WestAI

Service Center for Artificial Intelligence

Contact

- **Web**: westai.de (Joachim Köhler)
- **Mail**: contact@westai.de
- **LinkedIn**: @westai

WEST AI
KI-Servicezentrum

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

UNIVERSITÄT BONN

Fraunhofer

JÜLICH Forschungszentrum

RWTH AACHEN UNIVERSITY

technische universität dortmund

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

4 AI Service Centers in Germany

- Main focus: Transfer of AI technolgoies to users, research organisations, and industry
- Funded by the BMBF
  - hessian AI Service Center
  - KI-Servicezentrum Berlin-Brandenburg
  - KISSKI- Ein hochverfügbares KI-Servicezentrum für sensible und kritische Infrastrukturen
  - WestAI - Service Center for Artificial Intelligence
- Details:

KI Service Zentrum HPI

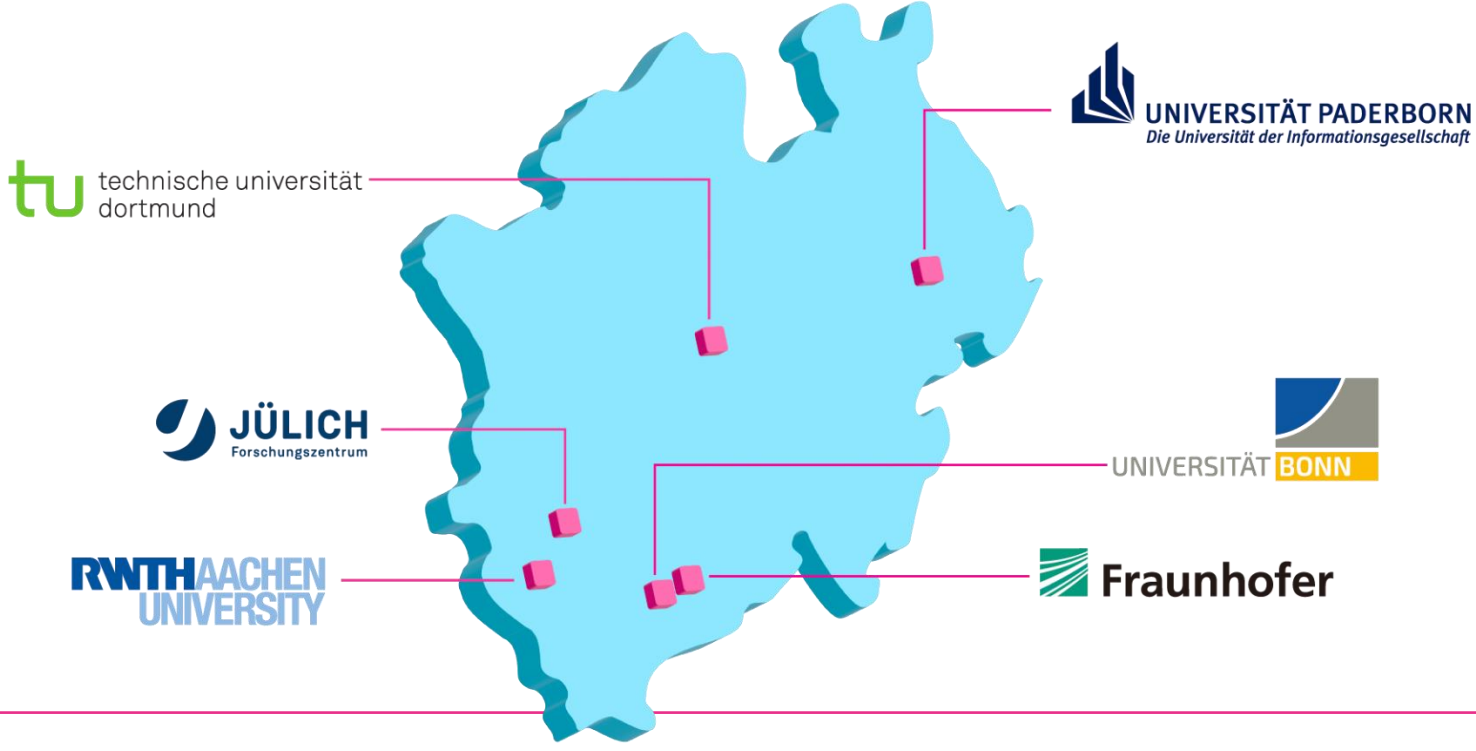
WEST AI KI-Servicezentrum

hessian.AI

KISSKI

WestAI Consortium

AI Competence from North Rhine-Westphalia



Our Mission

- **Large, transferable AI models** can add **trillions of dollar** in value to **global productivity** (McKinsey, 2023)
- Many German companies recognize AI as a relevant topic, but the **majority are not yet using AI applications** (ifo Institute, 2023)

