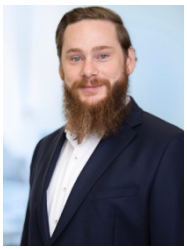
Dr. Gabriel Dax



83410 Laufen | gdax.science@gmail.com | linkedin.com/in/gadax

SKILLS

Languages: Python, C/C++ (C++20, OpenMP, pybind11, Boost), CUDA, Bash
Frameworks: PyTorch (DDP, NCCL, AMP), TensorFlow, Nvidia DALI, TensorRT, ModelOpt, VitisAI, OpenCV, MLflow
Infrastructure: Slurm, Docker, Apptainer, Distributed Training (up to 32 GPUs / 8 nodes), HDF5, WebDataset, Git
Hardware: HPC-Cluster, CPU/GPU Server, FPGA (Xilinx ZCU102 UltraScale+)
Spoken Languages: German (native), English (fluent)

PROFESSIONAL EXPERIENCE

Jul. 23 – Now — Machine Learning Engineer
Fraunhofer Institute for Integrated Circuits IIS – Munich, Germany

- Built a **data-parallel deep learning infrastructure** on a Slurm-based HPC cluster, scaling training up to 32 GPUs across 8 nodes via NCCL with mixed-precision training and containerized deployment (Apptainer).
- Achieved **5x runtime improvement** on a C++/OpenMP image analysis tool by redesigning the computation with parallelization directives for batch mean-value computation.
- Developed custom **CUDA kernels** for GPU-based image processing, tiled inference, and IoU computation.
- Built an AutoML-driven defect detection system for semiconductor manufacturing with hyperparameter optimization and automated model selection. Pruning reduced the inference time by 35% (same accuracy).
- Eliminated I/O bottlenecks in distributed training pipelines by implementing NVIDIA DALI-based data loading, reducing data ingestion time at scale.
- Designed a **SAM-based object segmentation pipeline** for industrial waste detection, covering data processing, different augmentation strategies, model training, and evaluation.
- Fine-Tuned and evaluated **vision-language models (CLIP) for zero-shot** classification of remote sensing imagery, benchmarking against supervised baselines to assess deployment feasibility.
- Developed a Python Package that **communicates with a REST interface** to send and receive data (AAS, EDC). The package is used in a demonstrator that is presented at Hannover Messe 2026.
- Built a GPU utilization monitoring tool for Slurm-managed clusters, enabling asynchronous tracking of compute resource usage across multi-node training jobs.

Oct. 20 – Jun. 23 — Research Associate
Technical University of Munich – Munich, Germany

- Designed a framework for parametric analysis of **image compression** configurations (across 3 CNNs and 2 Satellite image datasets), reducing average training time by 19.9% with <1% accuracy degradation.
- **INT8 Quantized deep learning models** and deployed them to FPGA (Zynq UltraScale+) via VitisAI, benchmarking accuracy across 11 CNN architectures and profiling power consumption and throughput.
- **Reduced CNN inference pre-processing time** by 25.2% on FPGA through input compression, demonstrating algorithm optimization for latency-critical edge hardware targeting satellite onboard processing.
- Developed the **image processing and classification component** of a wildfire detection framework (TensorFlow) for satellite onboard processing, to reduce downlink data volume.
- Implemented a **compression-based trajectory similarity metric** (Python/C++) using Bloom filters, achieving 8.9% accuracy improvement over baselines while reducing computational complexity.

Dec. 19 – Sep. 20 — Research Associate
Bundeswehr University Munich – Munich, Germany
Supervisor: Prof. Dr. Martin Werner

- Developed a **Python/C++ package** (pybind11, Boost, >1,000 lines C++ core) for in-memory processing of raster data, similarity distance computation using entropy, and binary operations.
- Designed a software architecture for **large-scale satellite data** collection and information extraction from remote sensing imagery.

Feb. 19 – Jul. 19 — Research Assistant
German Aerospace Center (DLR) – Oberpfaffenhofen, Germany
Supervisor: Prof. Dr. Mihai Datcu

- Designed and implemented a **change detection framework in Python and MATLAB** for satellite imagery analysis, processing 70,000 km² of remote sensing data using compression-based data mining.
- Achieved **36% runtime reduction** through computational overhead elimination and memoization.

EDUCATION

Oct. 20 – May 24 — Doctor of Natural Sciences in Aerospace and Geodesy
Technical University of Munich – Munich, Germany
Thesis: Aspects of Algorithmic Information Theory in Spatial Machine Learning
Supervisor: Prof. Dr. Martin Werner
Mentor: Jose Moreira (IBM)

- Contributed to the research community by publishing 18 papers (e.g., ACM CF, MLSys, IGARSS, JSTARS).
- Focus: Optimization of data-driven pipelines using compression to increase performance and scalability.

Sept. 17 – Oct. 19 — Diplom-Ingenieur in Information Technology and Systems Management
Salzburg University of Applied Sciences – Salzburg, Austria
Thesis: Supervised and Unsupervised Data Mining Methods in Remote Sensing

- Contributed to the research community by publishing 5 papers (e.g., Springer, SPIE).
- Relevant Coursework: Software Engineering, Software Architectures, Parallel Computing

Sep. 14 – Jul. 17 — Bachelor of Science in Information Technology and Systems Management
Salzburg University of Applied Sciences – Salzburg, Austria
Thesis: Performance Data Collection in a Distributed System for Rendering Cinema Movies

- Relevant Coursework: Software Development and Operating Systems, Algorithms and Data Structures

COURSES & CERTIFICATES

Node-Level Performance Engineering – LRZ, 2025 (Architectures, NUMA, SIMD)
Fundamentals of Accelerated Computing with Modern CUDA C/C++ – NVIDIA, 2023 & 2025 (CUDA C/C++, Thrust)
GPU Programming Workshop – LRZ, 2025 (OpenACC, NVIDIA Nsight Profiling)
Parallel Programming of High Performance Systems – LRZ, 2023 (OpenMP and MPI)

TEACHING & SUPERVISION

Teaching assistant for C++ programming (5 hrs/week), covering OOP, templates, STL, memory management, multi-threading, MPI, and MATLAB in bachelor's and master's programs at TUM. Supervised multiple master's theses in compression, computer vision, and quantum computing at TUM and Fraunhofer IIS.

SELECTED PUBLICATIONS

Dax, G., Nagarajan, S., Li, H., & Werner, M. (2023).
Compression Supports Spatial Deep Learning.
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)

Koch, A., Petry, M., Ghiglione, M., Raoofy, A., **Dax, G.**, Furano, G., Werner, M., Trinitis, C., & Langer, M. (2023).
Machine Learning Application Benchmark.
20th ACM International Conference on Computing Frontiers

Raoofy, A., **Dax, G.**, Serra, V., Ghiglione, M., Werner, M., & Trinitis, C. (2022).
Benchmarking and feasibility aspects of machine learning in space systems.
19th ACM International Conference on Computing Frontiers

Dax, G., & Werner, M. (2021).
Trajectory Similarity Using Compression.
22nd IEEE International Conference on Mobile Data Management