

D. V. Klopfenstein, Ph.D.

email dvklopfenstein@protonmail.com
GitHub <https://github.com/dvklopfenstein>
location Philadelphia, PA

ORCID iD 0000-0003-0161-7603
LinkedIn www.linkedin.com/in/dvklopfenstein
mobile +1.267.844.1719

Summary

I aim to shorten the time in which scientific advances are translated from biomedical and biochemical research to improvements in human health by working with a team of microbiologists and biochemists to make transformative discoveries.

I am a scientist who productively employs the command line, Vim, Linux/UNIX, git, and Python to accomplish tasks in mining big data, bioinformatics, exploration, literature searches, and scientific writing. Additionally, I featured a tool written in the R programming language in a paper published in *Scientific Reports*.

Education

Ph.D. in Biomedical Engineering
M.S. in Biomedical Engineering
B.S. in Electrical Engineering

Drexel University, Philadelphia, PA
Drexel University, Philadelphia, PA
Rensselaer Polytechnic Institute, Troy, NY

Experience researching and writing peer-reviewed papers during the Ph.D.

“Evaluating retrieval qualities of Google Scholar and PubMed” *First author – Two citations**
2021 *Research Synthesis Methods* 2019 – 2021

- Wrote a commentary regarding an academic search paper in a format that is only by editorial invitation.
- Evoked a response from the original paper's authors that was published in *Research Synthesis Methods*.
- Created a new method for managing PubMed literature searches utilizing citation data from the National Institute of Health's Open Citation Collection (NIH-OCC).

“Cyclin-D1 integrates G9a-mediated histone methylation” *Middle author – 15 citations**
2019 *Oncogene* 2018 – 2019

- Won first place poster in the 2014 Jefferson Kimmel Cancer Consortium Symposium.
- Created novel whole genome visualizations utilizing matplotlib to examine hypothesis-related data.
- Modified Jefferson's hypothesis from “G9a/cyclin-D1 molecules bind to LAD regions” to “G9a/cyclin-D1 molecules bind to the edges of LAD regions.”
- Performed genomic data analyses utilizing bioinformatics tools such as bwtool (a tool for working with genomic data stored in the BigWig format), the Broad Institute's interactive genome viewer, BEDTools (tools comparing genomic features), and the BigWig applications from the University of California, Santa Cruz (UCSC).

“GOATOOLS: A Python library for Gene Ontology (GO) analyses” *First author – 329 citations**
2018 *Scientific Reports* 2014 – 2018

- Conceived and wrote the manuscript, resulting in 13 coauthors receiving 329 citations* each.
- Provided a high degree of international client service by dramatically expanding functionality in the GOA TOOLS open-source project and conversing with researchers.
 - Created a new method for grouping GO terms.
 - Expanded research interest as gauged by GitHub stars rising from ~40 to over 470.
 - Created Jupyter Python notebooks to demonstrate library usage.
- Promoted from GitHub contributor to collaborator based on my GO parser speed improvements (3,500%) and new reporting features which create easy-to-read Microsoft Excel spreadsheets. Collaborators have full permission to write changes in the project source; contributors must request change approval.
- Defined and answered open questions about gene ontology enrichment analyses (GOEAs) by creating 100,000 stochastic simulations and examining and interpreting the data.

Electrical engineering industry experience

L-3 Communications

Camden, NJ

Lead FPGA and ASIC Hardware Verification Engineer

2009 – 2013

- Held a Secret-level security clearance.
- Accomplished first-pass success in an application specific integrated circuit (ASIC) by architecting and implementing a verification environment with a user interface designed for engineers, which allowed the hardware architect and designer to discover over 100 hardware design bugs in an eight-month period.
- Found over 40 bugs within five months, more than any other engineer, in a field programmable gate array (FPGA) by creating and using a verification environment written in C++, VHDL, SystemC, and Python. These included bugs manifesting in hardware error registers, which can be challenging to find.
- Worked productively in small, focused engineering teams.

SGI

Tinton Falls, NJ

Lead ASIC Hardware Verification Engineer

2001 – 2009

- Architected and implemented proprietary cache coherent non-uniform memory access (ccNUMA) protocol simulation models and supporting infrastructure. The models both validated SGI's ccNUMA protocol architecture, which connects thousands of Intel processors in SGI super-computers, and verified the implementation of the ASIC hardware register-transfer level (RTL) logic design.
- Conceived, designed, and wrote a message sequence chart generator to visualize and communicate details in complex scenarios when investigating newly discovered design issues. Published and presented results at the DesignCon conference.
- Designed an easy-to-use stimulus entry method that enabled both senior architects (experts in ccNUMA protocol) and engineers (ccNUMA beginners) to create complicated system scenarios rapidly.
- Worked productively in a large, multidisciplinary, geographically distributed engineering team.

Open-source projects

GOA TOOLS

480 stars, 176 forks

Collaborator

<https://github.com/tanghaibao/goatools>

- Created the GOA TOOLS Python library to manage GO data and added it to GOA TOOLS.
- The library is trusted by over 107 projects, including 21 packages, as their software prerequisite.
- Successfully multitasked to add features requested by international researchers using the GitHub issues feature while also adding features required for my own research.

ReactomePy

11 stars, 3 forks

Owner, architect, and implementer

<https://github.com/dvklopfenstein/ReactomePy>

- Created a Python library to explore the biomolecular pathways in the Reactome Neo4j graph database.
- Built tools to access GO terms annotated to pathways and their components, which include molecular complexes, catalyst activities, reactions, and genomic sequences.

PMID Cite

4 stars

Owner, architect, and implementer

<https://github.com/dvklopfenstein/pmidcite>

- Added functionality to download data from the National Institutes of Health (NIH) that shows the likelihood a paper will be cited by a future clinical trial, indicating that the scientific advances described in the paper may potentially have a transformative clinical impact.
- Improved the selection of the next papers to examine in PubMed searches by utilizing citation and reference data, which is also known as forward/backward citation chaining or snowballing.
 - Architected and implemented customizable grouping and sorting of citing papers utilizing citation counts and performance data from the NIH-OCC.
 - Integrated the NIH-OCC API, a RESTful interface that utilizes HTTP requests, and the National Center for Biotechnology Information (NCBI) Entrez Programming Utilities (E-utilities) API into the *pmidcite* Python libraries.