deep learning optimized chip designs**

OTHER BIG DATA PROJECTS & APPLICATIONS

Prof. Katrin Amunts – INM-1

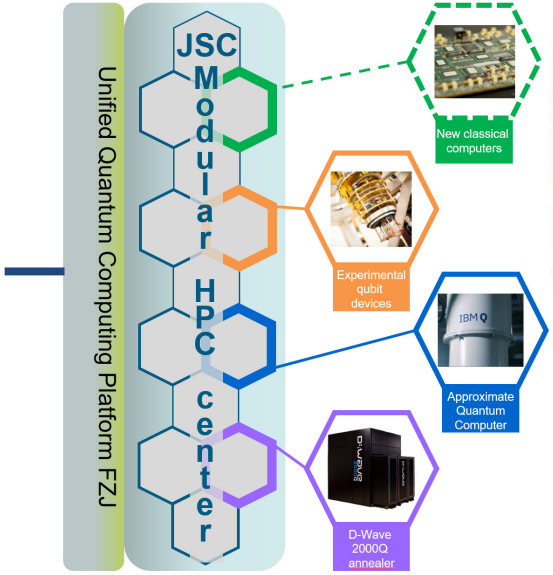
DEEP LEARNING FOR BRAIN MAPPING

How to map cellular resolution human brain architecture at high throughput?



Spitzer, Amunts, Dickscheid et al. (2018). Improving Cytoarchitectonic Segmentation of Human Brain Areas with Self-supervised Siamese Networks. MICCAI

Prof. Kristel Michielsen – JSC

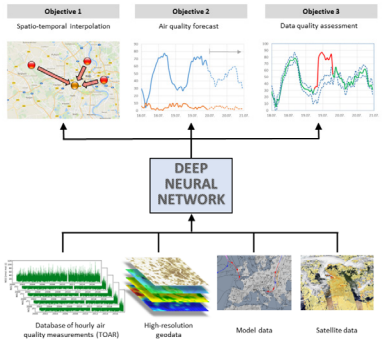


Dr. Martin Schultz – JSC

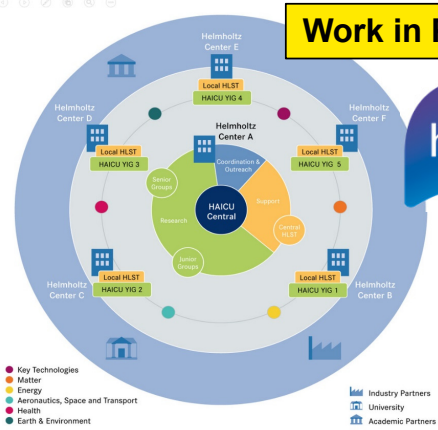
SPONSORED BY THE
Federal Ministry of Education and Research

IntelliAQ: Anwendung von Deep Learning zur Interpolation, Vorhersage und Qualitätskontrolle von Daten zur Luftqualität

erc
European Research Council
ERC-2017-ADG
#787576

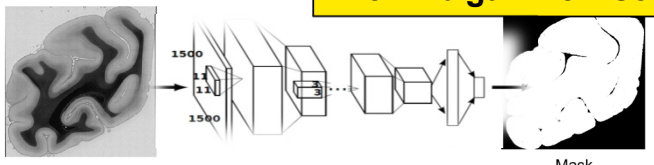


Work in Progress @ Helmholtz



haicu

Helmholtz Artificial Intelligence Cooperation Unit



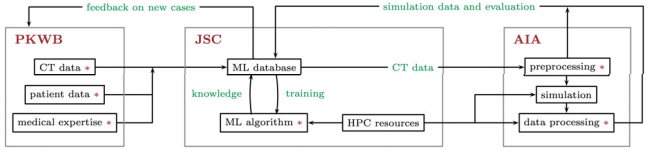
Image

Mask

Prof. Abigail Morrison – JSC

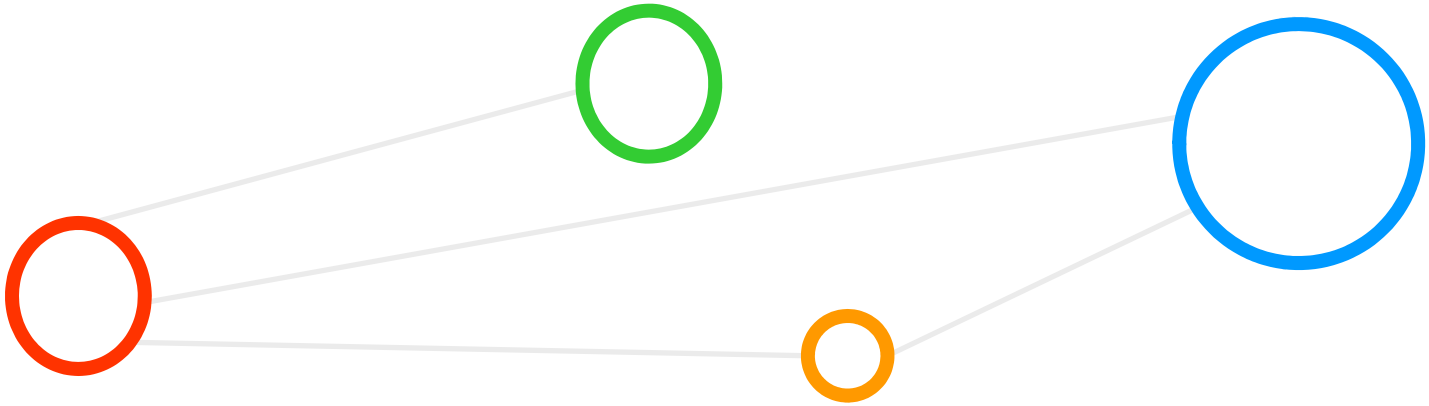
RHINO DIAGNOST

JARA|HPC



Dr. Andreas Lintermann – JSC/RWTH

REFERENCES



REFERENCES (1)

- [1] Helmholtz Association Web Page,
  Online: https://www.helmholtz.de/en/
- [2] DEEP Projects Web Page,
  Online: http://www.deep-projects.eu/
- [3] SMITH Projects Web Page,
  Online: http://www.smith.care
- [4] Alfred Winter et al., 'Smart Medical Information Technology for Healthcare (SMITH) – Data Integration based on Interoperability Standards', submitted to Journal of Methods, 2018, to appear
- [5] AIXCAPE Web Page,
  Online: http://www.aixcape.org/association
- [6] JSC Industry Relations Team (IRT) @ Juelich Supercomputing Centre,
  Online: http://www.fz-juelich.de/ias/jsc/EN/Expertise/IndustryRelations/_node.html
- [7] OpenFOAM Web Page,
  Online: https://www.openfoam.com/
- [8] M. Riedel, 'Deep Learning using a Convolutional Neural Network', Ghent University, Invited YouTube Tutorial,
  Online: https://www.youtube.com/watch?v=gOL1_YIosYk&list=PLrmNhuZo9sgZUdaZ-f6OHK2yFW1kTS2qF

REFERENCES (2)

