Jon Paul Janet

Resumé

I am a scientist applying data-driven techniques to drug discovery tasks at AstraZeneca. Previously, I was graduate student in the chemical and computational engineering Ph.D. program at MIT, advised by professors Heather Kulik and Youssef Marzouk. My research lies at the intersection of atomistic simulation, machine learning and optimization.

Experience

Research/Professional

2020– **Senior Scientist, Computational Chemistry & Machine Learning**, AstraZeneca.
develop[1] and apply generative reinforcement learning in combination with traditional structure[2]- and ligand-based[3] computational chemistry techniques to commercial early drug discovery projects. I also work with transformers for reaction informatics[4] and active learning to close the gap between simulation and generative modeling, and evangelize around data-driven drug discovery!

2015–2019 **Research Assistant, Machine Learning Materials Design**, Massachusetts Institute of Technology.
applied techniques from applied mathematics and machine learning combined with quantum chemical simulation for rational exploration and optimization in chemical space [5–17]. Featured in MIT News a, ACS Chemical & Engineering News b, c and MIT Energy Futures d and authored a book[18] titled Machine Learning in Chemistry as part of the ACS In Focus Series.

2014 **Internship, Computational Fluid Dynamics**, *Helmholtz-Zentrum Dresden-Rossendorf.*
investigated nucleation processes in dynamic 3D simulation of multiphase fluid flow[19]

2013 **Research and Teaching Assistant**, *University of the Witwatersrand.*
designed and conducted experiments on biological remediation of acid mine drainage[20]

2013 **Research Assistant**, *Process Modeling Group, University of Cape Town.*
designed reactor models for oil and gas processes

Teaching

2018 **Teaching Assistant**, *Massachusetts Institute of Technology.*
graduate computational chemistry course at MIT, 6.8/7 student rating

2017 **Teaching Assistant**, *Singapore University of Technology and Design.*
second-year differential equations and optimization course for engineers

2013 **Teaching Assistant**, *University of the Witwatersrand.*
first and second year chemical engineering core courses

2010-2013 **Mathematics & Physics Tutor**, *Intuition Tutors & Sum-It Maths.*
mathematics and physics tutor for final year high school students and university students

Education

2015–2019 **Ph.D. Chemical Engineering and Computation**, *Massachusetts Institute of Technology.*

2014–2015 **M.Sc. Scientific Computing**, *Technical University of Berlin.*

2013–2014 **M.Sc. Applied Mathematics**, *Royal Institute of Technology, Stockholm.*

2009–2012 **B.Sc. Chemical Engineering (Hons.)**, *University of Cape Town.*

Skills and projects

Mölndal – 43147 Västra Götaland, Sweden
jon.paul.janet@gmail.com • *jpjanet.io* • *jonpauljanet*

- **Programming**: High proficiency: Python, R, Matlab. Basic proficiency: C, MPI/high performance/parallel programming.
- **Machine learning**: Experience designing and training neural networks and kernel models and uncertainty quantification. Experienced with PyTorch, TensorFlow, Keras.
- **Computational Chemistry and simulation**: Experienced with quantum and classical simulation, calculation automation and integration with machine learning.
- **General**: Expert proficiency with LaTeX, TikZ, MS Office. Experienced with Linux, Windows & supercomputer administration, hardware and maintenance.
- **Open-source projects**: molSimplify – a python toolbox for inorganic molecular assembly and prediction & molSimplify Automatic design – a python toolbox for combining simulation and machine learning for molecular design
- Presented a workshop on machine learning in chemistry at Ben Gurion University (02/14/19), Hebrew University of Jerusalem (02/17/19) and MIT (03/29/19).

Selected Publications

[1] J. Guo, V. Fialková, J. D. Arango, C. Margreitter, J. P. Janet, K. Papadopoulos, O. Engkvist, and A. Patronov. Improving de novo molecular design with curriculum learning. *ChemRxiv*, 2021. doi: 10.33774/chemrxiv-2021-37m7w.

[2] J. Guo, J. P. Janet, M. R. Bauer, E. Nittinger, K. A. Giblin, K. Papadopoulos, A. Voronov, A. Patronov, O. Engkvist, C. Margreitter, and et al. Dockstream: A docking wrapper to enhance de novo molecular design. *ChemRxiv*, 2021. doi: 10.33774/chemrxiv-2021-qvhml.

[3] K. Papadopoulos, K. A. Giblin, J. P. Janet, A. Patronov, and O. Engkvist. De novo design with deep generative models based on 3d similarity scoring. *Bioorganic & Medicinal Chemistry*, 44:116308, 2021. doi: https://doi.org/10.1016/j.bmc.2021.116308.