Causes

- **Training large, transferable models** requires **powerful hardware**, which is almost exclusively provided by large commercial institutions
- **Large, transferable models** are often **not open and commercially free** available
- **Multimodal,** transferable models require **further research**
- The **effort required to train and utilize large, transferable models** is **too high** for the majority of institutions and therefore requires **support**

Mission

We provide easy access to **technologies and services** that enable research institutions and companies to **transfer large, multimodal AI models** to their specific use cases

Our Fields of Action

Long-term Establishment of WestAI as an AI Service Center

Hardware

Provisioning powerful, scalable computing resources for AI

Research

Enabling the next generation of large transferable AI models

Services

Driving the dissemination of AI through an easily accessible, broad, high-quality service portfolio

We bring together hardware, research and application expertise.

Hardware

HPC Systems @RWTH + Jülich

CLAIX-2023

- 15 Nodes†
- each 4 H100 GPUs*
- Serving 10+ projects since January 2024

JURECA

- 16 Nodes†
- each 4 H100 GPUs*
- Serving 4+ projects since May 2024
- Isolated and exclusive usage for WestAI

- Access via batch mode (SSH) and interactive mode (JupyterHub) possible
- Support for all common AI frameworks (PyTorch, Scikit-Learn, TensorFlow)
- Access to various file systems and community datasets
- Requesting compute time possible over Jards for academic and contact@westai.de for commercial usage
- When approved, access to 10.000 GPUh is provided

← Jards

†Infiniband Connected
*Full NVLINK, 94GB HBM2e

Research

Large-scale, Multimodal and Transferable AI Models and Datasets

Scalable data management & heterogeneous data integration

- Generation of datasets for training large models
- Development and application of methods for processing sensitive data

Pre-training models on large scale

- Training of foundation models transferable to different domains and tasks
- Derivation of scaling laws

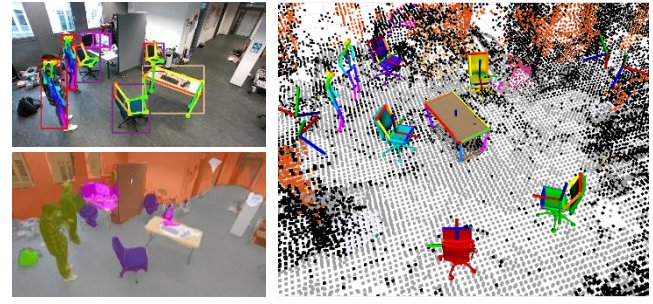
Transfer of pre-trained models to various domains

- Model transfer to various domains and tasks (speech, image, audio, video, ...)
- Investigation of influencing variables on the transfer behaviour

Transfer of models to hardware with low-resource constraints

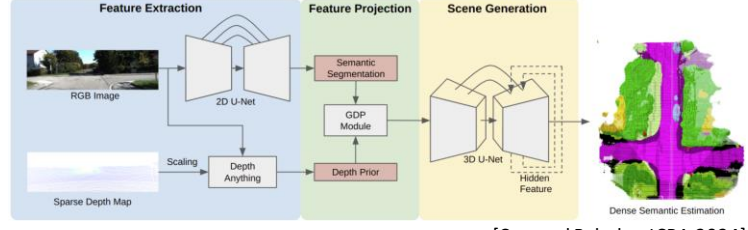
- Model compression
- Transfer of models to resource-constrained use cases

Object-level 3D Semantic Mapping with Smart Edge Sensors



[Hau, Bultmann, and Behnke, IRC 2022]

Sequential LiDAR-Camera Fusion for Semantic Scene Completion

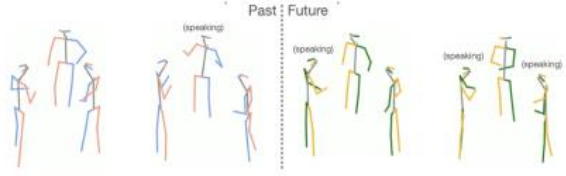


[Cao and Behnke, ICRA 2024]

Research

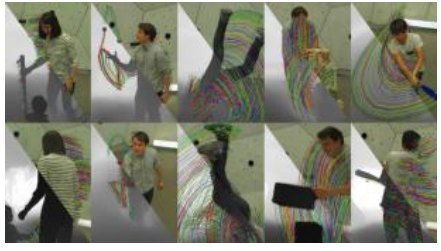
Large-scale, Multimodal and Transferable AI Models and Datasets