- **[9] SoccerWatch.TV,**
  Online: https://soccerwatch.tv/
- **[10] Smart Data Innovation Lab (SDIL),**
  Online: https://www.sdil.de/en/
- **[11] European Commission, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, EC, 2018, 2nd May**
  Online: https://ec.europa.eu/commission/sites/beta-political/files/communication-modern-budget-may2018_en.pdf?utm_source=POLITICO.EU&utm_campaign=e3a8a86cc6-EMAIL_CAMPAIGN_2018_05_02&utm_medium=email&utm_term=0_10959edeb5-e3a8a86cc6-189710085

ACKNOWLEDGEMENTS

Previous & current members of the High Productivity Data Processing Research Group

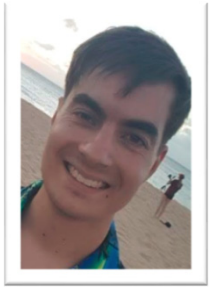
PD Dr.
G. Cavallaro

Senior PhD
Student A.S. Memon

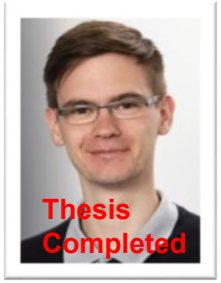
Senior PhD
Student M.S. Memon

PhD Student
E. Erlingsson

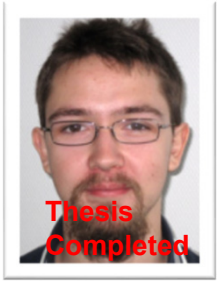
PhD Student
S. Bakarat



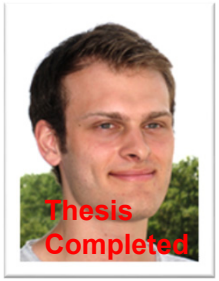
MSc Student
G.S. Guðmundsson
(Landsverkjun)



Dr. M. Goetz
(now KIT)



MSc M.
Richerzhagen

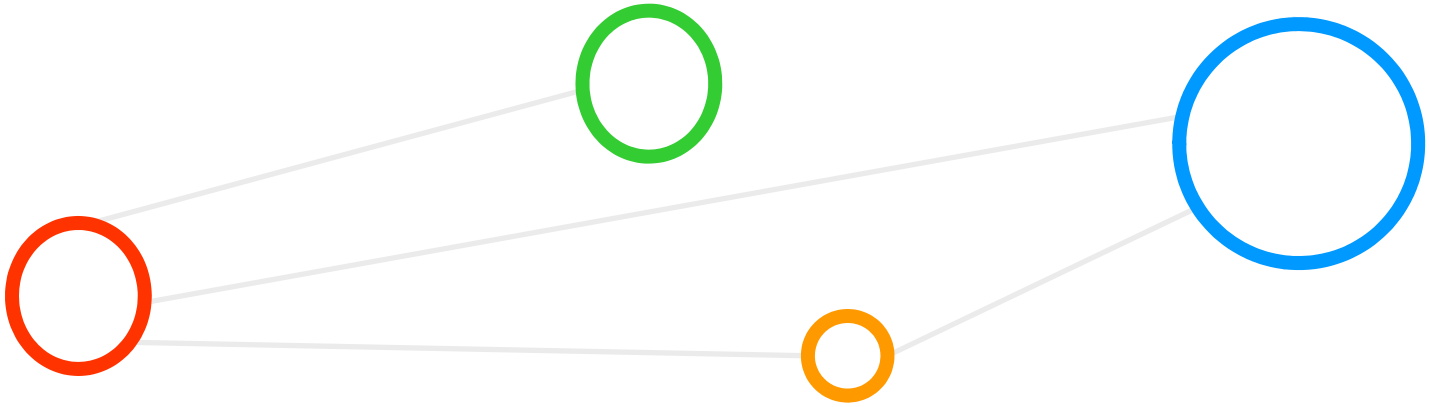


MSc
P. Glock
(now INM-1)

MSc
C. Bodenstein
(now Soccerwatch.tv)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 763558

BACKUP SLIDES



HPC & DATA SCIENCE: A FIELD OF CONSTANT EVOLUTION

Perspective: Floating Point Operations per one second (FLOPS or FLOP/s)

1.000.000 FLOP/s

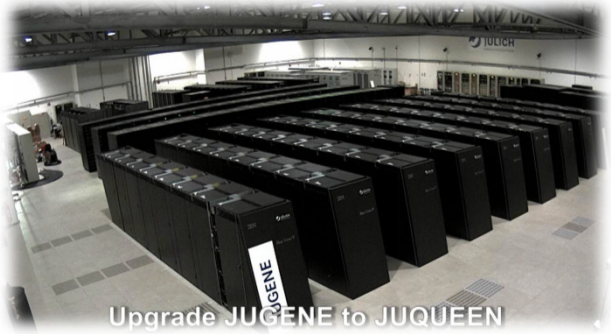
~1984

© Photograph by Rama, Wikimedia Commons

- 1 GigaFlop/s = 10^9 FLOPS
- 1 TeraFlop/s = 10^{12} FLOPS
- 1 PetaFlop/s = 10^{15} FLOPS
- 1 ExaFlop/s = 10^{18} FLOPS

1.000.000.000.000.000 FLOP/s

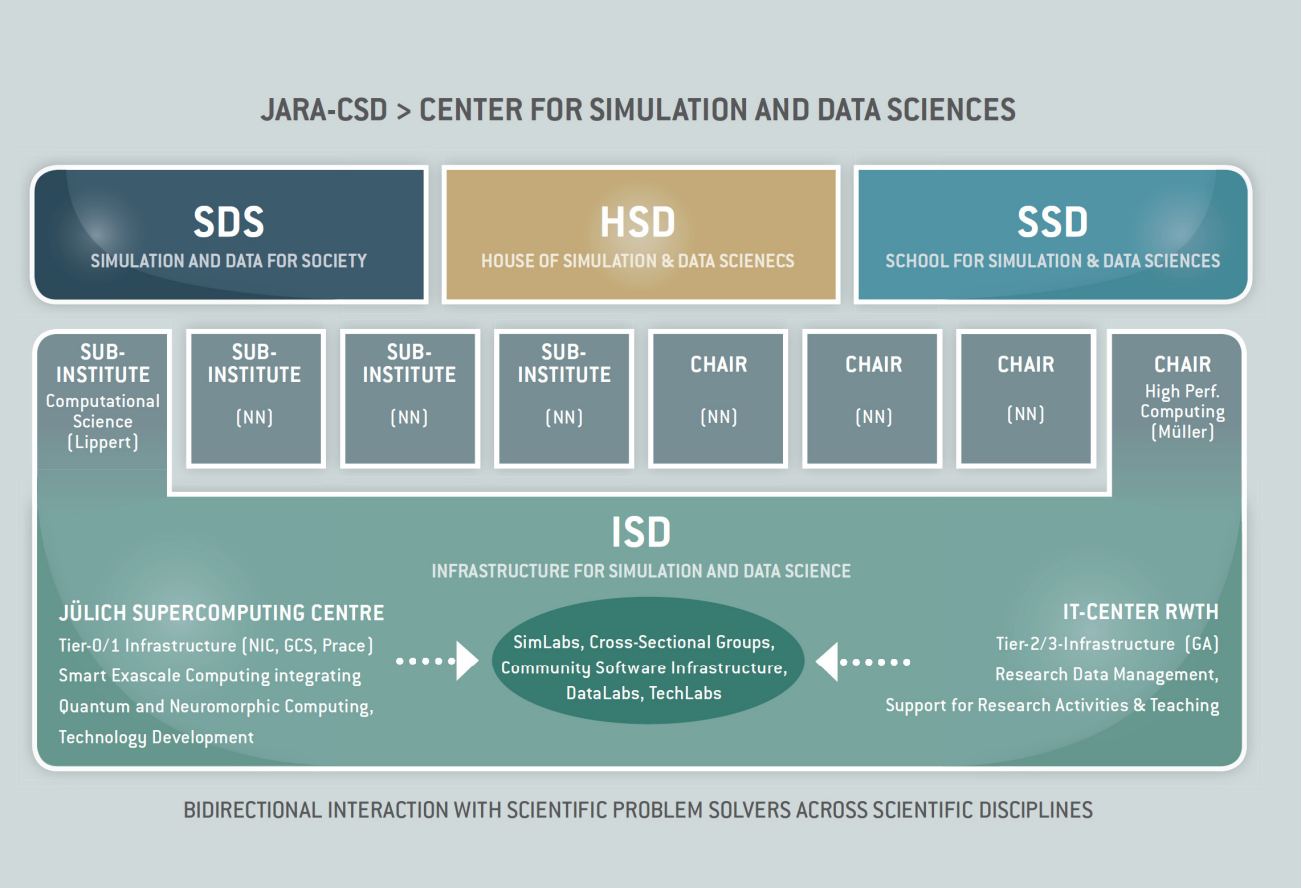
~295.000 cores~2009 (JUGENE)



