

Vangelis Kourlitis PhD

Data Engineering | Technical Leadership | ex-CERN

PROFILE

Data engineer and software developer with 10 years experience in various technical projects and leadership positions at CERN. Specialized in developing scalable pipelines in big data analytics and machine learning domains, with a proven track record of optimizing workflows and delivering impactful solutions. Skilled leader with extensive experience mentoring teams, managing R&D projects and authoring and reviewing technical reports. Now seeking to transition into tech industry to apply advanced technical expertise to business challenges.

CONTACT

Email: kourlitisV@hotmail.com

LinkedIn: vangelis-kourlitis

GitHub: ekourlit

SKILLS

Advanced
Python • C++

Intermediate
Unix • Git (CI/CD) • HTCondor •
Slurm • VTune

Novice
SQL • Docker • GCP • AWS • CUDA

EDUCATION

ATHENA RESEARCH CENTER

Athens Natural Language
Processing Summer School
2024 | Athens, GR

UNIVERSITY OF SHEFFIELD

PhD in Physics
2015 - 2019 | Sheffield, UK

ARISTOTLE UNIVERSITY OF THESSALONIKI

B.Sc. in Physics
2009 - 2015 | Thessaloniki, GR

EXPERIENCE

TECHNICAL UNIVERSITY OF MUNICH & CERN | Geneva, CH Analysis Model Group Coordinator

Oct 2023 - Dec 2024

- Led an international, diverse, and distributed 120-member group overseeing data analysis software (~1M open source LOC) and data formats, delivering features aligned with organizational goals.
- Drove the adoption of a lightweight data format ahead of schedule, offering 3x reduction in data storage costs.
- Rolled out streamlined software configuration paradigm, reducing onboarding time by over 95%.
- Planned and mentored multiple early-career developers projects, delivering innovative software products in data engineering on time and fostering professional growth.
- Promoted performant, array-based solutions for large-scale, unstructured, multi-modal data processing in advanced data analytics and AI workflows; organized tutorials using open source Python tooling across the US and EU.

Data Science Researcher

Mar 2023 - Sept 2023

- Awarded €40k grant to lead a 10-member team in modernizing legacy C++ ETL tools for large-scale (O(100TB)) data, implementing scalable array-based solutions that achieved 4x higher throughput and aligning with industry data science standards.

Python libraries: Awkward, Dask, Numba, CuPy

ARGONNE NATIONAL LABORATORY | Chicago, US

Data Science Researcher

Nov 2019 - Feb 2023

- Led a team of 7 researchers in applying advanced analysis algorithms and AI for complex data classification, increasing experimental reach by 12%.
- Facilitated collaboration among 10 international teams by establishing FAIR principles for AI models and documenting standardized methodologies for robust and explainable AI-driven results.
- Identified bottlenecks and doubled processing speed of large-scale C++ radiation simulations via computational optimizations.
- End-to-end developed 3D CNN computer vision model with multi-GPU training workflows to restore fast, low-accuracy sensor images and accelerate further simulation workflows by up to 20%.

Python libraries: PyTorch, TensorFlow, MLflow, Ray, PySpark

UNIVERSITY OF SHEFFIELD | Sheffield, UK

Doctoral Researcher

Nov 2015 - Oct 2019

- Authored the first technical report at CERN to open source a complete statistical model, setting a precedent for reinterpretation and transparency in the field.
- Developed and maintained C++ data analysis software for high-throughput computations across the full development lifecycle.

Python libraries: NumPy, pandas, SciPy, scikit-learn