# Samuel Wiley

678-596-1797 | samcwiley@gmail.com
linkedin.com/in/samcwiley | github.com/samcwiley

## EDUCATION

**Georgia Institute of Technology** *Atlanta, Georgia*
**Bachelor of Science** in **Physics**, GPA: 3.42 *May 2021*
Minor in **Chemistry**
Zell Miller Scholar, Leddy Family Scholar
**Certification** in Data Science, CDC Data Science Upskilling Program *Jul. 2024*

## EXPERIENCE

**Centers for Disease Control and Prevention (CDC)** *Jul. 2023 - Present*
*ORISE Fellow, Influenza Division Bioinformatics*

· Contributed to codebases and pipelines to support CDC's influenza surveillance and vaccine development
· Collaborated with internal and external partners to fulfill CDC's public health mission
· Presented on new skills and topics in bioinformatics for colleagues

**Georgia Institute of Technology, CHAOS Lab** *Jul. 2016 - Mar. 2020*
*Research Assistant, NSF Frontiers Grant, President's Undergraduate Research Award*

· Designed and conducted electrophysiology experiments to study heart arrhythmias and heart attacks
· Received grant for researching cardiac dynamics of regenerated zebrafish heart tissue
· Assisted in upkeep of lab equipment and materials

**Georgia Tech School of Physics** *Aug. 2016 - Dec. 2019*
*Teaching Assistant*

· Led lab and recitation sessions for two introductory physics courses and physics freshman seminar
· Graded assignments for Stellar Astrophysics, introductory classes, and physics freshman seminar

**Paper (Tutoring Company)** *Feb. 2022 - Jul. 2023*
*Tutor, Level 3*

· Provided structured academic support in math and sciences to over 4,200 students in grades 1 - 12
· Conducted quality assurance reviews of peer tutoring sessions
· Mentored peers to more effectively help students and progress the company's mission

## PROJECTS

**UDF-bioutils: Bioinformatics-Related User-Defined Functions for Cloudera Impala** Ongoing
*https://github.com/CDCgov/udf-bioutils*

· Developed and optimized custom user-defined functions in C++ for influenza SQL databases
· Implemented database functions based on user feedback and requests, including including allele sorting, Tamura-Nei pairwise distance calculation, and amino acid substitution site sorting

**IRMA-Core: Iterative Refinement Meta-Assembler** *Oct. 2024 - Present*
*Open-sourcing in progress*

· Collaborated with a team of four and stakeholders to re-write core components of the IRMA data pipeline in Rust for increased efficiency and accuracy
· Optimized algorithm for efficient primer and barcode trimming from DNA sequence reads and handled user command-line arguments

### SSWSort2 Multithreading, Grid Substitution, and Configuration     Nov. 2023 - Dec. 2024
*A tool for viral genome classification via Striped Smith-Waterman Algorithm in Rust*

· Implemented multithreaded and batch cluster job submission for efficient classification of large datasets
· Designed configuration files and configuration parser to store and implement user preferences for tool used in production

### Visualizing Influenza Evolutionary Trajectories in Fitness Landscapes   Sep. 2023 - Jul. 2024
*CDC Data Science Upskilling Program, 2024 Cohort*

· Collaborated with a team of four to create a data pipeline for analyzing trends between genomic, proteomic, and antigenic characteristics of influenza viral strains
· Designed and created an interactive data visualization dashboard using Plotly/Dash for communicating findings with Influenza Division leadership
· Presented on findings and results at the 2024 DSU Symposium and to Influenza Division leadership
· Completed four week-long bootcamps on data science, machine learning, modeling, and data visualization in Python and R

### Zoe Distances Module     Ongoing
*Zoe: An Open-source Rust Bioinformatics Crate*

· Contributed to open-source Rust crate for handling common bioinformatics functions and files for internal CDC pipelines and external users
· Created and optimized module for calculating nucleotide substitution distances including Jukes-Cantor, Felsenstein, Tamura-Nei, and General Time-Reversible (in progress, including linear algebra submodule) models

### Computational Compost: A Simulational Approach to Compost Thermodynamics   2018
*The University Physics Competition, Silver Medal*

· Developed a cellular diffusion model; created a computational framework to simulate a compost pile in Python with visualizations in MatPlotLib
· Worked in team to research, design and conduct simulation, and write a paper within 48 hr time limit

## SKILLS

**Languages, Scripting, and Mark-up:** Rust, Python, C++, Scala, R, Perl, BASH, SQL, LaTeX
**Software/Frameworks:** MATLAB, Linux, Git, Plotly/Dash, pandas, Fusion360
**Bioinformatics:** Sequence Alignment, Nucleotide Substitution Models, NGS Primer Trimming, Proteomics
**Computer Science:** Algorithm Development, Fuzz Testing, Unit Testing, Benchmarking, Data Analytics **Physics:** Computational Methods, Linear/Nonlinear Oscillations, Reaction-Diffusion Mechanisms,
**Equipment/Tools:** Electronics/Circuits, Microelectrode, Optical Mapping, Wood shop tools
**Chemistry:** Computational Methods, Density Functional Theory, Molecular Dynamics

## LEADERSHIP AND ACTIVITIES

**Bagpipes:** *Pipe Sergeant, Member at Large*, and *instructor* for the Atlanta Pipe Band
         Played at the 2024 Grade 1 World Championships with the City of Dunedin Pipe Band
**GT Society of Physics Students:** *Secretary*, Outreach Committee Member
**Singing:** *Section Leader* of GT Glee Club; *Staff Singer* and soloist at All Saints Episcopal Church