>5.900.000.000.000.000 FLOP/s
~ 500.000 cores
~ 2017

JUELICH AACHEN RESEARCH ALLIANCE (JARA)

Work in Progress & Developments of a Center for Simulation and Data Sciences (CSD)

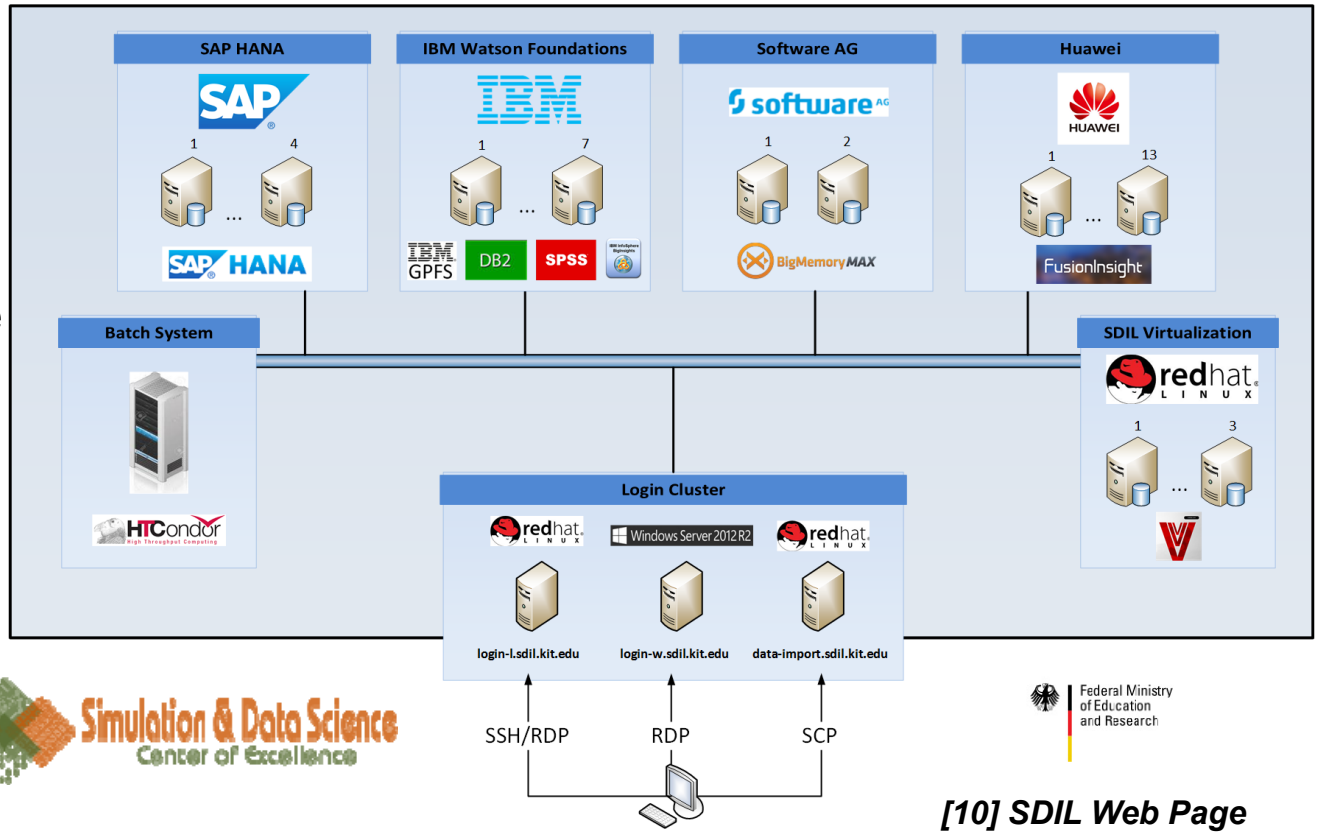


INDUSTRY EXAMPLES – TRANSLAB & TOOLS

Smart Data Innovation Lab

Joint Use of Smart Data Innovation Lab (SDIL) Platform

- Technology Platform for Data Analytics
  - Key technologies from vendors w.r.t. commercial parallel & scalable machine learning tool platforms
  - SAP Hana, IBM DB2 & SPSS, Software AG BigMemory MAX, Huawei FusionInsight, etc.
  - Data-driven SIMDAS projects can leverage the platform (small proposal needed / case)



