

ErUM-Data Consortium

Federated Unified Storage Environment

ErUM Data-Hub Onboarding Meeting

Torsten Harenberg, Christian Voß
Zoom, November 25th 2025

Involved Partners

Sites and Communities Participating in FUSE

Following sites are participating or associated in FUSE

- RWTH Aachen, AIP Potsdam, DESY, University of Göttingen, HAW Hamburg, LMU Munich, University of Mainz, University of Wuppertal; associated: University of Bonn, CERN, KIT

This covers the communities

- Particle Physics (incl. theory), Astronomy, Computer Science
- Associated: Photon Science and Accelerator Research and Operations

Cover two main work packages

- “Federated Storage with a multi-site common directory structure” (Aachen, DESY, Göttingen, HAW, Mainz, LMU, Wuppertal)
- “Caching Services and Cloud Storage Integration” (AIP, DESY, LMU, Wuppertal)

Main contacts for further questions and collaboration

- Torsten Harenberg (harenberg@physik.uni-wuppertal.de), Christian Voß (christian.voss@desy.de)

Work Packages in FUSE

Detailed Overview of Individual Task Areas

"Federated Storage with a multi-site common directory structure"

- Development and deployment of a central site acting as a hub to plug in external storage infrastructure (based on dCache)
- Integration of storage for "permanent" data (hook up storage at universities, access and namespace managed by hub site)
- Integration of "opportunistic" storage (transparent caches at HPC sites, tape resources at sites without middleware support)
- Monitoring and self-healing capabilities for such a federated system

"Caching Services and Cloud Storage Integration"

- Further Development of existing and independent caching solutions with focus on further developments on monitoring
- Bridging the gap between Cloud Compute data access standards and existing scientific storage infrastructure allowing cloud-like access to data stored on classical HPC or Grid-SE storage infrastructure
- Adding Sync&Share capabilities to analysis facilities (requested by the not funded FAIRUM consortia) → discuss future of this task area, maybe limit support/capabilities to selected clusters, e.g. NAF at DESY and LMU user cluster

Thank you

Adversarial Learning for Robust & Physics-Preserving AI in Particle Physics

Partners

- RWTH Aachen: Wiebusch (IceCube), Schmidt (CMS)
- Uni Bonn: Flek (NLP), Schott (GravNet, ATLAS)

Motivation

- Adversarial learning (AL) enhances neural-network robustness to noise and modeling uncertainties
- GANs and other adversarial generative models offer fast, energy-efficient simulation alternatives.

Project Vision

- Advance Adversarial AI for Physics (AALearning) through physics-preserving methods, systematic-uncertainty quantification, and safe generative modeling—with transferability to other scientific domains.



AirMon-VLT (Airglow Monitor for the VLT observations)

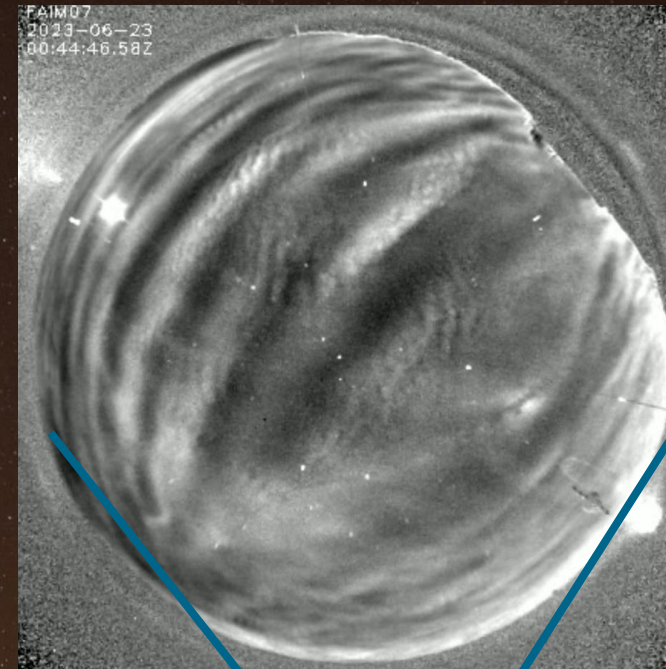
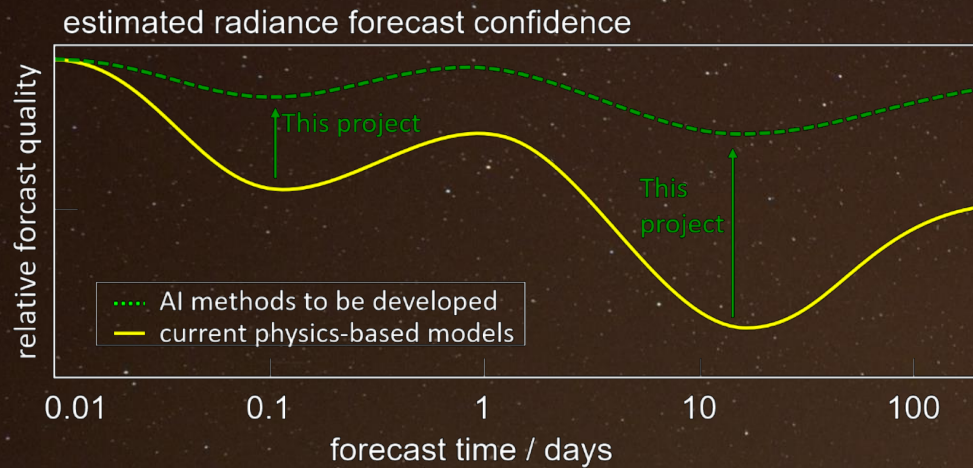
Patrick Hannawald¹, Rainer Lienhart²

and associated partners Alain Smette (ESO) and Jay Stephan (UKATC)

1. German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany

2. Universität Augsburg (UniA), Institute of Computer Science – Chair of Machine Learning and Computer Vision, Augsburg, Germany

With funding from the:



ErUM-Data

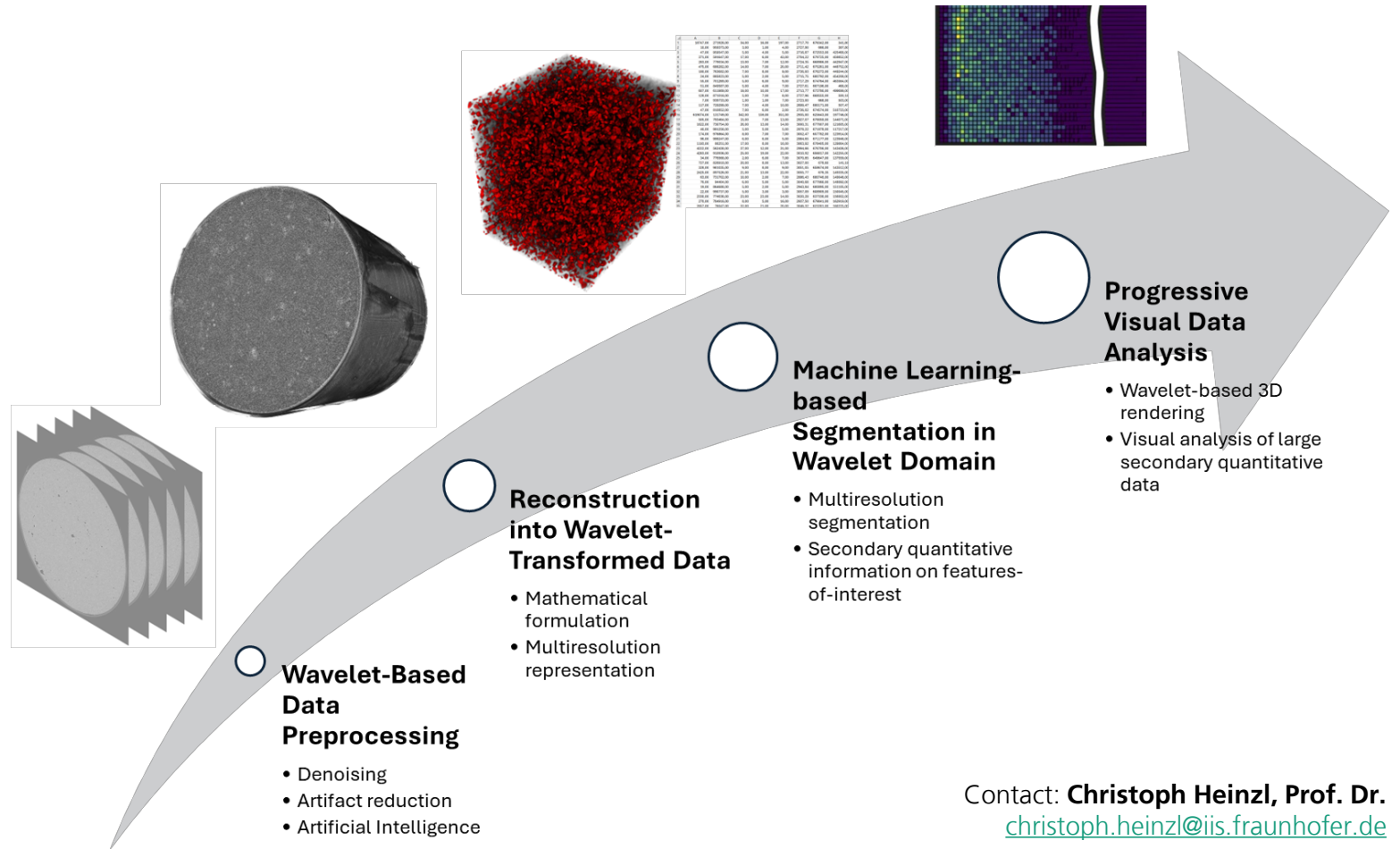
BeyondVoxels

BeyondVoxels

Hierarchical Analysis, Processing and Visualization of Large-Scale Synchrotron Radiation Imaging Data