Social Diffusion: Long-term Multiple Human Motion Anticipation



[Tanke et al., ICCV, 2023]

Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis



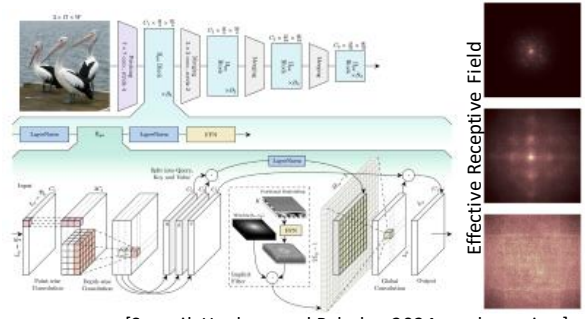
[Luiten et al., 3DV, 2024]

Training framework for multimodal large language models that leverages PyTorch's native capabilities (e.g. FSDP)

Learning subsurface scattering solutions of tightly-packed granular media using optimal transport

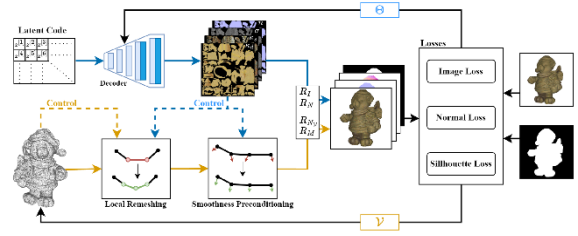
[Zingsheim and Klein, Computers & Graphics, 2024]

HyenaPixel: Global Image Context with Convolutions



[Spravil, Houben, and Behnke, 2024, under review]

ROSA: Reconstructing Object Shape and Appearance Textures by Adaptive Detail Transfer



[Kaltheuner et al., 2024, under review]

Services

Support (Large-Scale) AI Uptake and Lower Technology Barriers

Scalable and simple access to computing resources

- Provision of easily accessible computing and storage resources
- Best practices for using scalable computing resources

AI Resource Hub and Experiment Space

- Development and open provision of software, pre-trained models and datasets
- Ensuring data sovereignty

Development service & AI consulting

- AI exploration consultation
- Support in realization consultation free of charge for industry and academia using research collaborations