Simulation & Data Science Center of Excellence

Federal Ministry of Education and Research

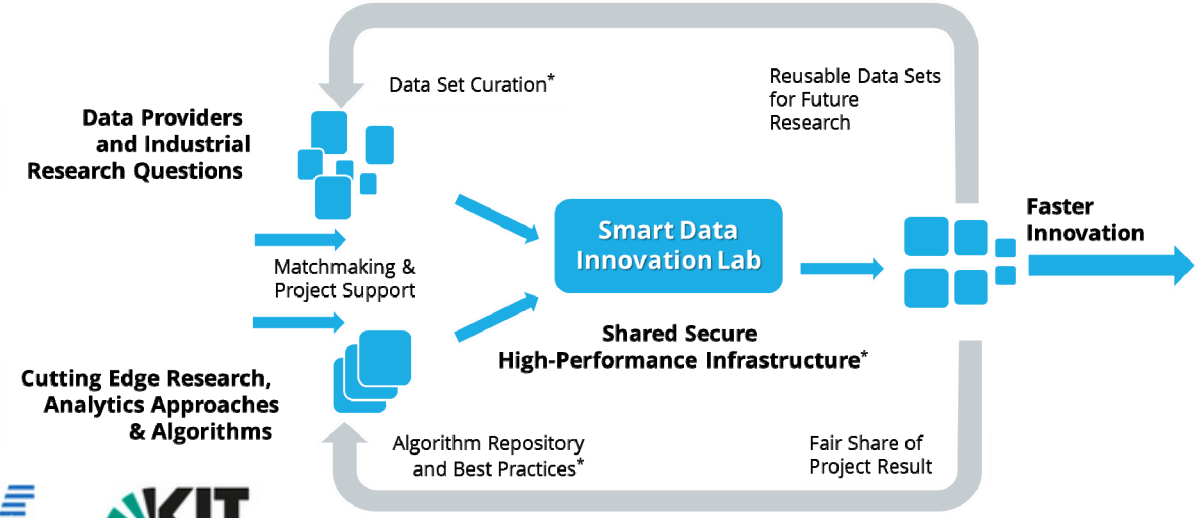
[10] SDIL Web Page

INDUSTRY EXAMPLES – TRANSLAB & TOOLS

Smart Data Innovation Lab

Joint Use of Smart Data Innovation Lab (SDIL) Platform

- SDIL Partners
  - Key players in German industry
  - Head of community Medicine (Prof. M. Riedel & Prof. A. Schuppert; both SMITH ASIC Use Case partners)



BAYER · BOSCH · DFKI · EnBW · IBM · KIT Karlsruher Institut für Technologie · Fraunhofer IAIS · HUAWEI · JÜLICH Forschungszentrum · SAP · SIEMENS · software AG · TRUMPF · bitkom · BDV · GI · SDSC BW · LNI4.0 Labs Network Industrie 4.0

Federal Ministry of Education and Research

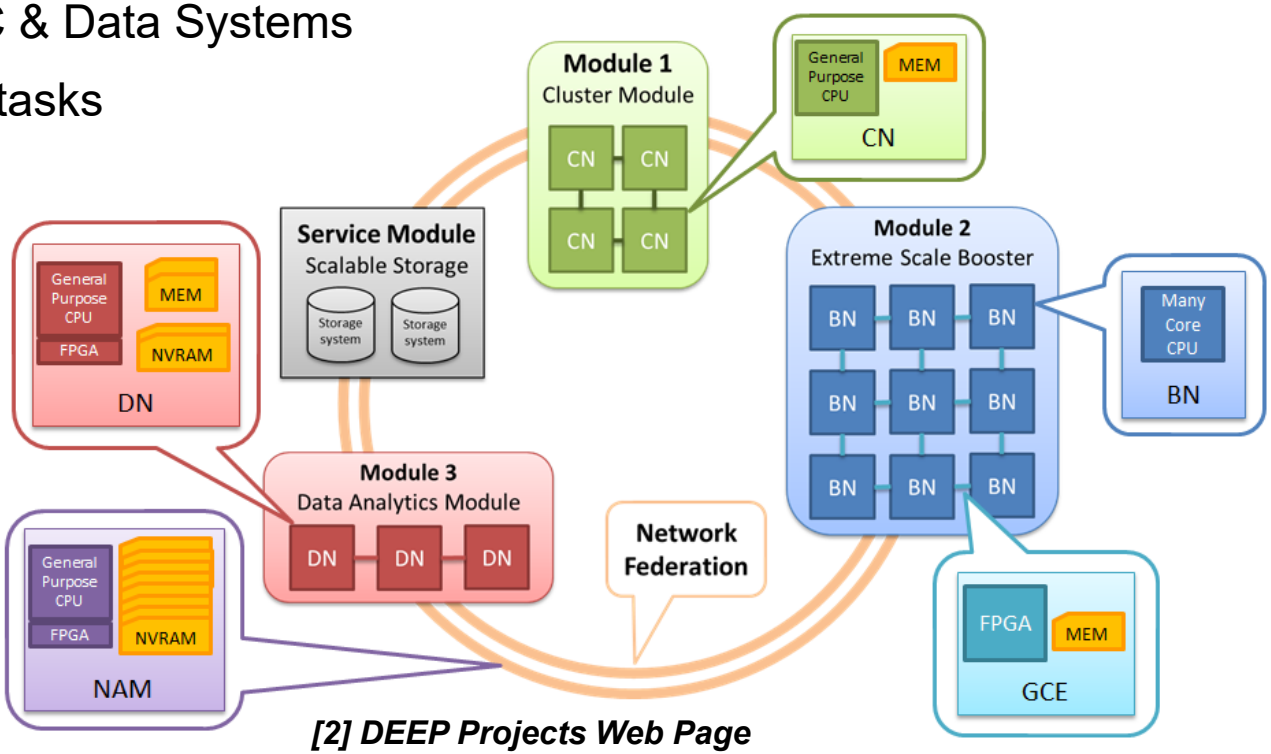
[10] SDIL Web Page

TOWARDS EXASCALE

DEEP Projects Series

- Flop/s metric will become increasingly less(!) relevant
  - Driven by application co-design of HPC & Data Systems
  - Support for less regular computational tasks
  - Significantly larger memory footprint
  - Extreme data processing capabilities
  - Improved/optimized data transport capabilities & specialized analytics
  - Scalable visualisation capabilities
  - Management of complex work-flows