Project Information

- **ErUM-Topic: Software and Algorithms**
- **Duration: 01.11.2025 - 31.10.2028**
- **Project Partners**
 - Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V.
 - University of Passau
 - Helmholtz-Zentrum Hereon
- **Associated Partners**
 - European Synchrotron Radiation Facility
 - Technische Universität Braunschweig
 - Helmholtz Zentrum Berlin
 - Humboldt-Universität zu Berlin



Contact: **Christoph Heinzl, Prof. Dr.**
christoph.heinzl@iis.fraunhofer.de

Contact

Christoph Heinzl, Prof. Dr.
Head of Research Group
Knowledge-based Image Processing and Visualization
Tel. +49 851 509 4475
Fax +49 851 509 3102
christoph.heinzl@iis.fraunhofer.de

Fraunhofer IIS/EZRT
Innstraße 43
D-94032 Passau
www.fraunhofer.de

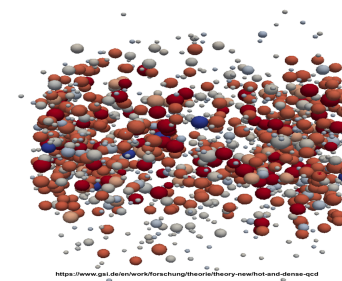
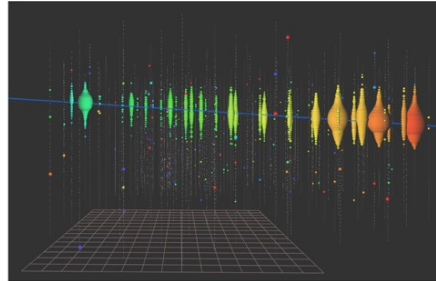
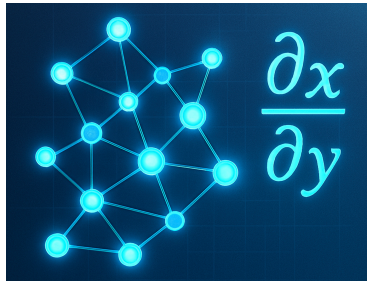
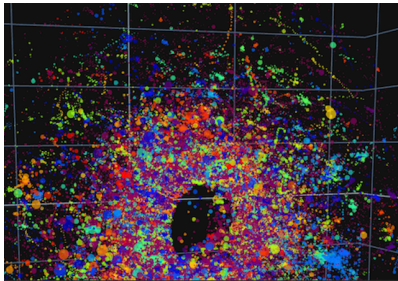
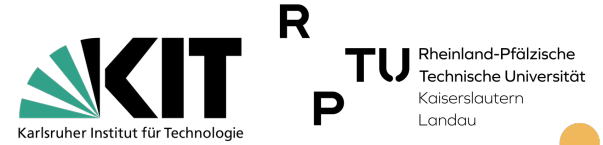


Entwicklungszentrum Röntgentechnik
des Fraunhofer-Instituts für
Integrierte Schaltungen IIS



Bridging Domains: AI-Driven Agnostic Reconstruction Frameworks for Irregular High-Dimensional Data (BRAID)

- Jan Kieseler (speaker), **KET**, KIT
- Nicolas Gauger, **Computer Science**, RPTU
- Christian Glaser, **KAT**, TU-Do
- Anastasios Belias, **KHuK**, GSI



- AI has demonstrated transformative power across a wide range of applications.
- BRAID aims at a paradigm change and to **unify reconstruction in particle, astro- and nuclear physics**.
- Goal: a flexible, detector-agnostic framework for irregular, high-dimensional data from experiments such as **HL-LHC, FAIR, and IceCube-(Gen2)**.
- Embed **physics constraints** and ensure **robustness** across varying detector conditions and data regimes.
- Integrate advanced, **gradient-driven dimensionality-reduction methods** to maximise efficiency and resource sustainability.
- Develop **open-source software** and cross-domain **public datasets** to support transparent, reproducible research.
- Demonstrate generalisability through applications beyond HEP, including **medical imaging and monitoring**.

CmarT

Development of Local Adaptive Tomography for the Study of Hierarchical Complex Matter

- multi-scale imaging using synchrotron radiation
- across multiple experimental stations
- optimized data acquisition and data flow
- interior tomography problem
- probe informed phase retrieval

Universität
Rostock



Fraunhofer
IMTE



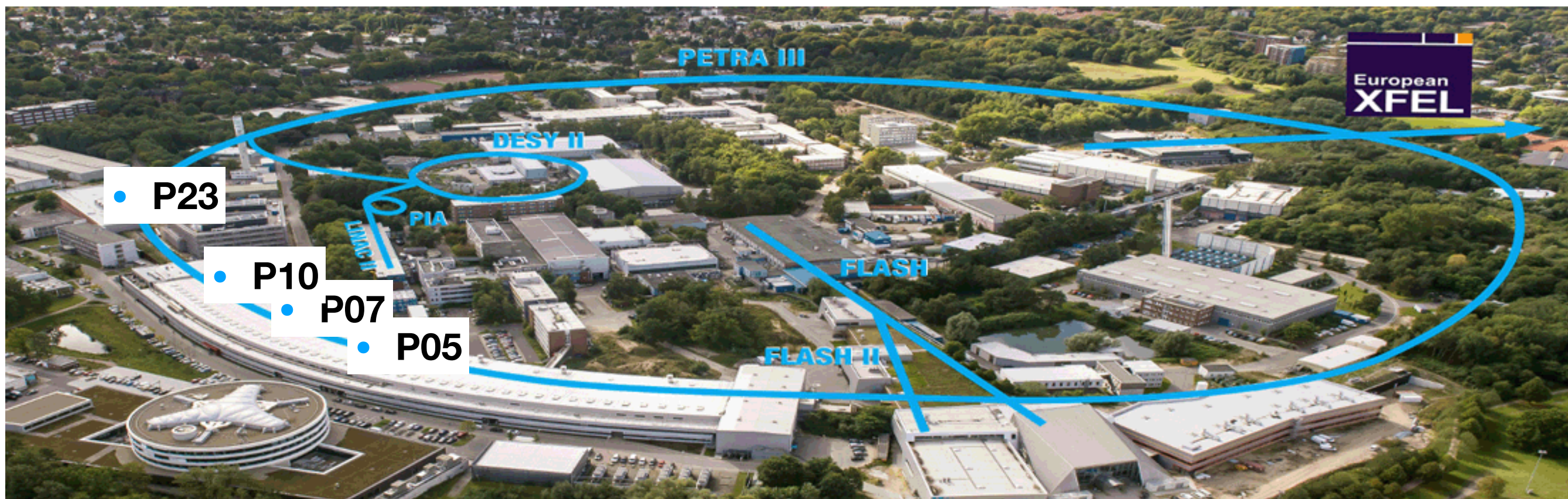
Helmholtz-Zentrum
hereon



Karlsruher Institut für Technologie



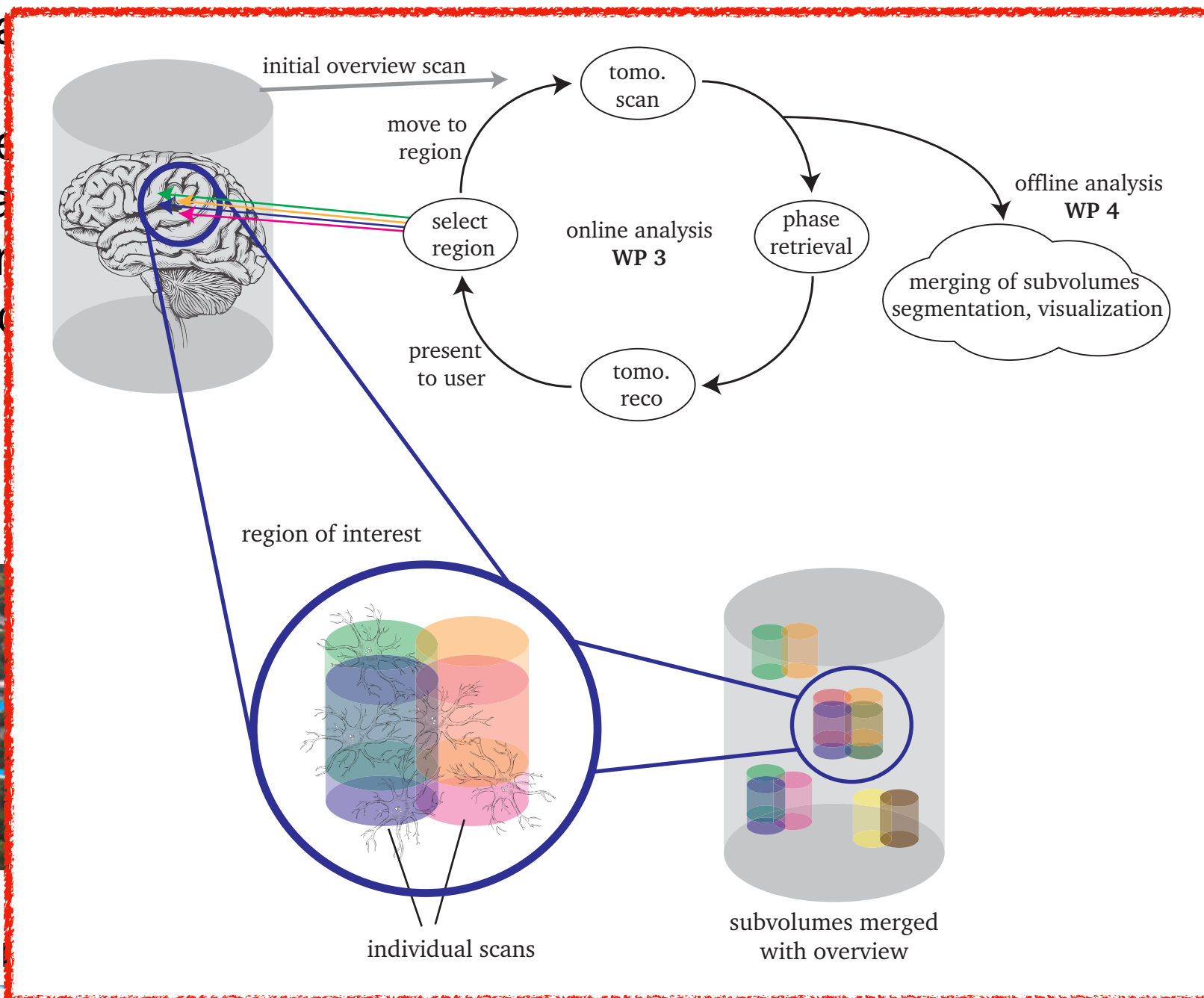
GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN



