Dr. James Kahn

Senior AI Consultant

Machine Learning | Big Data | Python | HPC/Grid Computing

Melbourne, VIC
Australian citizen

Email: kahn.jms@gmail.com
Phone: 0413 241 335

[LinkedIn](#) | [GitHub](#) | [Homepage](#)

SUMMARY

- Ex-particle physicist focusing on AI for energy research over the last 3 years
- Big data experience with High Performance Computing (HPC) and expert understanding of Linux systems
- 10+ years of software development, scientific computing, and project leadership
- Highly creative problem solver, "think outside the box" mentality
    - State-of-the-art solutions in multiple fields: 5 delivered research projects (published, open-sourced) in last year
    - Invented novel deep learning processes used by 1000+ member international physics collaboration today

SKILLS

Languages:	Python as primary, Bash, SQL, C/C++
Frameworks:	Pytorch, Tensorflow/Keras, Graphcore
Day-to-day:	Linux, CI/CD, Git, Docker/Singularity, Slurm/HTCondor
AI Fields	Computer Vision, Graph Learning, Automation & Optimization
Technologies:	Numpy, Scikit-learn, Pandas, Matplotlib, Optuna, MLFlow, ArcGIS

EVENTS/TALKS

- **NHR-Computational Physics Symposium,** 11-2022: BaumBauen in Belle-II
- **AI Hub @ Karlsruhe**, 10-2022: AI Consulting in Energy Research
- **Zero-To-Hero Summer Academy**, 09-2022: Tutor for 3-day workshop
- **Spring '22 Helmholtz AI Consultants Hackathon Retreat**, 05-2022: Host of 3-day hackathon
- **Introduction to Machine Learning @ KIT**, 01-2022: Tutor for 3-day workshop
- Annual guest lecture for **Modern Methods of Data Analysis** Master's course at Karlsruhe Institute of Technology with 30+ attendees
- **NVIDIA AI Podcast**, 05-2021: Deep Learning's Collision Course with Particle Physics (Ep. 143)
- **ZDF Volle Kanne**, 05-2021: Project feature on German national TV program

WORK EXPERIENCE

Helmholtz AI - Senior AI consultant, Germany   **OCTOBER 2020 - DECEMBER 2022**
> AI consulting unit for Germany's largest scientific organisation, servicing 40k+ employees

Responsibilities: Providing project management and technical leadership consulting to Helmholtz researchers nationwide in the field of energy. I have supervised, built, and delivered projects that represent more than 30% of the unit's total research output.

Technology Landscape: 45+ AI consultants, 6 teams, and 2 supercomputers (HoreKA, JUWELS Booster) utilising Red Hat Enterprise Linux (RHEL), GPUs/IPUs, Slurm, Docker/Singularity, MPI, and more.

Key contributions:
- **Reported directly to the Unit Head**, partnered with energy researchers nationwide to **manage and deliver** 17 consulting projects across two years.
- **Acquired 100k€ funding** as Principal Investigator (P.I.) **leading** 5 person research team developing AI-based smart grid vulnerability framework.
- **Invented 3 new AI technologies from scratch** to solve data simulation/analysis bottlenecks and **implemented and open-sourced** all in PyTorch.
- **Led by hands-on example**: organised annual consultant team-building retreats, led task-oriented hackathon for fellow consultants to learn new ML technologies**.**
- **Coordinated and managed** Concentrated Solar Plant (CSP) maintenance schedule optimization project utilising Boosted Decision Trees, resulting in a 15% reduction in wasted cleaning operations.
- **Mentored 6 doctoral students** to grow their technical knowledge and independence as researchers, and **conceived and supervised** multiple master's and bachelor research projects.
- **Produced** field leading AI research, resulting in 9 publications and 5 code bases

Steinbuch Centre for Computing - Postdoctoral fellow, Germany   **MAY 2019 - SEPTEMBER 2020**
> A national centre for data-intensive computing servicing researchers across southern Germany

Responsibilities: 50% position working as liaison and technical operations support for GridKa Tier-1 computing site. Responsible for ensuring reliable delivery of grid computing resources and expanding available resources, as well as representing site at triannual steering board meetings.