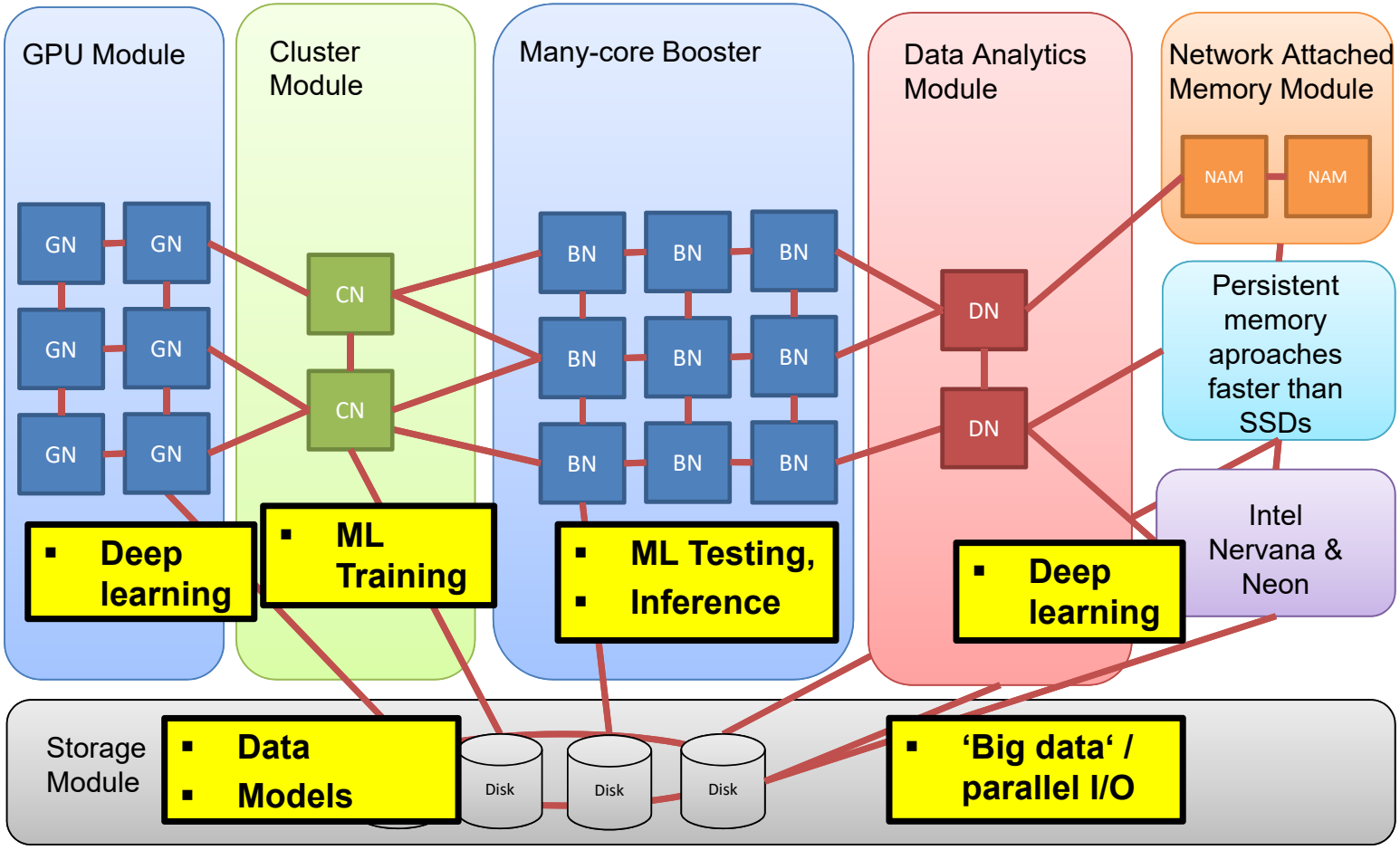
- **One plausible answer to those facts is the modular supercomputer archiecture driven by the JSC & DEEP projects**



[2] DEEP Projects Web Page

MODULAR SUPERCOMPUTING ARCHITECTURE

JSC Roadmap



- **Innovative Ideas, e.g. trained models in memory**
- **Innovative memory, e.g. persistent RAM**
- **Innovative chips, e.g. use of deep learning optimized chip designs**

PROJECT EXAMPLES – HEALTH AREA

smith
Smart Medical Information Technology for Healthcare

Bayer AG & RWTH Aachen + University Hospital & Forschungszentrum Juelich

- SMITH ASIC Use Case
  - ASIC: Algorithmic Surveillance of Intensive Care Unit (ICU) Patients & Focus on Acute Respiratory Distress Syndrome (ARDS)
  - University Clinic Aachen (UKA): Machine Learning for patient stratification & virtual ICU patient & risk patterns
  - FZJ: Parallel & Scalable Machine Learning & Statistical Modelling via HPC
  - Bayer AG: clinical trial optimization in prevention studies & virtual ICU patients & organ models

UKA-Schuppert	FZJ-Riedel	Bayer-Lippert	UKA-Haferkamp
Phänotypen Identifikation und Integration in PheP VP Identifikation und Integration in das ASIC Expertensystem (WP 5/6/8)	Phänotypen und VP Identifikation und Modellreduktion (WP 5/8)	VP Identifizierung, Erstellung des VP Modells und Validierung Patienten Subgruppen-Stratifizierung (WP 6, 8)	DIZ-Datenverfügbarkeit Anbindung Forschungsdatenbank

UNIKLINIK RWTH AACHEN JÜLICH Forschungszentrum BAYER Bayer UNIKLINIK RWTH AACHEN

[3] SMITH Project Web Page

[4] SMITH Methods Journal

SPONSORED BY THE
Federal Ministry of Education and Research

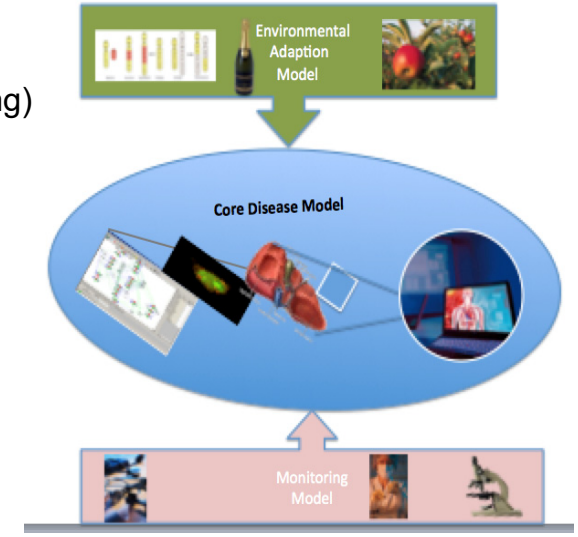
PROJECT EXAMPLES – HEALTH AREA

smith
Smart Medical Information
Technology for Healthcare

SPONSORED BY THE
Federal Ministry of Education and Research
MEDIZIN INFORMATIK INITIATIVE

UNIKLINIK RWTH AACHEN
ukb universitäts klinikum bonn
UKD Universitätsklinikum Düsseldorf
Universitätsklinikum Essen

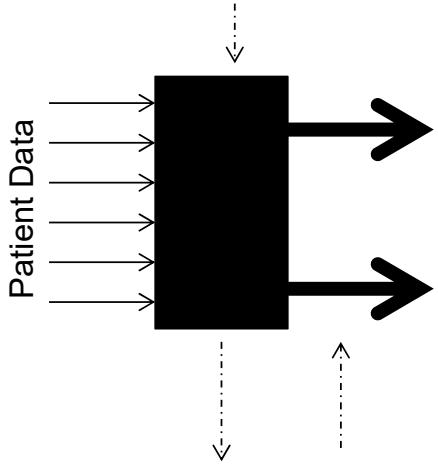
Patient + state paramters:
individualisation
(Machine Learning)



generic model
(mechanistic)

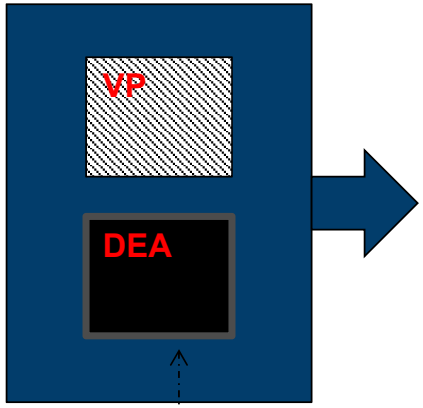
adaption of generic model
to available data (Machine Learning)

Unsupervised Patient Stratification
- Dynamic clustering
- Critical state detection