[4] J. P. Janet, A. Tomberg, and J. Boström. Reusability report: Learning the language of synthetic methods used in medicinal chemistry. *Nature Machine Intelligence*, 3(7):572–575, July 2021. ISSN 2522-5839. doi: 10.1038/s42256-021-00367-2.

[5] J. P. Janet, C. Duan, A. Nandy, F. Liu, and H. J. Kulik. Navigating transition-metal chemical space: Artificial intelligence for first-principles design. *Accounts of Chemical Research*, 54(3):532–545, 2021. doi: 10.1021/acs.accounts.0c00686.

[6] J. P. Janet, S. Ramesh, C. Duan, and H. J. Kulik. Accurate multiobjective design in a space of millions of transition metal complexes with neural-network-driven efficient global optimization. *ACS Central Science*, 6(4):513–524, 2020. doi: 10.1021/acscentsci.0c00026.

[7] J. P. Janet, C. Duan, T. Yang, A. Nandy, and H. J. Kulik. A quantitative uncertainty metric controls error in neural network-driven chemical discovery. *Chemical Science*, 10:7913–7922, 2019. doi: 10.1039/C9SC02298H.

[8] S. Gugler, J. P. Janet, and H. Kulik. Enumerating de novo small inorganic complexes for machine learning and chemical discovery. *Molecular Systems Design & Engineering*, 5:139–152, 2020. doi: 10.1039/C9ME00069K.

[9] J. P. Janet, F. Liu, A. Nandy, C. Duan, T. Yang, S. Lin, and H. J. Kulik. Designing in the face of uncertainty: Exploiting electronic structure and machine learning models for discovery in inorganic chemistry. *Inorganic Chemistry*, 58(16):10592–10606, 2019. doi: 10.1021/acs.inorgchem.9b00109.

[10] C. Duan, J. P. Janet, F. Liu, A. Nandy, and H. J. Kulik. Learning from failure: Predicting electronic structure calculation outcomes with machine learning models. *Journal of Chemical Theory and Computation*, 15(4):2331–2345, 2019. doi: 10.1021/acs.jctc.9b00057.

[11] J. P. Janet, L. Chan, and H. J. Kulik. Accelerating chemical discovery with machine learning: Simulated evolution of spin crossover complexes with an artificial neural network. *The Journal of Physical Chemistry Letters*, 9(5):1064–1071, 2018. doi: 10.1021/acs.jpclett.8b00170.

[12] A. Nandy, C. Duan, J. P. Janet, S. Gugler, and H. J. Kulik. Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry. *Industrial & Engineering Chemistry Research*, 57(42):13973–13986, 2018. doi: 10.1021/acs.iecr.8b04015.

[13] J. P. Janet and H. J. Kulik. Predicting electronic structure properties of transition metal complexes with neural networks. *Chemical Science*, 8:5137–5152, 2017. doi: 10.1039/C7SC01247K.

[14] J. P. Janet and H. J. Kulik. Resolving transition metal chemical space: Feature selection for machine learning and structure–property relationships. *The Journal of Physical Chemistry A*, 121(46):8939–8954, 2017. doi: 10.1021/acs.jpca.7b08750.

[15] J. P. Janet, T. Z. H. Gani, A. H. Steeves, E. I. Ioannidis, and H. J. Kulik. Leveraging cheminformatics strategies for inorganic discovery: Application to redox potential design. *Industrial & Engineering Chemistry Research*, 56(17):4898–4910, 2017. doi: 10.1021/acs.iecr.7b00808.

[16] A. Bajaj, J. P. Janet, and H. J. Kulik. Communication: Recovering the flat-plane condition in electronic structure theory at semi-local dft cost. *The Journal of Chemical Physics*, 147(19):191101, 2017. doi: 10.1063/1.5008981.

[17] J. P. Janet, Q. Zhao, E. I. Ioannidis, and H. J. Kulik. Density functional theory for modelling large molecular adsorbate–surface interactions: a mini-review and worked example. *Molecular Simulation*, 43(5-6):327–345, 2017. doi: 10.1080/08927022.2016.1258465.

[18] J. P. Janet and H. J. Kulik. *Machine Learning in Chemistry*. American Chemical Society, Washington, DC, USA, 2020. doi: 10.1021/acs.infocus.7e4001.

[19] J. P. Janet, Y. Liao, and D. Lucas. Heterogeneous nucleation in cfd simulation of flashing flows in converging–diverging nozzles. *International Journal of Multiphase Flow*, 74:106–117, 2015. doi: 10.1016/j.ijmultiphaseflow.2015.04.005.

[20] J. P. Janet, K. Harding, C. Sheridan, and D. Drake. Increasing pumping depth in the long-term management of acid mine drainage. In *WISA 2014: Water Institute of Southern Africa*, South Africa, 2014.