Dakota Folmsbee

Computational Chemist | Data Scientist | Materials Scientist

(+1) 802-683-4502 | dfolmsbee@gmail.com | dlf57.github.io | dlf57

Summary

Ph.D.-trained scientist with over 7 years of experience applying computational chemistry, machine learning, and data science to complex molecular and materials problems. Skilled in developing predictive models, analyzing simulation and chemical data, and applying cheminformatics tools. Proficient in Python, DFT, molecular dynamics, and modern ML methods. Strong collaborator and communicator with a track record of translating scientific insights into practical solutions.

Experience

Computational Materials Scientist *April 2024 - Oct. 2025*

Prometheus Materials, Longmont CO

- Developed machine learning models to predict material properties for concrete and biomineralization systems.
- Applied modeling and optimization methods to advance discovery of sustainable material compositions.
- Developed interactive ML dashboards for performance prediction and design optimization.
- Built and maintained data pipelines and centralized databases for R&D data.
- Collaborated cross-functionally to translate data insights into actionable materials solutions.

T32 Postdoctoral Scholar – Computational Drug Discovery *March 2022 - Feb. 2024*

University of Pittsburgh - Koes Group

- Built computational pipelines for molecular dynamics simulations of ion channels in membranes.
- Implemented enhanced sampling approaches to produce protein structures for binding interaction studies.
- Applied cheminformatics to analyze structure-function relationships affecting drug efficacy.
- Utilized pharmacophore searches and data analysis techniques to identify and prioritize promising drug-like compounds.
- Directed undergraduate research teams conducting computational modeling studies.

Graduate Researcher – Computational Chemistry & ML *Jan. 2017 - Feb. 2022*

University of Pittsburgh - Hutchison Group

- Conducted first-principles (DFT) calculations to predict optical, electronic, and dielectric properties of organic materials.
- Benchmarked ML models against ab initio methods for property prediction, enabling speed-ups in screening workflows.
- Designed novel molecular representation methods (chemreps) to improve ML accuracy for quantum chemical properties.
- Developed and applied genetic algorithms for material property optimization and dielectric material discovery.
- Implemented distance geometry method using quantum torsion as an alternative to crystal-structure-based methods.

Education

University of Pittsburgh *Aug. 2016 - Feb. 2022*
Ph.D. in Physical Chemistry *Pittsburgh, PA*

Clarkson University *Aug. 2012 - May 2016*
B.S. in Chemistry *Potsdam, NY*

Certifications

IBM Deep Learning with PyTorch, Keras and Tensorflow
Credential ID SJZLWDAK8H6T

Google Data Analytics
Credential ID WOADFVJ190RU

Projects

diamondfp *Aug 2025 - Present*

Developer

- Developed a Python library for generating statistical fingerprints of baseball players to enable similarity searches and ML.
- Technologies: Python, NumPy, Pandas, Polars, Scikit-Learn, Matplotlib, Git, Pytest
- https://github.com/dlf57/diamondfp

Concrete Mix Optimization Dashboard *April 2024 - July 2025*

Developer, Prometheus Materials

- Developed and deployed a web dashboard for exploring ML-driven predictions of concrete mix properties.
- Technologies: Python, Scikit-learn, Pandas, Plotly Dash

chemreps *Aug. 2018 - March 2020*

Developer

- Created an open-source Python package for generating molecular representations for ML models in chemistry.
- Technologies: Python, NumPy, cclib, qcelemental, Git, Pytest
- https://github.com/chemreps/chemreps

QM/MM Study Group *July 2018 - Dec. 2018*

Instructor & Organizer

- Designed and led a course teaching computational chemistry and Python fundamentals to graduate students.
- Technologies: Python, PySCF, Psi4NumPy, Git
- https://github.com/shivupa/QMMM_study_group

MolecularLearning *Sept. 2017 - Aug. 2018*

Developer

- Implementation of PhD work for creating custom molecular representations for ML.
- Technologies: Python, Open Babel, Scikit-Learn, Git, Pytest
- https://github.com/dlf57/MolecularLearning

Publications

2023

Folmsbee, D., Koes, D., Hutchison, G. Systematic Comparison of Experimental Crystallographic Geometries and Gas-Phase Computed Conformers for Torsion Preferences. *J. Chem. Inf. Model.* 2023. https://doi.org/10.1021/acs.jcim.3c01278

2022

Hiener, D., Folmsbee, D., Langkamp, L., Hutchison, G. Evaluating Fast Methods for Static Polarizabilities on Extended Conjugated Oligomers *Phys. Chem. Chem. Phys.* 2022. https://doi.org/10.1039/D2CP02375J

2021

Matlock, M., Hoffman, M., Dang, N., Folmsbee, D., Langkamp, L., Hutchison, G., Kumar, N., Sarullo, K., Swamidass, S. J. Deep Learning Coordinate-Free Quantum Chemistry. *J. Phys. Chem. A.* 2021. https://doi.org/10.1021/acs.jpca.1c04462

2021

Folmsbee, D., Koes, D., Hutchison, G. Evaluation of Thermochemical Machine Learning for Potential Energy Curves and Geometry Optimization. *J. Phys. Chem. A.* 2021. https://doi.org/10.1021/acs.jpca.0c10147

2020

Folmsbee, D., Hutchison, G. Assessing conformer energies using electronic structure and machine learning methods. *Int J Quantum Chem.* 2020. https://doi.org/10.1002/qua.26381

2019

D. Folmsbee, S. Upadhyay, A. Dumi, D. Hiener, & D. Mulvey. *chemreps/chemreps: Molecular Machine Learning Representations (Version 0.1.1).* 2019.Zenodo. http://doi.org/10.5281/zenodo.3333856

Skills

Programming Languages	Python, R, SQL, Bash, C++, Julia, LaTeX
Computational Chemistry	RDKit, Open Babel, Avogadro, Gaussian, ORCA, PySCF, xTB
Computational Biology	Amber, GROMACS, MDAnalysis, PyMOL, VMD, CHARMM-GUI, AlphaFold, GNINA
Python Libraries	Scikit-learn, PyTorch, Keras, TensorFlow, Pandas, NumPy, Plotly, Dash, Matplotlib
Data Science	Data cleaning, visualization, feature engineering, model evaluation & deployment
General	Git, GNU/Linux, SLURM, HPC environments, VIM, Jupyter, Microsoft Office Suite, Tableau