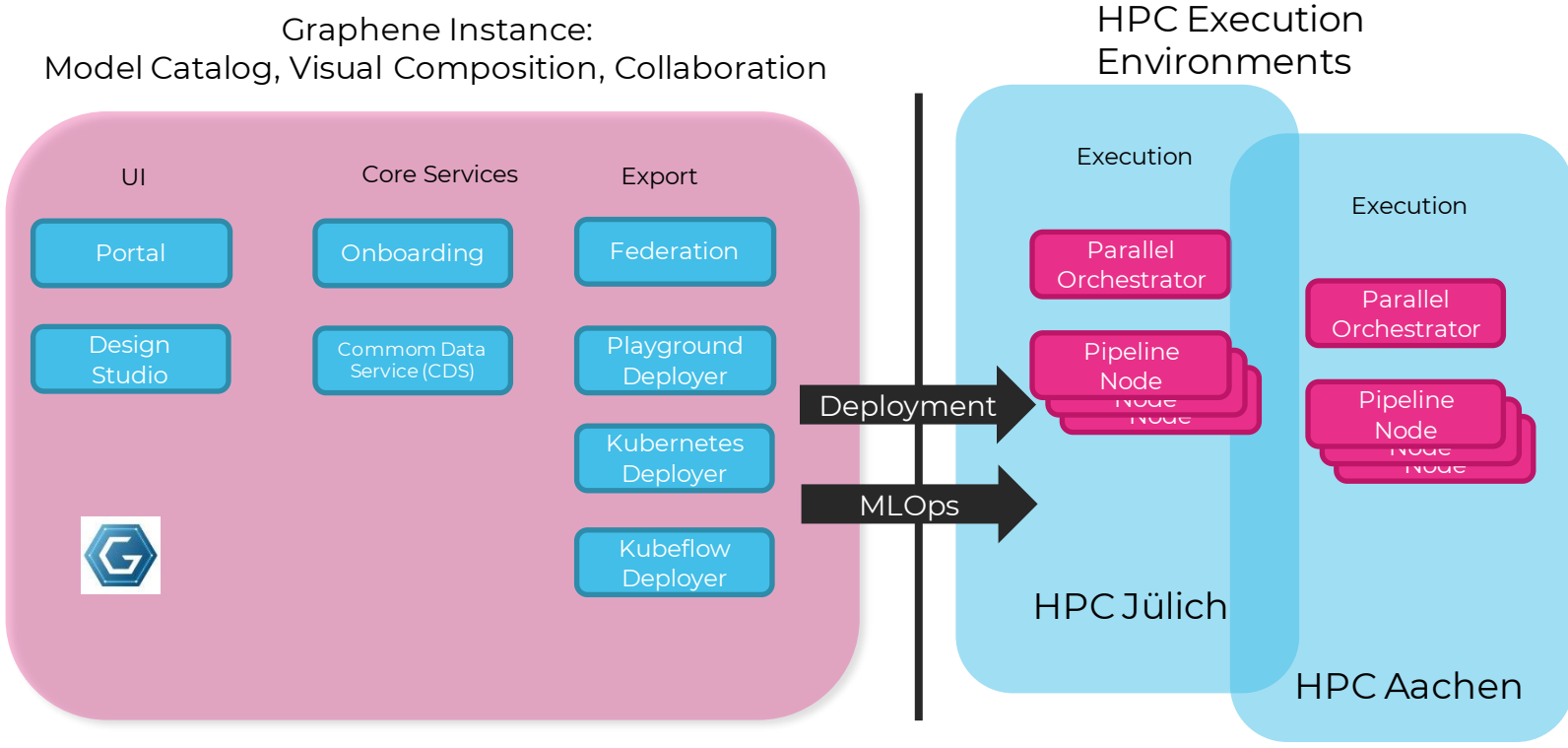
Courses, training, and education

- Organisation of training courses

Communication and visibility

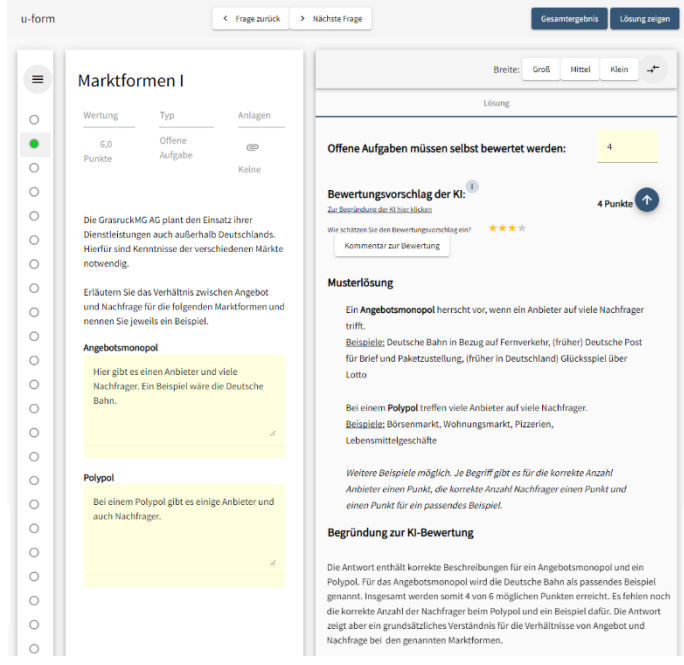
- Scientific communication
- Organization of events
- Community management
- Collaboration with partner initiatives
- Access to expertise of the strong open-source grassroot ML research community via LAION

Low effort Deployment of Models and Pipelines to HPC



WestAI Service Success Stories

- Research collaboration between Forschungszentrum Jülich and the U-Form-Verlag
- Several requests from external parnters https://westai.de/ using an internal ticketing system (> 30 requests during last months)
- Huge amount of courses and trainings: https://westai.de/schulungen/
- Easy access to compute by Aachen and Jülich (H100 nodes)
- The project integrates **various language models**, comparing commercial ones with open-source models, Hugging Face space: https://huggingface.co/WestAI-SC (first models released)



Outlook and Networking with National/European Initiatives

Outlook

- Conference of the German AI-Servicecenters (KonKIS 24, 18 / 19 Sept. 2024) →
- Research base model pre-training efforts and research model deployment
- Execution of further **realization projects** with commercial partners
- Expansion of the **service portfolio** (https://westai.de/services/)

Networking with National/European Initiatives

- Collaboration with regional AI activties (KI.NRW, Lamarr Institute: https://lamarr-institute.org/)
- Strong network with other Germain AI Service centers
- Synergies between DeployAI (European-AI-On-Demand Platform), EDIHs, upcoming AI Factories

Get in Touch with us!

- **Web**: westai.de
- **Mail**: contact@westai.de
- **LinkedIn**: @westai