

THEODOROS PANAGIOTAKOPOULOS

Ph.D Physicist ~ Data Engineer

SUMMARY

Data Scientist and **Ph.D.** Physicist building scalable **Deep Learning** models and end-to-end **data pipelines** to accelerate scientific discovery and reduce computational cost in **HPC** environments.

 TheoPhD.com

 teosfp@hotmail.com

 321 202 3216

 theodorospanagiotakopoulos

 Orlando, FL, USA

 TheoPhD

SKILLS

Languages: Python, Julia, R, C/C++, C#, SQL, Bash

AI Tools: Neural Networks, Decision Trees, Convolutional Neural Networks

Platforms: Linux, Git, HPC

EDUCATION

08/2019 - present **Ph.D: Computational Physics**
GPA: 4.0/4.0

University of Central Florida

10/2017 - 07/2019 **M.S.: Physics**
GPA: 9.2/10, Valedictorian

National and Kapodistrian University of Athens

10/2011 - 07/2017 **B.S. Physics**

National and Kapodistrian University of Athens

PROFESSIONAL EXPERIENCE

5/2025 – 8/2025 **Modeling Product Engineer Intern**

ASML, Silicon Valley, CA

- Engineered a **terabyte-scale CNN** analysis **pipeline**, scaling from **ThreadPoolExecutor** for I/O and **ProcessPoolExecutor** for compute to **Dask** to overcome DataFrame limits, **cutting analysis time** by **80%** and exposing critical overfitting.
- Trained a **ResNet-18 from scratch** in **PyTorch** with **metric learning (triplet loss)** to remediate uncovered **overfitting** by detecting **redundant or near-duplicate images** and rebalancing the **train/validation split**, **achieving target performance**.
- Implemented a **Physics-Informed Neural Network** to solve the 2D Helmholtz equation achieving high simulation accuracy as a **surrogate** for traditional solvers.

Python / C/C++ / Bash

5/2024 – 8/2024 **Modeling Product Engineer Intern**

ASML, Silicon Valley, CA

- Optimized** a geometric simulation parameter, reducing **runtime 5%** and **memory usage 10%** in production software used by **ASML customers**.
- Designed and executed** optimization experiments for a core optical parameter, cutting **runtime 10%** and **memory usage 34%** in **ASML's commercial release**.
- Developed **Python libraries** and **ETL pipelines** that were adopted by ASML's product engineering teams, which standardized large-scale simulation **data cleaning, preprocessing**, and **analysis**.

Python / C/C++ / Bash

8/2019 – present **Research Assistant**

University of Central Florida

DOE - NSF Funded

- Developed a two-stage **PyTorch** surrogate for **large-scale energy prediction** and **critical-threshold identification**, trained on high-fidelity numerical baselines and pivotal in securing **NSF** research funding.
- Implemented a **3D CNN in PyTorch** on **voxelized atomic data** to predict structural outcomes of deposition processes, **accelerating simulations** by producing accurate results in **seconds instead of hours**.
- Orchestrated a **distributed PySpark pipeline** converting **10K+ atomic configurations into 600K+ 3D voxel tensors**, eliminating data-ingestion bottlenecks for **3D CNN training**.
- Built a **Graph Neural Network pipeline** for incomplete graphs, achieving **94% accuracy** on sparse data and reducing computational overhead in material design simulations.
- Engineered a **high-performance C++ solver**, using hierarchical parallelism with **Kokkos** and **TBB** and compile-time unit safety with **Boost.Units**, delivering **strong scaling across CPUs and GPUs**.
- Created a centralized **SQL database** for multi-GB simulation datasets, improving data accessibility and accelerating validation across research teams.
- Automated **data preprocessing, feature engineering, and model training** by designing end-to-end **ETL and ML pipelines** in **Python** and **Bash**, improving iteration speed.
- Authored and maintained three **Python libraries** for fabrication modeling, electrochemical simulations, and 3D-to-2D visualization, adopted by university engineering teams.
- Designed and deployed an **end-to-end data pipeline** to evaluate CO₂ reduction catalysts, delivering **key insights** that guided experimental work and contributed to securing **DOE funding**.
- Established **analytical workflows** for complex **HPC** simulation data, uncovering **novel reaction mechanisms** and identifying **root causes** of enhanced catalytic performance in clean energy processes.

Python / Julia / C/C++ / Bash / SQL

GitHub

TECHNICAL SKILLS

- **Libraries & Frameworks:** TensorFlow, PyTorch, Scikit-Learn, XGBoost, Pandas, NumPy, SciPy, StatsModels
- **Big Data & Distributed:** Spark, Hadoop
- **Databases:** SQL (ETL, data integration, optimization)
- **Algorithms & Optimization:** Data structures, algorithm design, linear programming
- **Tools & Platforms:** Git, CI/CD, Linux, HPC

SELECTED - PUBLICATIONS

Effect of Ammonium-Based Cations on CO₂ Electroreduction

 Kaige Shi, Duy Le, **Theodoros Panagiotakopoulos**, Talat S. Rahman, Xiaofeng Feng

 2025  ACS

 link

Electronic structure of cobalt valence tautomeric molecules in different environments

 **Theodoros Panagiotakopoulos**, Esha Mishra, Thilini K Ekanayaka, Duy Le, Talat Shahnaz Rahman, Ping Wang, Kayleigh McElveen, Jared Paul Phillips, Zaid Zaz, Saeed Yazdani, Alpha T. N'Diaye, Rebecca Y. Lai, Robert Streubel, Ruihua Cheng, Michael Shatruk and Peter A. Dowben

 2022  Nanoscale

 link

Exploring Simulated Residential Spending Dynamics in Relation to Income Equality with the Entropy Trace of the Schelling Model

 **Theodoros Panagiotakopoulos**, George-Rafael Domenikos , Alexander V. Mantzaris

 2022  MDPI

 link

Direct and indirect detection of dark matter

 **Theodoros Panagiotakopoulos**, Vasilios Spanos

 2019  Pergamos library, National and Kapodistrian University of Athens

 link

Description of the method development for separating the Dalitz from the normal π^0 in the CDF detector

 **Theodoros Panagiotakopoulos**, Arkadios Manousakis

 2017  Pergamos library, National and Kapodistrian University of Athens

 link