CmarT

Development of Local Adaptive Tomography for the Study of Hierarchical Complex Matter

- multi-scale imaging
- low radiation
- across multiple scales
- optimized data acquisition
- interior tomography
- probe informed



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN



• P23

• P10

ErUM-Data DEEP: Data Efficiency in Embedded Processors

We are DEEP!
erumdata-deep.de

Prof. Annika Eichler
(TU Hamburg-Harburg)



DESY

Prof. Klaus Desch
(University Bonn)



Prof. Torben Ferber*
Prof. Jürgen Becker
(Karlsruhe Institute of Technology)
*project coordination



Prof. Marco Gersabeck
(University Freiburg)

SICK AG



Dr. Karim Guirguis
(SiMa.ai)

- Develop missing tools and build up world-leading expertise for **advanced real-time ML** (μ s to minutes) on **heterogeneous hardware** focussing on AMD Versal with AI cores, and SiMa.ai Modalix SoCs
- Interdisciplinarity: Particle Physics, Systems Control, Electrical Engineering
- Research infrastructures: Belle II, LHCb, AMBER, PETRA III/IV, XFEL, FRM II, and ESS

DIAMOND

Deep Inverse Software Architectures for Multi-Layer Optical Coating Design

Applications of optical multi-layer coatings



Free-electron laser & synchrotron facilities:
Optics for pump-probe lasers, XUV / x-ray beam lines, ...



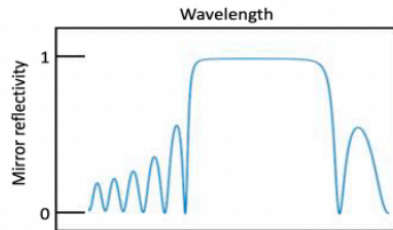
Optics-based manufacturing:
Optics for semiconductor production, additive manufacturing, ...



Broad end-user applications
Optics for electronic displays, eye correction, ...

Inverse design challenge

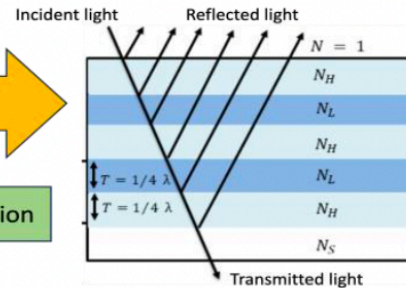
Example optical properties



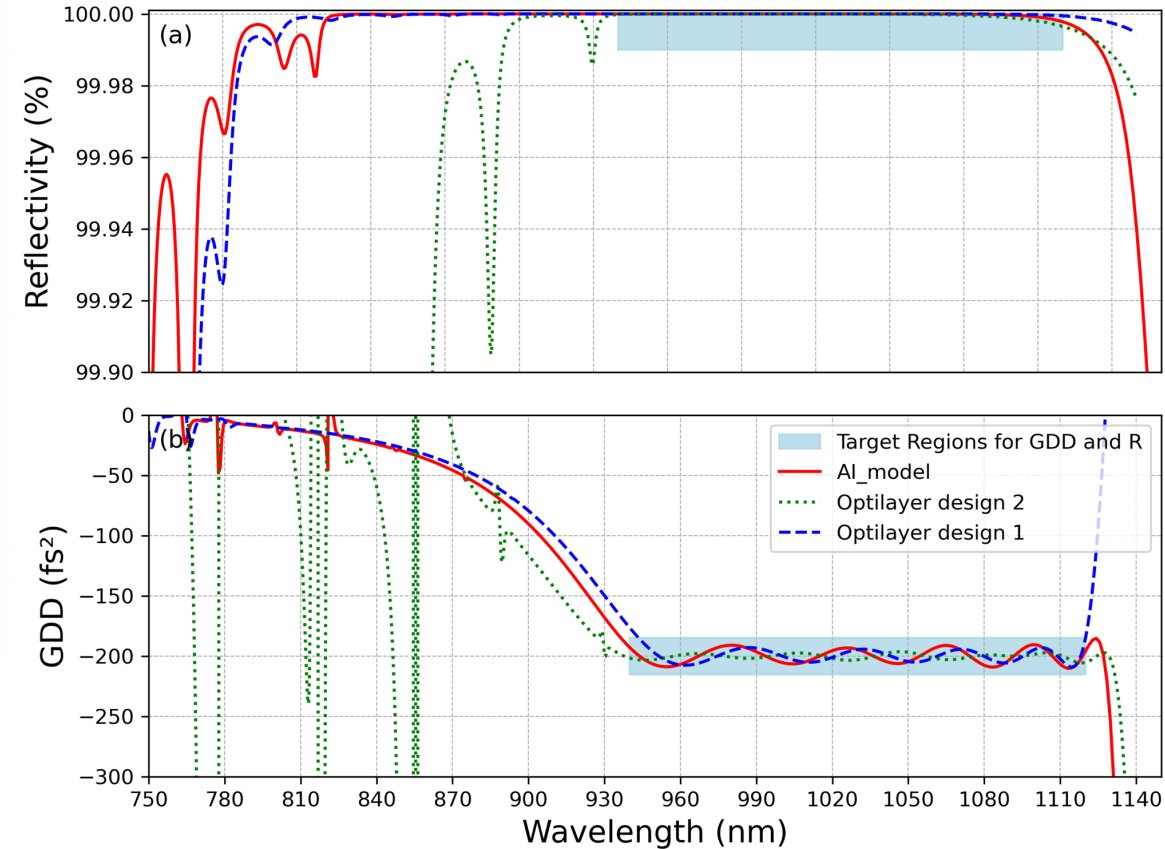
Challenge

Easy computation

Example multi-layer stack



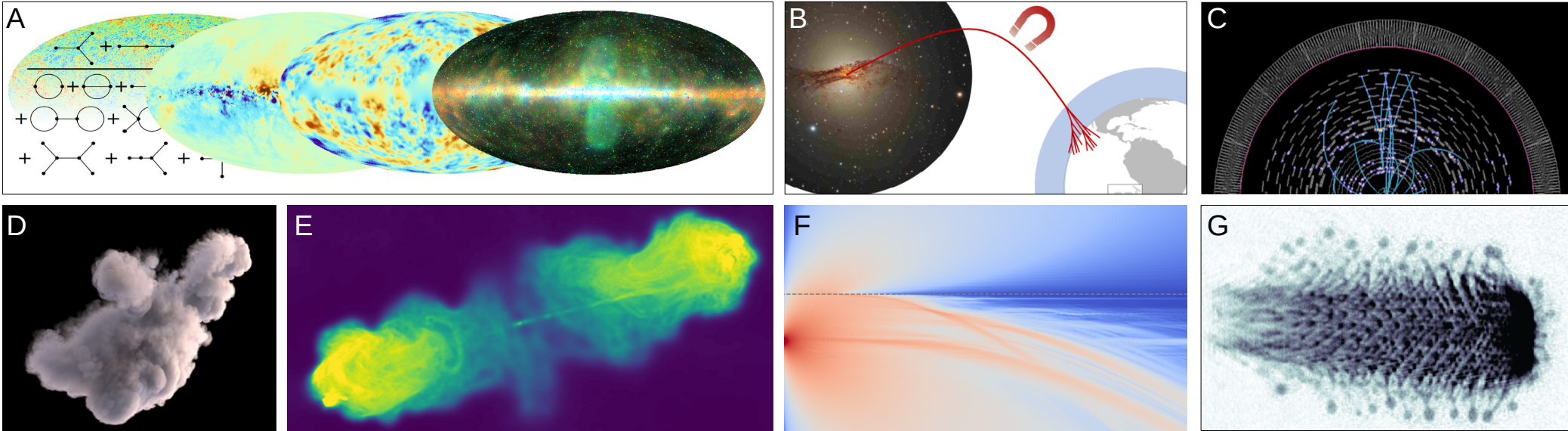
- Escaping local minima to find global optima
- Matching state-of-the-art performance
- Unlocking novel, non-intuitive designs beyond human intuition



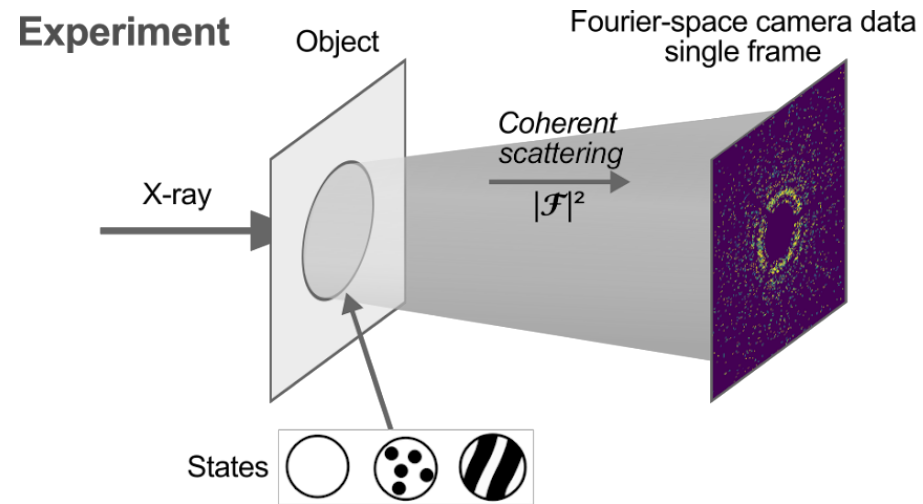
U. Chattopadhyay, et al, "Efficient optical coating design using an autoencoder-based neural network model", . *Phys. Photonics* **8** 015007 (2025) DOI 10.1088/2515-7647/ae1b53