Patient subgroups
& classifiers

Predictive modelling Machine for Algorithmic Surveillance of ICU Patients



Machine Learning, Patient association,
Subgroup specific prediction

Prognosis for
Individual patient

PROJECT EXAMPLES – HEALTH AREA

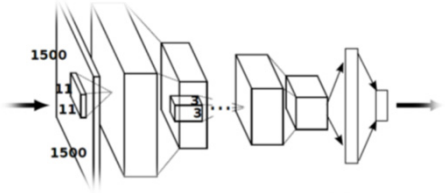
smith
Smart Medical Information Technology for Healthcare

Bayer AG & RWTH Aachen + University Hospital & Forschungszentrum Juelich

- MCMC with VP models
  - 1 run for a patient ~ 2 sec (1 core)
  - 10^6 runs required for MCMC
  - 1 patient ~ 1000 core-h
  - Good news: 100% parallel → nice scalability and useful for booster
  - More data is helpful for more concrete feature selection w.r.t. ARDS

- **Markov Chains Monte Carlo (MCMC) with Virtual Patient models**
- **In clinical practise not feasible with today's computer technology => model reduction is necessary**
- **Virtual Patient (VP) model mapped into a deep learning network**

- Compute-intensive part of the Virtual patient model will be mapped onto a deep-learning (DL) network
  - DL-network has is numerically hard to train, but fast to simulate
  - Mapping strategy has been evaluated and applied at partner Bayer
  - Requires HPC (scanning of the full parameter space)