Technology Landscape: ~300 computer scientists and 4 national data/HPC facilities using Red Hat Enterprise Linux (RHEL), Slurm, Git, Docker, Puppet, Foreman, Redmine, Grafana, and more.

Key contributions:
- **Continuously administered** compute resource availability totalling 30k+ CPUs, 24PB Disk, and 46PB Tape as part of scientific data management team, using **industry-standard best practices**.
- **Expanded available opportunistic resources** to incorporate 1000+ extra available job slots (~28k+ cores) into HTCondor cluster from 4 sites across Germany.
- **Coordinated and completed international transfer** of 2PB raw experimental data, and **established local hosting** access of transferred data for researchers around the globe.

- **Acquired and Integrated** 8 NVIDIA V100 GPUs into local high throughput cluster, including **implementing** necessary HTCondor modifications to manage GPU resources. This served as the framework for installation of an additional 32 V100 GPUs.
- **Maintained** access points for all participating organisational units of grid centre, responsible for interactive access for members of all Large Hadron Collider experiments.

Belle II collaboration - Postdoctoral fellow, Germany   **MAY 2019 - SEPTEMBER 2020**
> An international particle physics experiment studying the fundamental laws of nature.

Responsibilities: 50% position as researcher, head of machine learning group, and supervisor for local Belle II particle physics team. Responsible for the conception of research projects and mentoring of the next generation of physics researchers.

Technology Landscape: 1000+ physicists utilising a central analysis software (Python API, C++ backend), international grid-computing resources, and Atlassian (Bamboo CI/CD, JIRA, Bitbucket, Confluence).

Key contributions:
- **Arranged and supervised** 4 Master's research project focused on deep learning, covering graph neural networks, uncertainty quantification, and fast simulation. All students went on to undertake PhD research projects on related topics.
- **Identified unnecessary caching** in collaboration's Decision Tree-based reconstruction software training procedure and **implemented an automated cleanup** resulting in a memory reduction of over 50% and training time reduction of ~10%.
- **Pioneered and led** the machine learning group within the Belle II experiment, organising regular meetings, guest speakers, and a central area for members to communicate and collaborate on research projects.
- **Reported regularly** at General Meetings in Tokyo, Japan, including **representing the machine learning group** with updates on projects and achievements.
- **Presented research output** at international conference venues for computing in high-energy physics.

EDUCATION

Ph.D., Physics (Grade: magna cum laude)   **2016 - 2019**

Ludwig-Maximilians-Universität München, Germany

Master of Science, Physics   **2013 - 2015**

The University of Melbourne, Australia

PUBLICATION HIGHLIGHTS

Deep Learning Approaches to Building Rooftop Thermal Bridge Detection from Aerial Images
Automation in Construction, 2023

Soiling Determination For Parabolic Trough Collectors Based on Operational Data Analysis and Machine Learning
Solar Energy (in review), 2022

Thermal Bridges on Building Rooftops (TBBR)
Nature Scientific Data (provisionally accepted), 2023

Learning Tree Structures from Leaves For Particle Decay Reconstruction
Machine Learning: Science and Technology, 2022

Punzi-loss: A non-differentiable metric approximation for sensitivity optimisation in the search for new particles
The European Physical Journal C, 2022

Accelerating neural network training with distributed asynchronous and selective optimization (DASO)
Journal of Big Data, 2022

HOBBIES AND INTERESTS

- Guitar: electric, blues, neo-psychedelia
- Photography: shot the social media campaign of a now sitting Greens federal MP in Germany
- Sports: Bouldering
- Technologies: Vim, Linux, Raspberry Pi