ErUM-IFT-2 Information field theory for experiments on large-scale research facilities

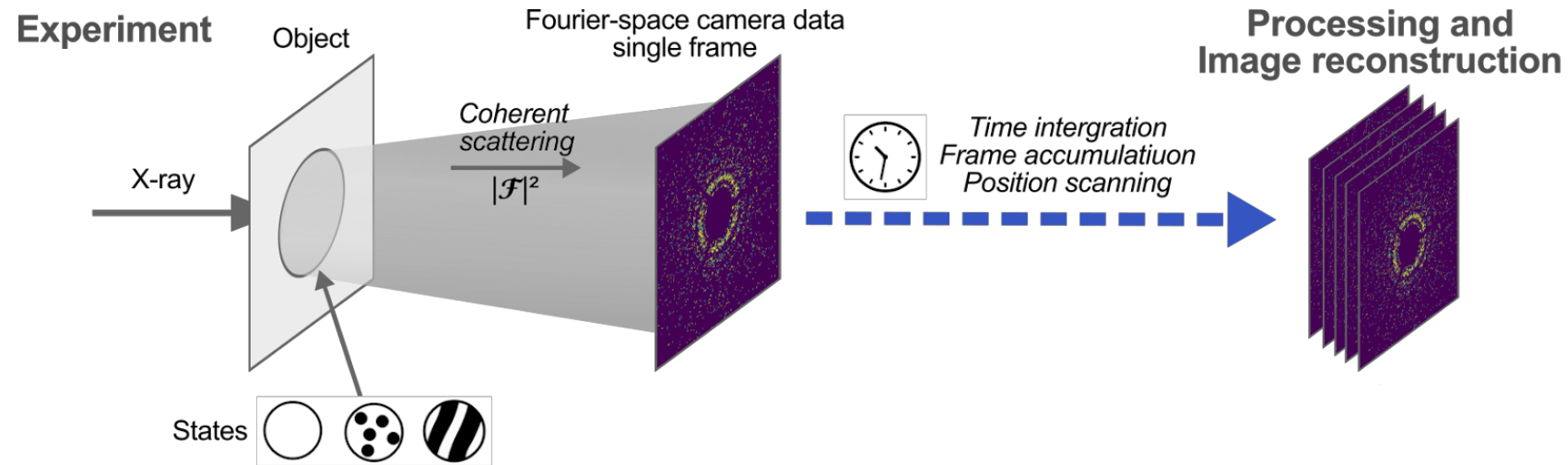
Vincent Eberle, Max Planck Institute for Astrophysics, Garching



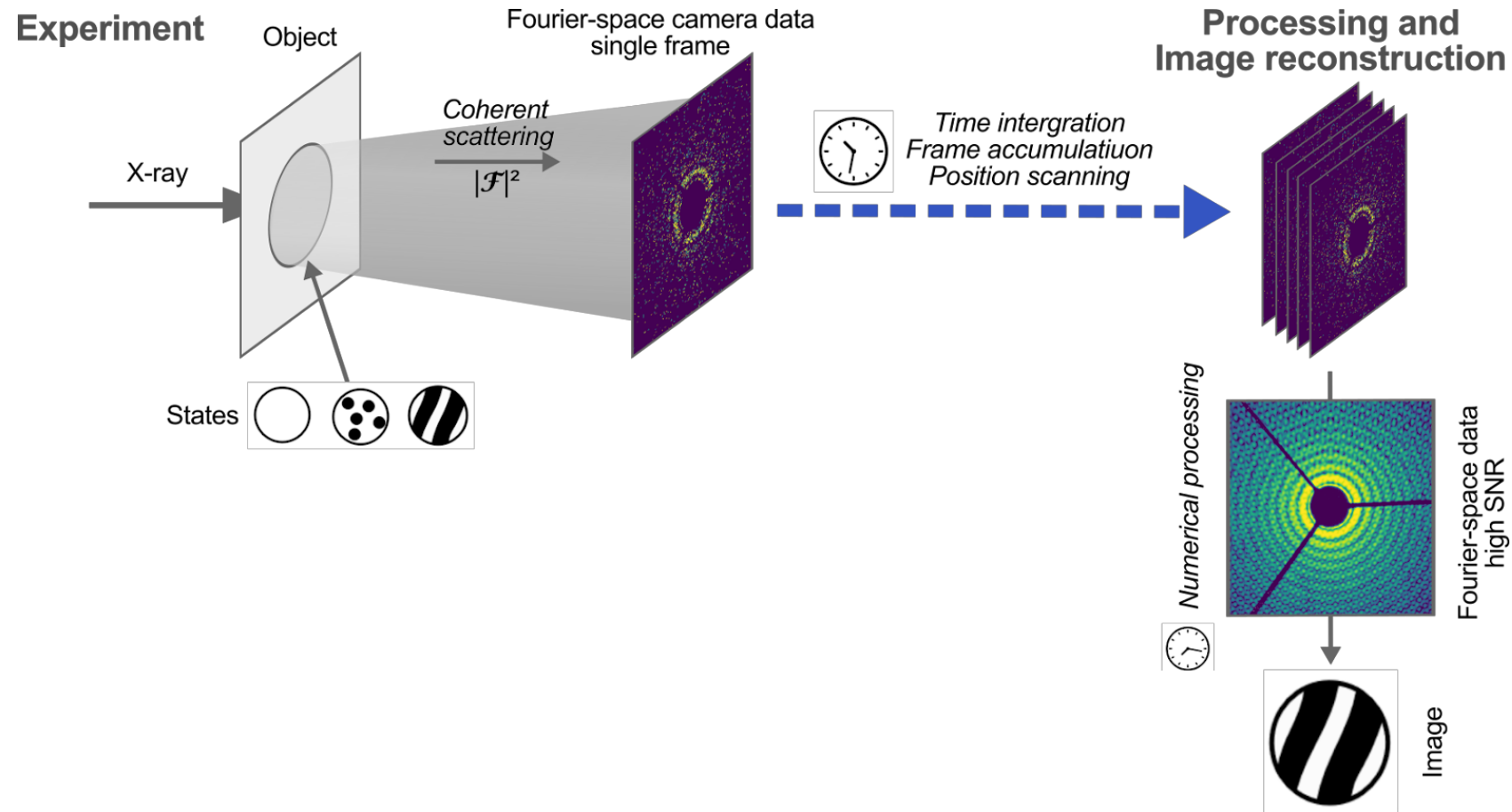
FoMoCID: Foundation model for real-time evaluation of raw coherent imaging data



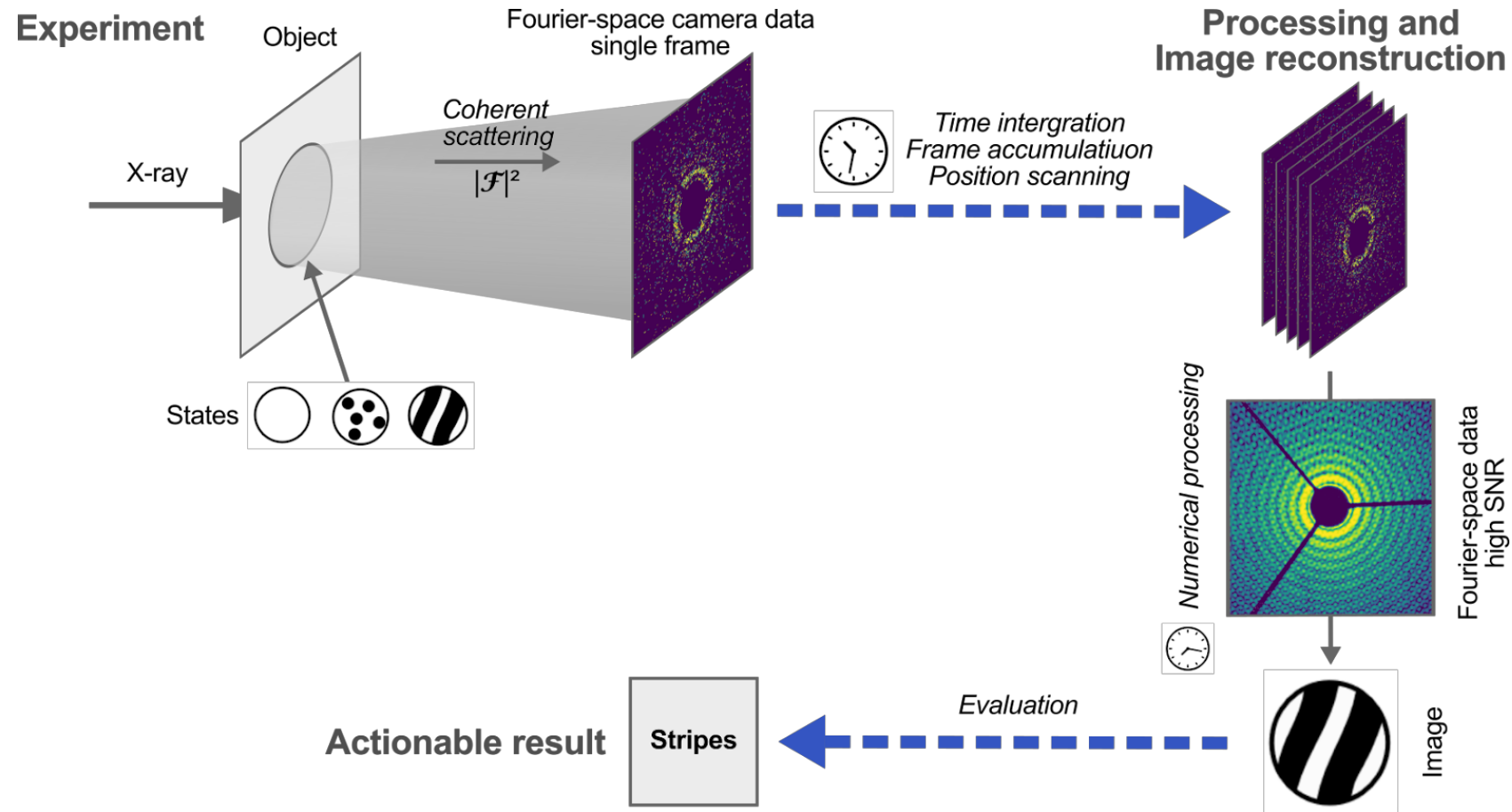
FoMoCID: Foundation model for real-time evaluation of raw coherent imaging data