PROJECT EXAMPLE – ARTIFICIAL INTELLIGENCE

ON4OFF

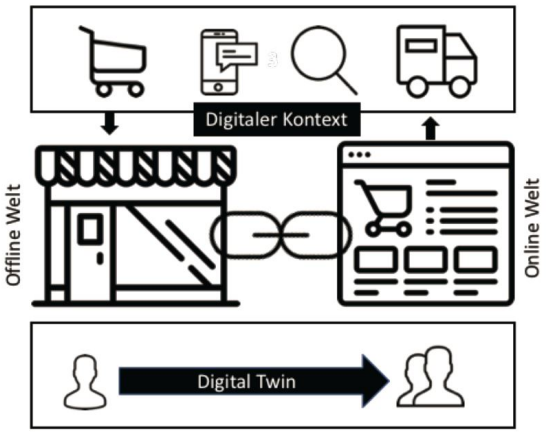
UNIVERSITÄT DUISBURG ESSEN

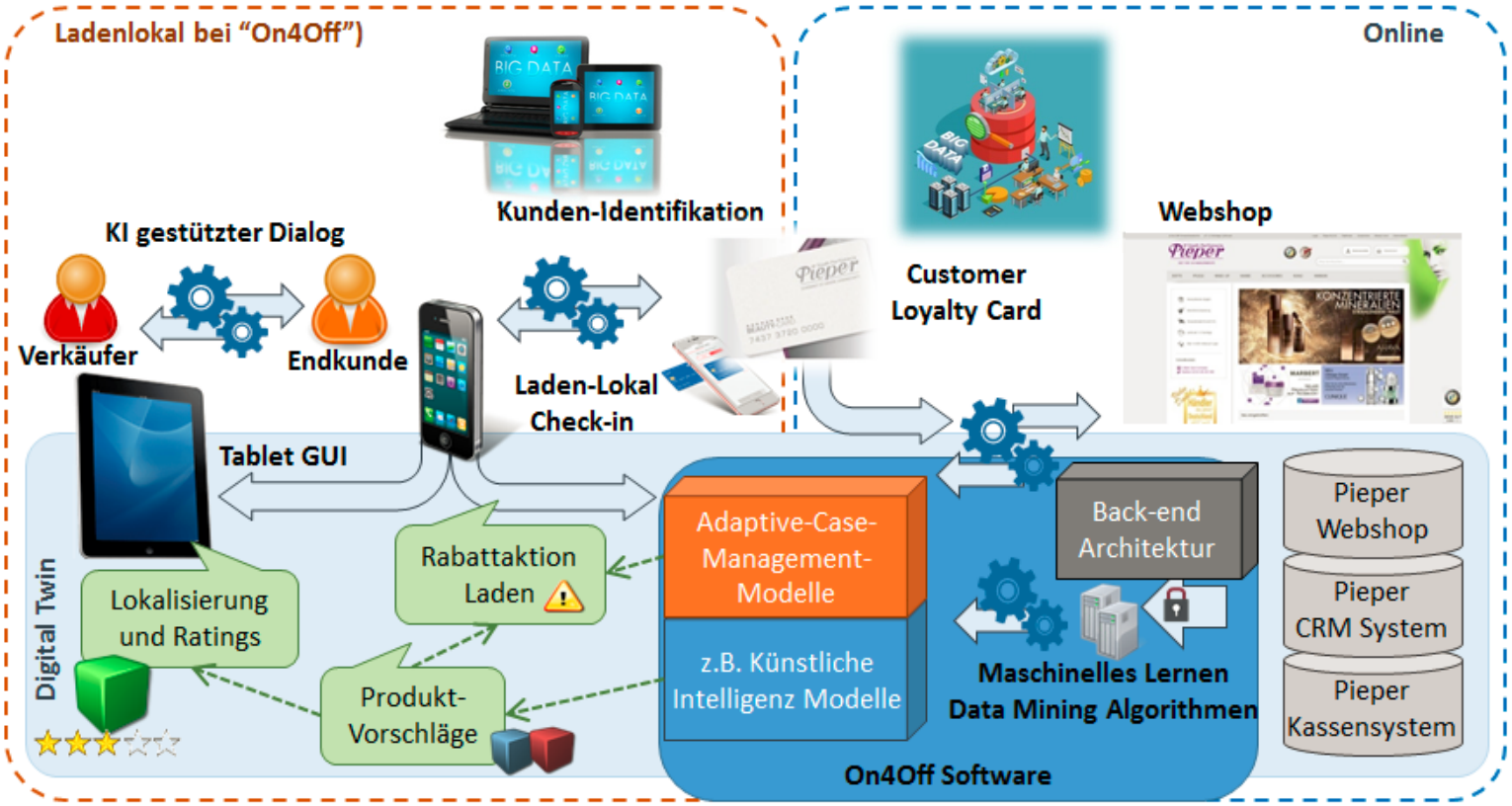
IN-telegence

adesso

Stadt-Parfümerie Pieper SEIT 1931 IN FAMILIENBESITZ

Hochschule Niederrhein
University of Applied Sciences





Leitmarkt Agentur.NRW

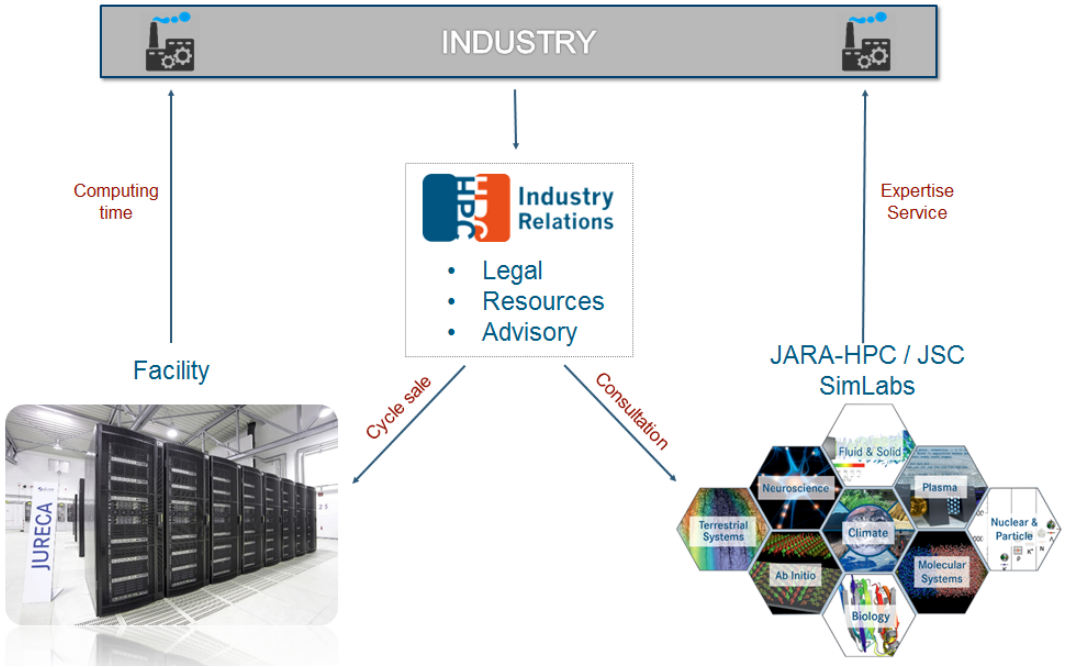
EUROPÄISCHE UNION
Investition in unsere Zukunft
Europäischer Fonds
für regionale Entwicklung

EFRE.NRW
Investitionen in Wachstum
und Beschäftigung

INDUSTRY RELATIONS TEAM (IRT) @ JSC

Examples: Selling Computing Time & Offering Code Optimization

- **SIEMENS**
  - Long-term cooperation with Siemens Power & Gas Department
  - Simulation of combustion processes in turbine systems
  - Computing time on JSC Jureca HPC system
  - Take advantage of application support team @ JSC
  - Bilateral cooperations and partners in big publicly funded (Germany BMWI) project consortium
- **Outotec**
  - Global leader in minerals & metals processing technology
  - Computing time on JSC Jureca HPC System
  - OpenFOAM computational fluid dynamics (CFD) computations
  - Take advantage of application support team @ JSC



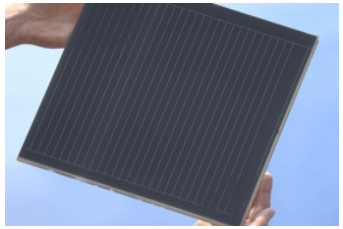
[6] JSC Industry Relations Web Page

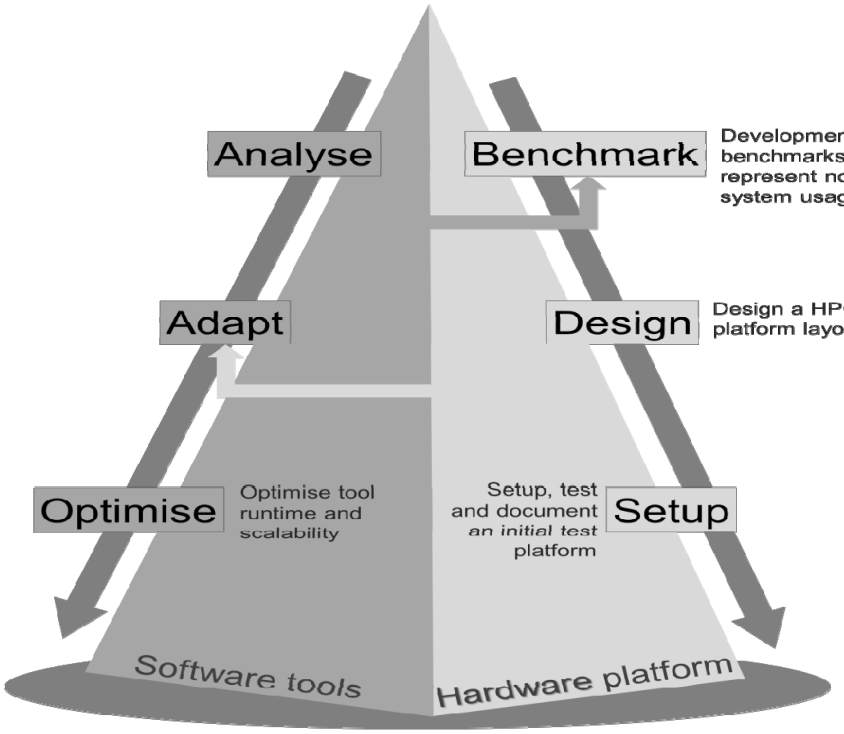
[7] OpenFOAM Web Page

INDUSTRY RELATIONS TEAM (IRT) @ JSC

Examples: Selling Computing Time & Offering Code Optimization

- TenneT Taking power further
  - One out of four German Transmission System Operators (TSOs) & designs, builds, and operates high voltage grids
  - Selected Consultation & Expertise Services from JSC included software & hardware guidance & support
  - Software: code analysis, optimization plan, work-flows
  - Hardware: support of a purchase decision & cluster testing





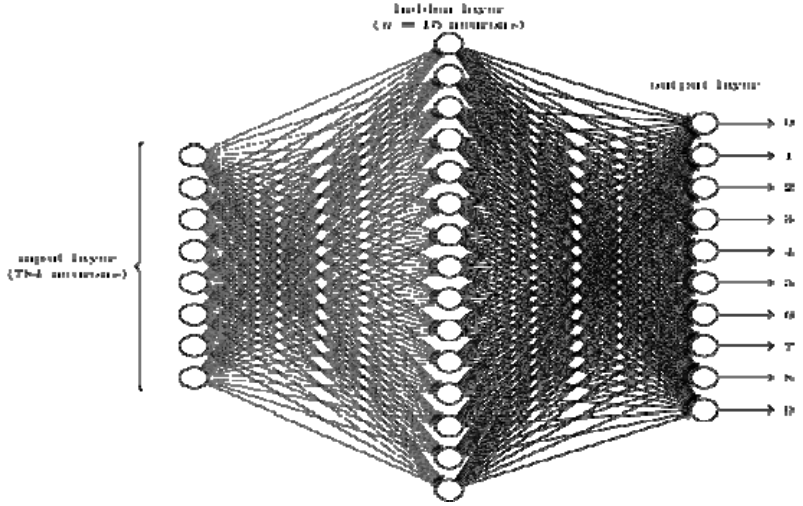
- **Benefit for FZJ/JSC: Work on real industry problem with high societal relevances & follow-up projects discussed like German BMBF projects**

