Alessio Valentini, Ph.D.

acuzzio@gmail.com +1 (650)250-9263
www.alessiovalentini.com - linkedin.com/in/alessio-valentini-a6157479

Software engineer and computational chemist with 20+ publications and 2,400+ citations, combining deep scientific expertise with modern software engineering. Experienced in building molecular modeling platforms and ML pipelines, and in transforming academic prototypes into production-ready, high-performance tools for industry innovation.

Education

Ph.D. in Computational Chemistry. "Semiclassical molecular dynamics of photoactive molecules". University of Alcalá de Henares *(Spain, 2011-2015)*
M.Sc. in Chemistry. "Automatic generation of QM/MM models for photoactive proteins". University of Siena *(Italy, 2008-2011)*

Skills

Programming: Python, C++, Bash, CUDA, Cython, pybind11, Haskell, Fortran
Data & Analytics: Pandas, NumPy, Pydantic, BigQuery, PostgreSQL, SQLAlchemy, Alembic, JupyterLab, NetworkX
Visualization & UI: Matplotlib, Seaborn, Plotly, D3.js, Streamlit
Machine Learning: PyTorch, TensorFlow, scikit-learn, LightGBM
Infrastructure: AWS, GCP, Docker, Kubernetes, GitLab CI/CD, GitHub Actions, Snakemake
Computational Chemistry: GROMACS, AMBER, RDKit, OpenEye, Tinker, OpenMolcas, TeraChem, ORCA, xtb

Professional Experience

Principal Engineer, CreyonBio *(North Carolina 2023-Present)*

- Scaled an electronic-structure based featurization pipeline into a production-ready CPU/GPU workflow for ML models (Docker, Snakemake, Kubernetes on GCP), achieving 200x cost savings and 400x faster runtimes
- Migrated ML featurization workflow from ad-hoc file storage to a structured database system (SQLAlchemy, Pydantic), ensuring reproducibility and clean data lineage
- Designed custom chemistry tooling (Python, RDKit) to generate and align novel oligomer structures; enabled rapid implementation of new chemistries with built-in QC
- Delivered SELEX data-to-dashboard pipeline (Pandas, Streamlit, Dockerized on GCP), adopted by wet-lab groups for cross-run aptamer analysis
- Re-engineered ML training pipelines into modular MLOps framework; enforced metadata/versioning, accelerated hyperparameter search, and standardized interpretability (SHAP/GAMs)
- Combined scientific R&D with software engineering by developing a package for force-field optimization, translating ad-hoc researcher workflows on clusters into reproducible Kubernetes pipelines

Research Associate, Stanford University *(California 2020-2023)*

- Led development of the *Reactions Template Studio*, a graph-theory platform for automated exploration of chemical synthesis pathways. Integrated with *ab-initio* packages, enabling discovery of novel reaction routes under a multimillion-dollar ONR MURI grant
- Implemented polarization of QM/MM force fields by interfacing C++ and Fortran codebases, enabling GPU acceleration of both QM and MM components. Reduced compute times significantly and expanded system sizes accessible to simulation
- Launched an Alexa-powered cloud application to perform quantum chemical calculations requested via voice by end users

Postdoctoral Fellow, University of Liège *(Belgium 2016-2019)*

- Engineered a full Python/Cython package for quantum wavepacket propagation, delivering C-like performance with a Python API
- Analyzed and visualized terabytes of raw data with Python (NumPy, pandas, Jupyter, Matplotlib, Blender), producing interpretable insights into nonlinear quantum phenomena

Postdoctoral Fellow, University of Siena *(Italy 2015-2016)*

- Streamlined QM/MM workflows by automating complex input generation with a Python pipeline, reducing setup time from several days to a few hours
- Integrated multiple chemistry packages (ProPKA, GROMACS, Tinker) to allow high-level objectives (e.g., PDB screening, mutant generation) and deliver insights on fluorescence and excitation energies

Complementary Skills

- Scientific visualization and 3D rendering (Blender) - recipient of two awards and multiple conference presentations

- Functional programming enthusiast: contributions to molecular dynamics and data analysis libraries in Haskell
- Linux/HPC administration: managed two university clusters (50+ nodes)
- Advocate for SOLID principles and clean, maintainable code

Personal

- Fluent in Italian, English, and Spanish
- International experience: lived in 6 different countries