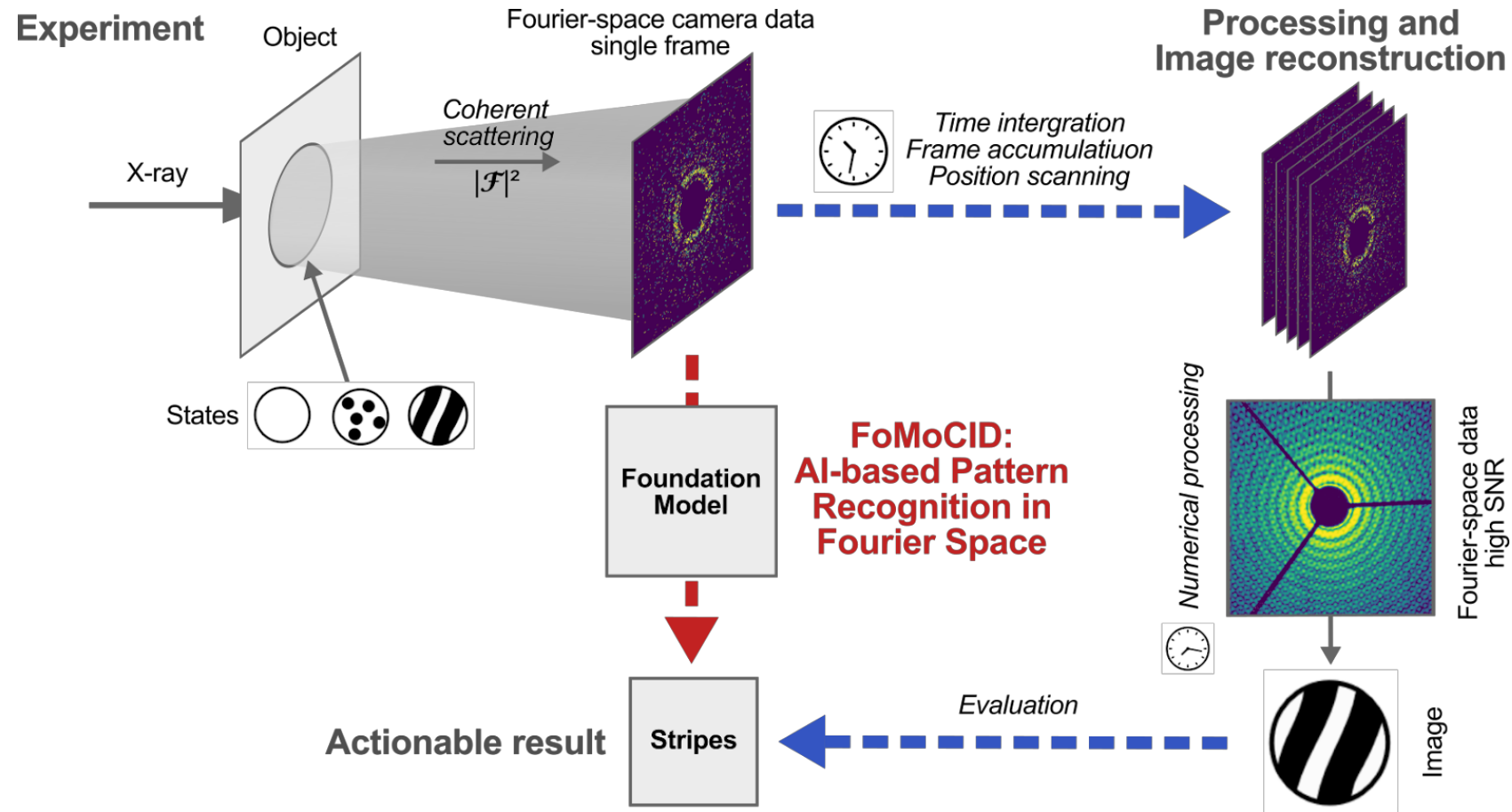
FoMoCID: Foundation model for real-time evaluation of raw coherent imaging data



FoMoCID: Foundation model for real-time evaluation of raw coherent imaging data



FoMoCID: Foundation model for real-time evaluation of raw coherent imaging data



The Görlitz Astronomical Data Centre

Purpose: Establishing a *sustainable* Virtual Observatory node (i.e., astronomical data centre) in Germany.

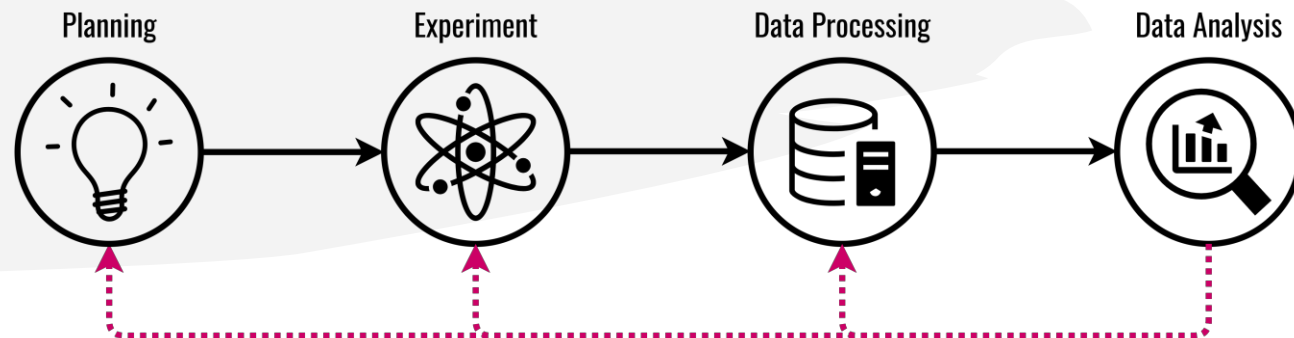
Method: Consolidate initiatives run in Hero Mode (e.g., in Heidelberg, Potsdam) so far and hire sufficient staff to take over

- Science oversight
- Server-side development (e.g., DaCHS)
- Client-side development (e.g., pyVO)
- Data curation
- Stringing (VO education, data acquisition. . .)
- Operations support

The DZA has signalled they would probably take over staff hired for the project.

OPTIX

Optimized Processing and Training for Intelligent Experimentation



Adapted from: Dingel, K. (2022). In *Organic Computing – Doctoral Dissertation Colloquium 2021*, pp. 86–98.



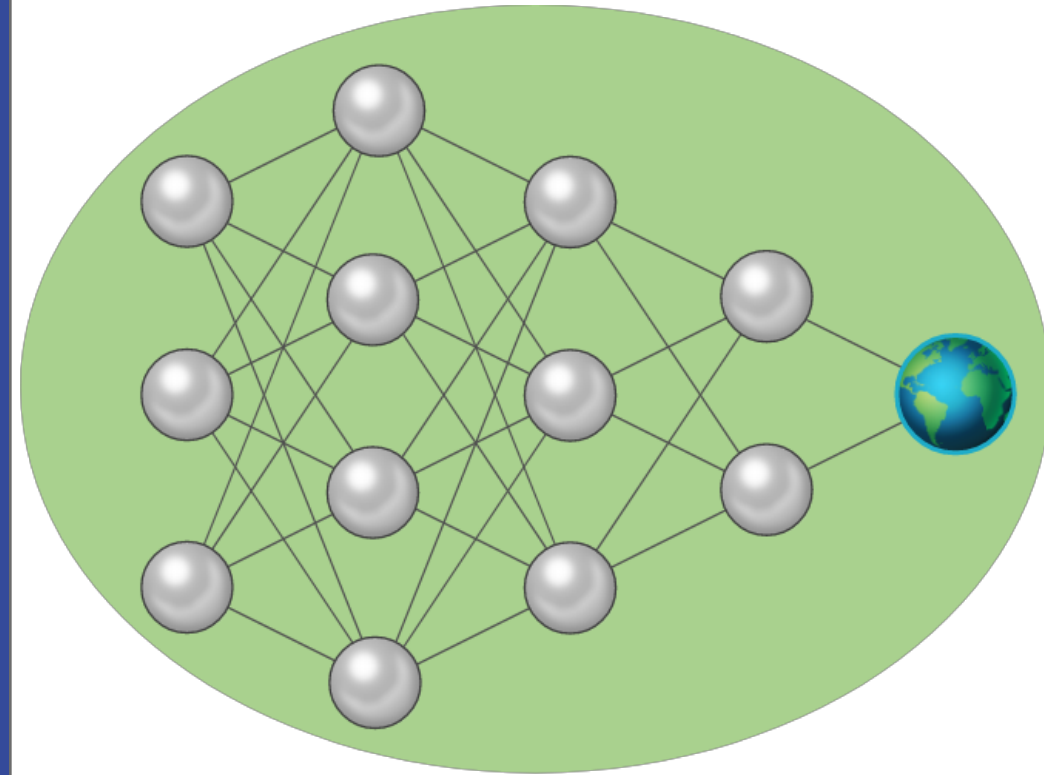
Taken from: Dingel, K. et al. (2022). Artificial intelligence for online characterization of ultrashort X-ray free-electron laser pulses. **Scientific Reports**, 12, 1–14.