[6] JSC Industry Relations Web Page

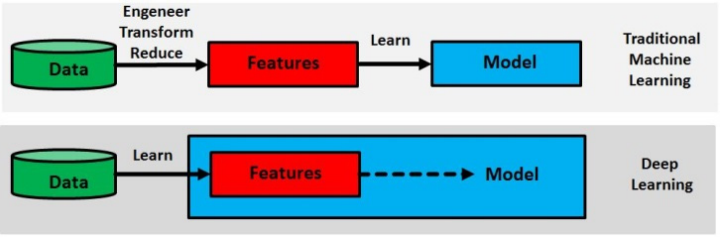
INNOVATIVE DEEP LEARNING TECHNOLOGIES

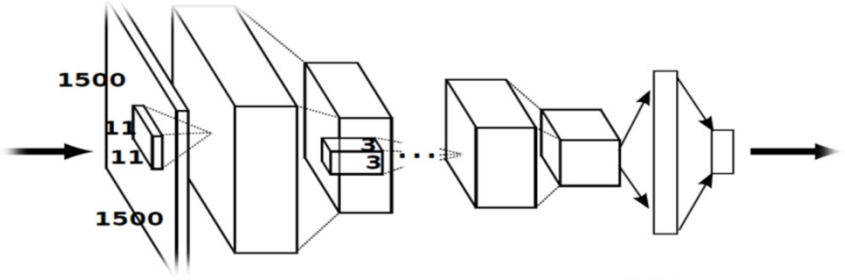
Short Overview & Role of Team Deep Learning for SIMDAS & Juelich Supercomputing Centre

- Innovative & disruptive approach



[8] M. Riedel, Invited YouTube Tutorial on Deep Learning, Ghent University

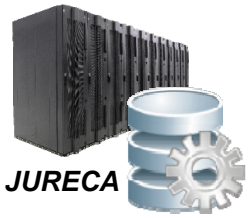


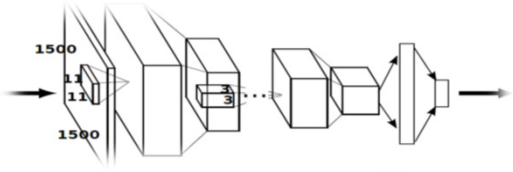


- **Provide deep learning tools that work with HPC machines (e.g. Python/Keras/Tensorflow)**
- **Advance deep learning applications and research on HPC prototypes (e.g. DEEP-EST, etc.)**
- **Engage with industry (industrial relations team) & support SMEs (e.g. Soccerwatch)**
- **Offer tutorials & application enabling support for commercial & scientific users (e.g. YouTube)**

RELATIONSHIP BIG DATA & ARTIFICIAL INTELLIGENCE (AI)

#GPUs	images/s	speedup	Performance per GPU [images/s]
1	55	1.0	55
4	178	3.2	44.5
8	357	6.5	44.63
16	689	12.5	43.06
32	1230	22.4	38.44
64	2276	41.4	35.56
128	5562	101.1	43.45





Model Performance / Accuracy

small training sets

manual feature engineering changes ordering

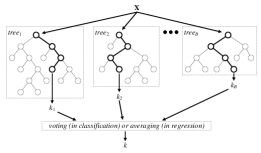
Large Deep Learning Network

Medium Deep Learning Network

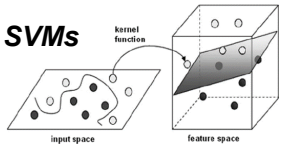
Small Neural Network

Traditional Learning Models

Random Forests



SVMs



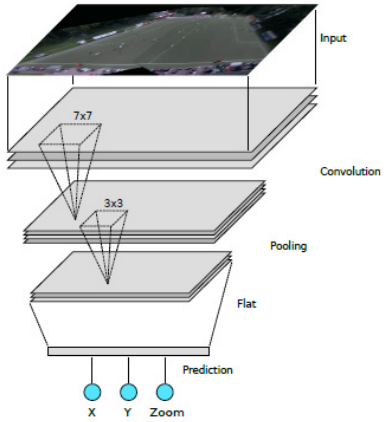
Dataset Volume

→ 'Big Data'

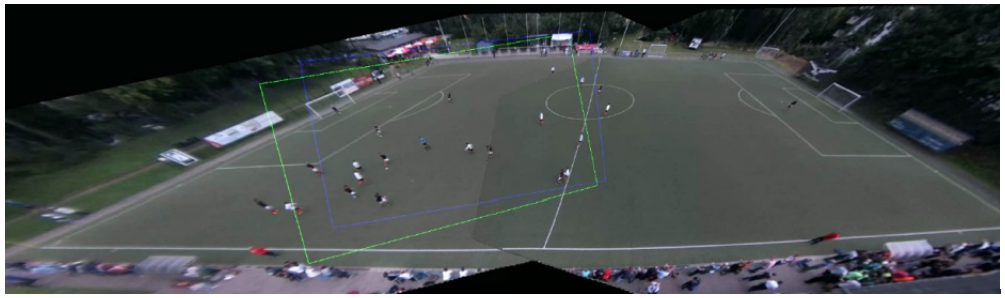
INDUSTRY EXAMPLES – INNOVATIVE START-UPS

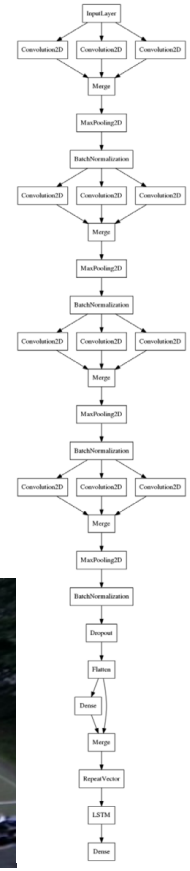
Collaboration in Applying Deep Learning in Commercial Scenarios & Small Start-Up Guidance

- SoccerWatch.TV
  - Start-up: created/joined by a ‘exit-ing‘ PHD Student @ JSC
  - Besides upper leagues: 80k matches/week
  - Recording too expensive (amateurs) with camera man needed
  - Approach: Find X,Y center and zoom on panorama camera using Deep Learning
  - Investor grant (1,5 mio €) from Adesso AG



[9] SoccerWatch.TV Web page





- **Letter of intent/support already requested and relevant joint selected work elements have been already discussed**
- **Further German BMBF project has been submitted (NRW-HUB) with relevance to SIMDAS & Retail (with Adesso)**

BIG DATA STORAGE INFRASTRUCTURE @ JSC

- JUST Storage Cluster
  - IBM Spectrum Scale file system (GPFS)
  - 75 PB gross capacity
  - 5th generation
  - Parallel access
- Tape Libraries
  - Automated cartridge systems
  - 300 PB
  - 3 libraries (in 2 buildings)
  - 60 tape drives
  - 35,000 tapes

JSC & CO-DESIGN APPROACH

Drive Technology Innovation in Different Roles

- Exascale Labs
  (or Competence Centres)
  - Long-term collaboration with suppliers
  - POWER Acceleration and Design Center
  - Collaboration between Forschungszentrum Juelich, IBM and NVIDIA
  - Mission statement: Provide support to scientists and engineers to target the grand challenges facing society in the fields of energy & environment, information & health care
- Co-Design Projects
  - E.g. DEEP projects & application use cases

(Selected JSC collaboration partners)

THANKS

Talk shortly available under www.morrisriedel.de