universität freiburg

Precision in Energy-Aware AI Research for Low-carbon Solutions

Valerie Lang

On behalf of the PEARLS consortium

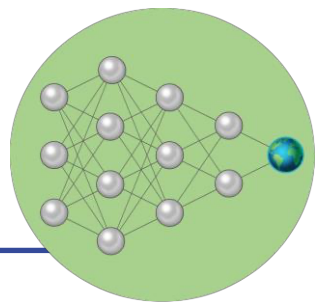


PEARLS

Precision in Energy-aware AI Research
for Low-carbon Solutions

ErUM-Data Onboarding Meeting, 25 Nov 2025

Precision in Energy-aware AI Research for Low-carbon Solutions – PEARLS –



Goal: Develop AI towards improved reliability and reduced carbon footprint

- Consortium of DESY and the University of Freiburg (UFR)
 - Fully led by junior researchers: Dr. P. Niknejadi, Dr. S. Gurdasani, N.K. Bhalla, Dr. V. Lang (spokesperson)

Methods:

- Reduction of the size of AI and machine learning (ML) applications → DESY
 - Pruning → Investigate critical pathways in Neural Networks (NNs) using information flow mapping
 - Data encoding → Investigate tailored image-like data formats for more native ML workflows
- Managing the performance and carbon footprint of AI and ML applications → UFR
 - Carbon accounting → Investigate electricity use during training and inference
 - Uncertainty estimation → Test Bayesian inference methods + Retrieval-Augmented Generation models
 - Improve your footprint → Expand on Know-your-footprint calculator, implementing PEARLS' results

→ If interested in collaborating, please get in contact: valerie.lang@physik.uni-freiburg.de

Reusable Data: the PHOTONIC project

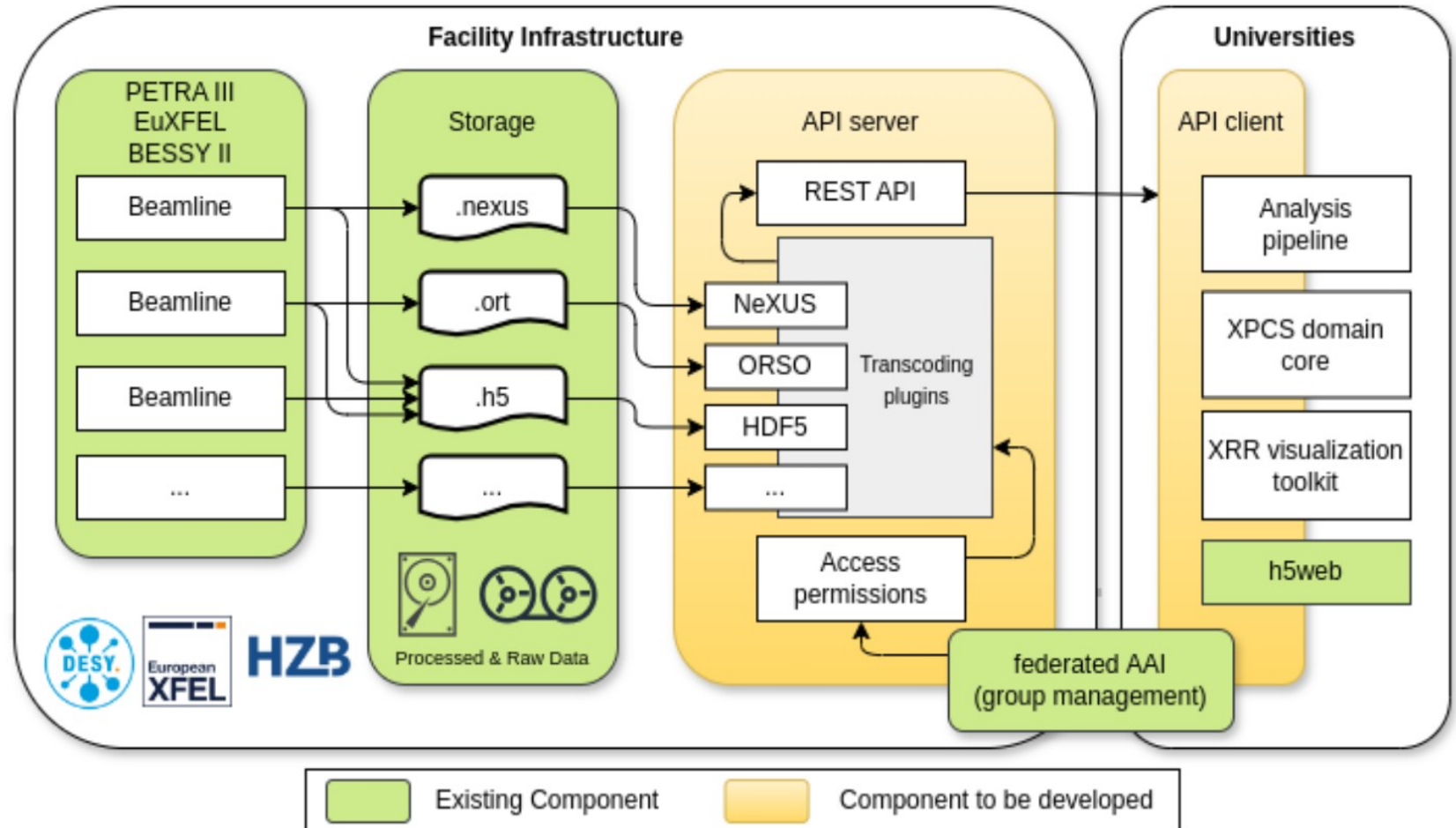
Real-time, online and standardized access to Photon Science data

Key facts:

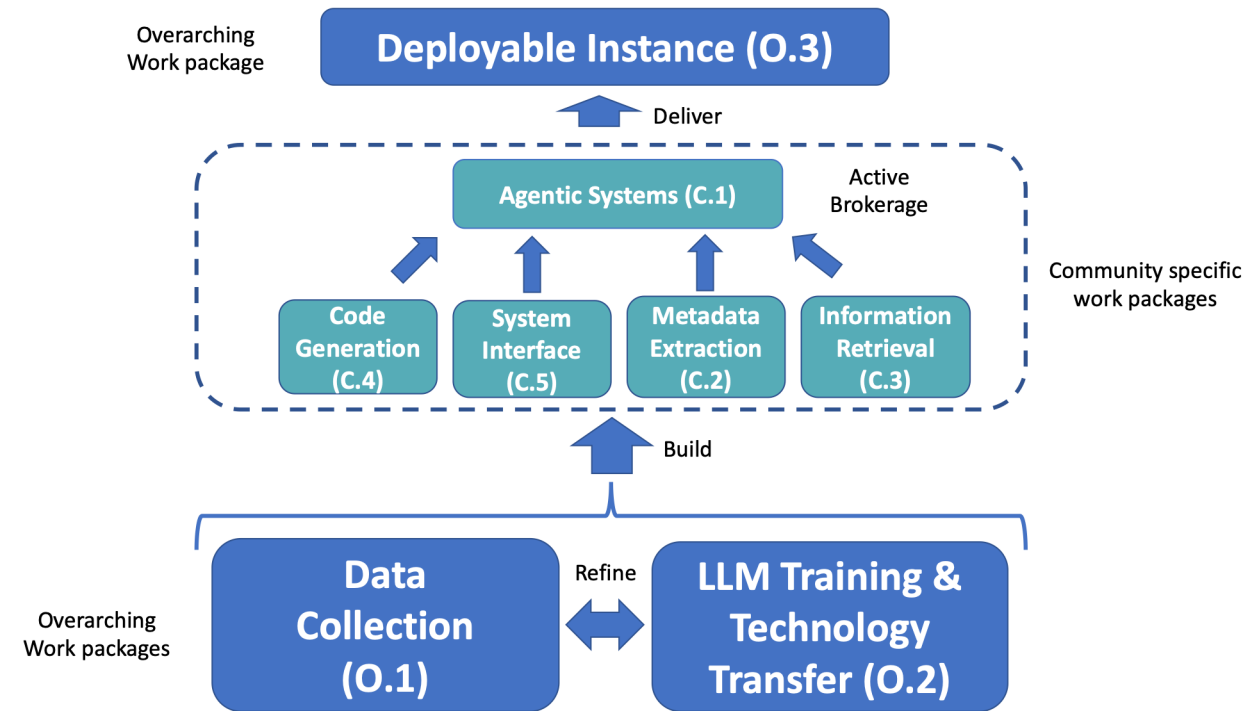
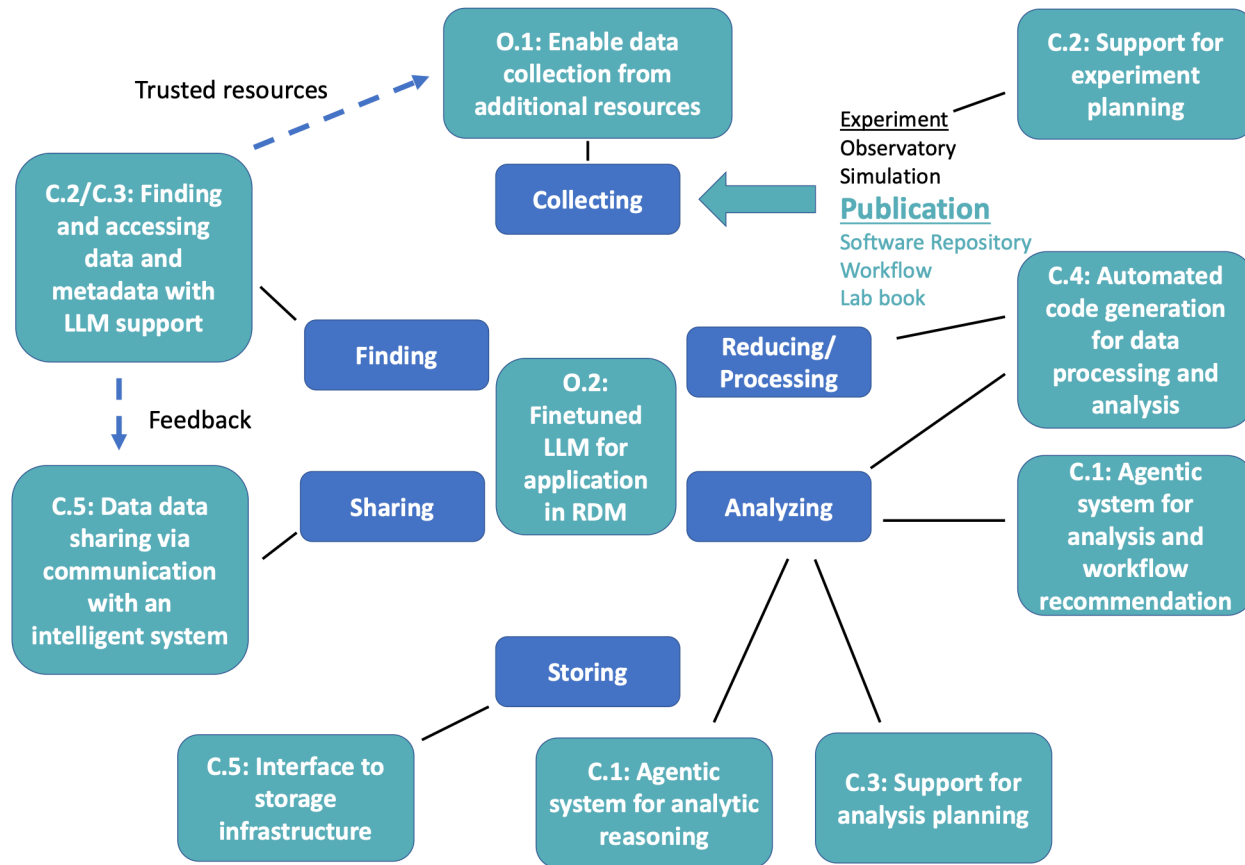
- 3 years project
- start date: 01.11.2025
- Partners: Siegen University, CAU, DESY + HZB and EuXFEL

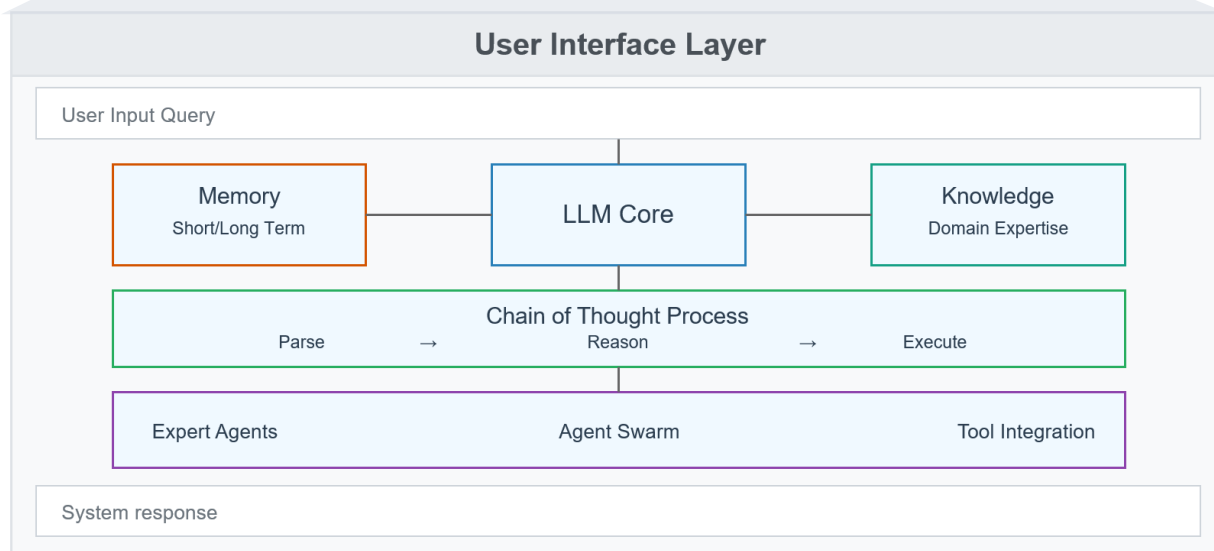
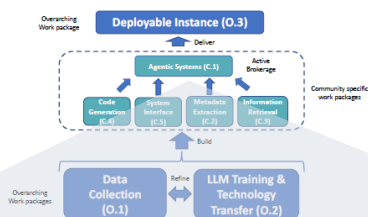
Objectives:

- Unify different data formats from various beamlines and techniques
- Enable live data access during experiments e.g. on a personal laptop connected to eduroam
- Visualise and access only the data of interest



From Big Data to Smart Data with *Physics-LLM*



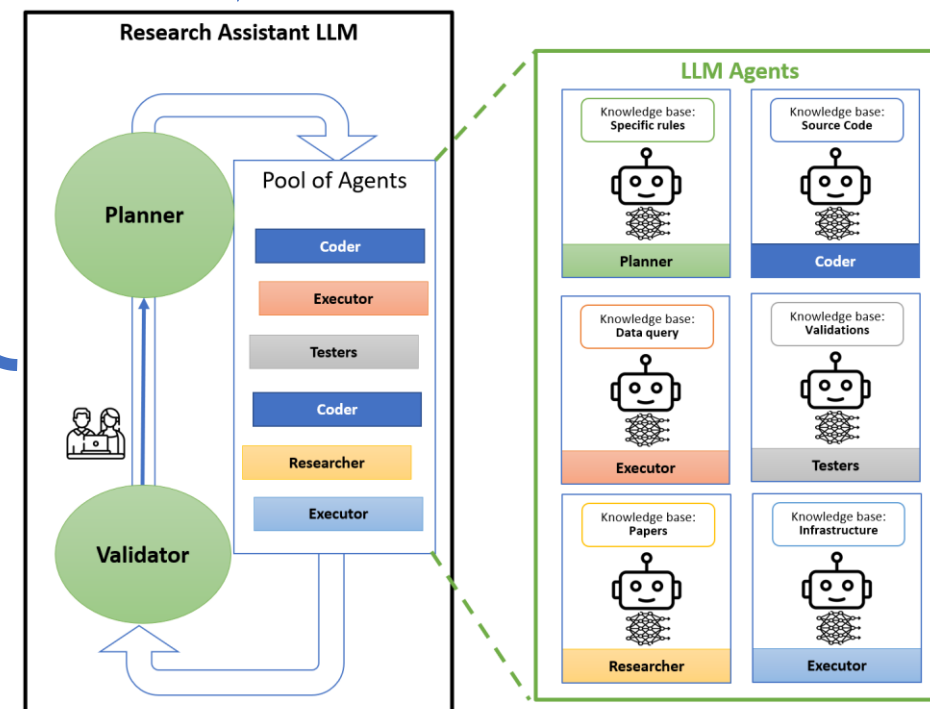


Developing an agentic system with specific domain knowledge

- Agentic system prototype pipeline ready for Astro-community
- Adding new workflow knowledge bases from other domains
- Deployment in production for all ErUM communities



User input:
Show an example of REANA pipeline for Starhorse datasets from 2024, CMD data is extracted and plotted interactively



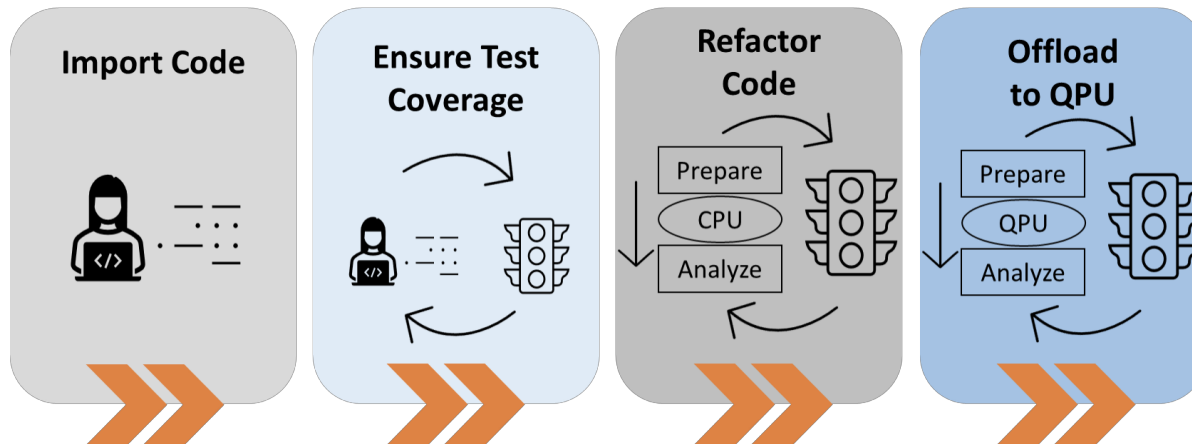
QC4ErUM - Quantum Computing for the ErUM Community.

Novel methods and tools for your research → get ready today !

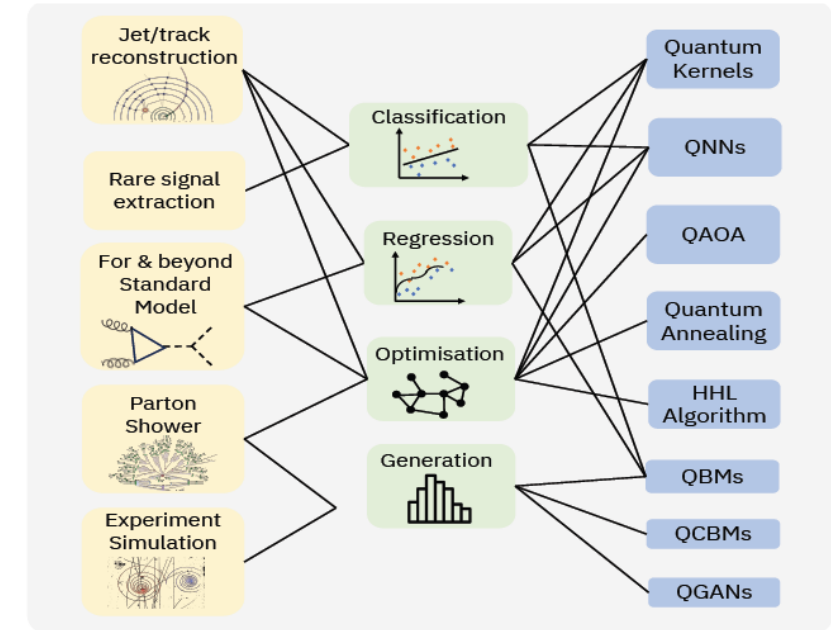
- Scalable and innovative research software is a high priority for the ErUM
 - Quantum Computing hardware and software show a rapid evolution, have intrinsic high potential for challenging problems in optimization, analyzing and simulating complex and highly correlated data.
- integration into research software is mandatory for ambitious research.

QC4ErUM: 2026-2028

systematically supports ErUM in **transforming existing research software into hybrid quantum-classical software processing**: partially on classical hardware (CPUs) and partially on quantum hardware (QPUs).



In addition: **training courses** and a **QC networking forum** about experiences, use cases, techniques and latest developments.

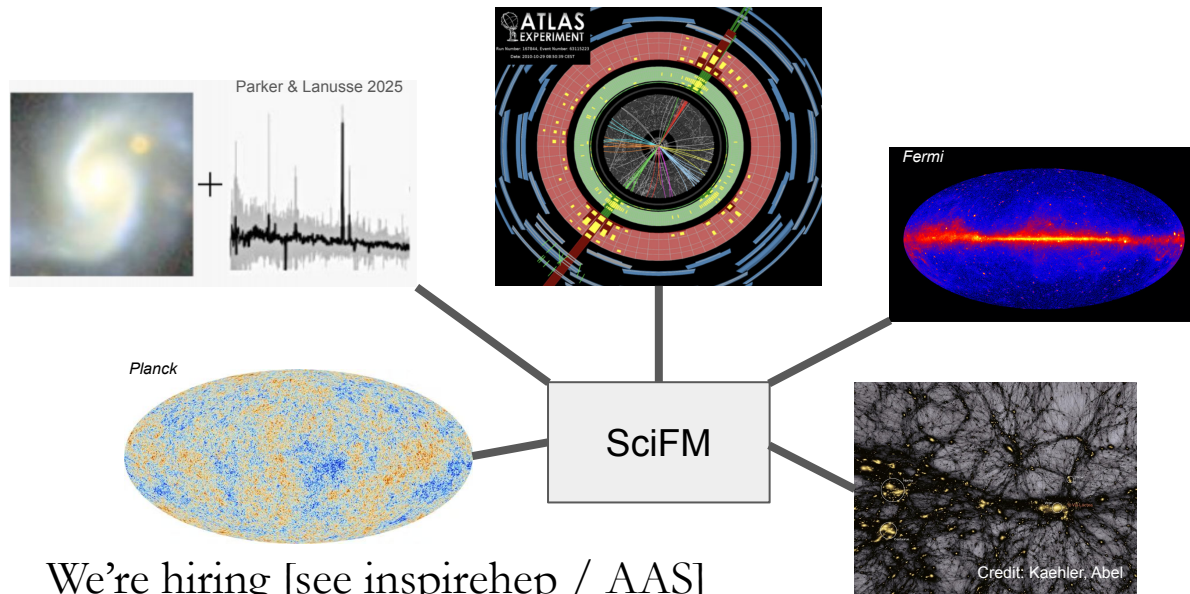


Quantum Computing for High-Energy Physics: State of the Art and Challenges
Published in PRX Quantum:
<https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.5.037001>

Foundation models for Astrophysics and Particle Physics

Self-supervised compression is all you need!

$$\frac{\text{Sensory information rate}}{\text{Behavioral throughput}} \approx \frac{1 \text{ Gbit/s}}{10 \text{ bit/s}} = 10^8$$



Lucie-Smith



Kasieczka



CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Hamburg



Grün



Heinrich



München



Heneka



Plehn



Heidelberg



Mertsch



Krämer



Aachen

SUSFECIT



Sustainable Federated Compute Infrastructures

Markus Schumacher
on behalf of but not blessed by the SUSFECIT Research Network
(Aachen University, Bonn University, CERN, DESY, Freiburg University,
Göttingen University, KIT, Siegen University, Wuppertal University, Öko-Institut e.V.)

With funding from the:



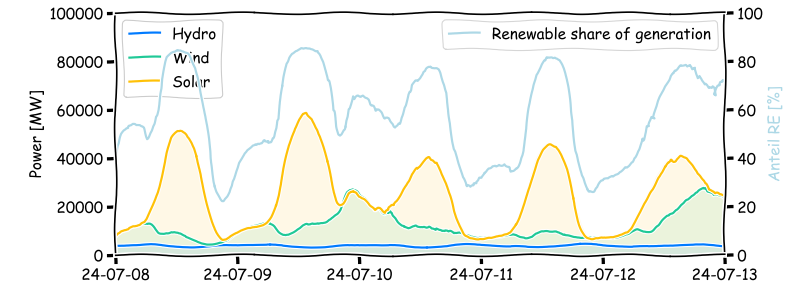
Federal Ministry
of Research, Technology
and Space

Sloppy: „Perform calculations preferably (long term exclusively) when and where sun is shining or wind is blowing“

➤ **A bit more precise:**

Adapt amount of calculations & energy consumption to the temporal and local availability of renewable energies (RE)

e.g. exploit that compute jobs are “easily“ dispatchable in time and space



➤ **Means of adaption:**

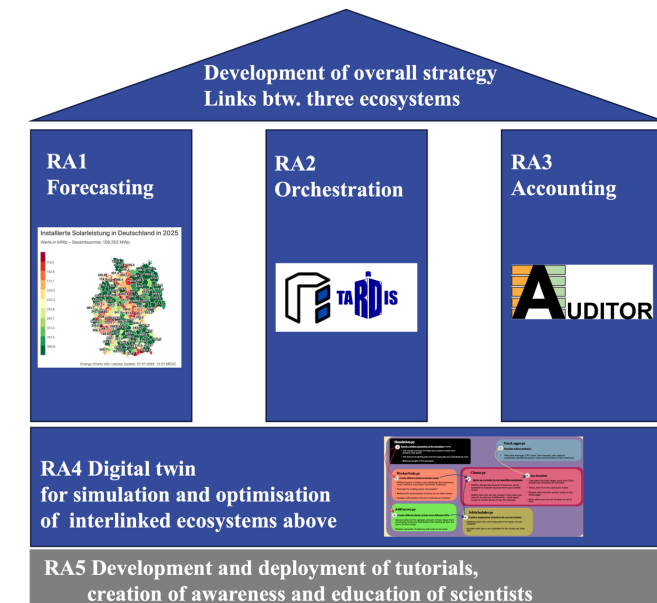
- Shut down/ ramp up (parts) of cluster
- Reduce clock frequency of CPUs (GPUs)
- Transfer jobs to sites where enough RE is available

➤ **Requirements on hardware to compensate for volatility of RE and dark doldrums**

- Overprovisioning of resources
- Inclusion of local energy storage in concept

➤ **Requirements on strategy of operation/ software ecosystems:**

- Forecasting of available RE with good temporal and local resolution
- Orchestration of jobs in time and on compound of sites according to available RE and compute demands
- Accounting of elapsed compute time, consumed energy, produced CO₂ footprint (incl. LCA), used water, ...
- Digital twin to model and optimize overall system



Partners



Achim Stahl



Martin Gasthuber



Philip Bechtle



UNIVERSITÄT
GÖTTINGEN

Arnulf Quadt



Karlsruher Institut für Technologie

Manuel Giffels



Jens Gröger



Markus Schumacher (spokesperson)

Associated Partners



Simone
Campana
universität freiburg



Marc Schumann
Dirk von Suchodoletz



Erik Bründermann
Achim Streit
Kathrin Valerius



Christian Gutt



BERGISCHE
UNIVERSITÄT
WUPPERTAL

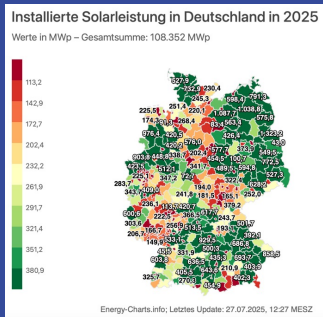
Christian Zeitnitz

Back Up / Additional Information

The „House“ of SUSFECIT, its Research Areas (RAs) and Goals

Development of overall strategy
Links btw. three ecosystems

RA1 Forecasting



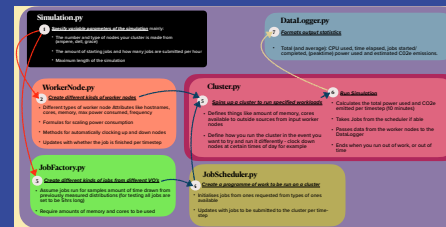
RA2 Orchestration



RA3 Accounting



RA4 Digital twin for simulation and optimisation of interlinked ecosystems above



RA5 Development and deployment of tutorials, creation of awareness and education of scientists

Goals:

- Increase the environmentally sustainable use and operation of federated compute resources.
- Raise awareness of the impact of scientific computing on the CO₂ footprint
- Train talented early career scientists in cutting-edge technologies

Dissemination:

- First to be developed and deployed for ATLAS, CMS and ET
- Later:
 - KATRIN, XENONnT, LXZD (neutrino and astro-particle physics)
 - Experiments at PETRA-III, EU-XFL, ESRF (synchrotron radiation physics)
- Perhaps: Further use cases identified

Links to [PUNCH4NFDI](#), [EGI](#), [WLCG](#) via partners and associated partners
[Associated partners](#) will act as critical reviewers, alpha-users and multipliers