
Kye Stapleton-Gray

Bioinformatics / Software Engineering / Data Science / Data Engineering

New York City Metropolitan Area

kyevsg@gmail.com

<https://kye.org>

<https://github.com/kyevsg>

<https://www.linkedin.com/in/kyestapletongray/>

EDUCATION

Carnegie Mellon University, 2021, Pittsburgh, PA

B.S. in Chemical Engineering with a minor in Biomedical Engineering. Presidential Scholar, Dean's List

- ❖ Peer Tutor and EXCEL Leader for Calculus and Chemistry
- ❖ Completed many graduate classes, including Engineering Molecular Cell Biology, Entrepreneurship, and Image-Based Computational Modeling and Analysis, as well as special interest classes (Leadership Development)
- ❖ Ballroom Dance Club President (previously Lessons Coordinator, Events Coordinator, Public Relations Officer)

TECHNICAL PROFICIENCY

Python (including PySpark, pandas, Flask, Biopython, Scikit-learn, Matplotlib, Selenium, RDKit), R (including Seurat, vcfR, Bioconductor), MATLAB, SQL, HTML, CSS, JavaScript, AWS, Docker, Git, Linux, Unix, Bash, and more

EXPERIENCE

Specialist, Data Engineer / Axsome Therapeutics, June 2024 - Present, New York City, NY

- ❖ Develop ETL data pipelines on AWS Glue for field medical and market access data analysis
- ❖ Implement machine learning and AI strategies for medical insights and improving efficiency/productivity
- ❖ Improve AWS data lake house infrastructure, governance, and cost efficiency
- ❖ Work with internal and external clients to optimize company data solutions

Bioinformatics Software Engineer / Sangamo Therapeutics, November 2021 - January 2024, Remote

- ❖ Developed bioinformatics tools, libraries, and internal web apps to improve research efficiency and quality using Python, HTML, and other languages, and interfaced the programs with external software
 - Architected, developed, and deployed an app for qPCR analysis to be regularly used by 40+ researchers within the company
- ❖ Built automated testing for internal web apps with Selenium and Airflow
- ❖ Restructured data storage and querying tools for the internal protein archive (MongoDB and SQL databases)
- ❖ Presented a poster session on the internal Flask app I developed for qPCR analysis

Bioinformatics Intern / Nkarta, June 2021 - September 2021, South San Francisco, CA

- ❖ Analyzed and validated company genome sequencing analysis pipelines (single cell and whole genome sequencing)
- ❖ Used R and R packages including Seurat and vcfR to perform clustering analysis, confirm the accuracy of multiple pipelines by comparing our analyses to Illumina and 10x data, and identify discrepancies that were communicated back to Basepair to improve the company's genome sequencing analysis pipelines
- ❖ Presented findings to the translational group within the company

Data Analytics Intern / Phosplatin, May 2020 - August 2020, New York City, NY

- ❖ Analyzed study data from participants treated with phosplatin anti-cancer therapeutics, with Excel and Python scripts
- ❖ Identified key discrepancies and points of interest in data sets in order to improve the quality of data and present key findings to the company's executive leadership
- ❖ Developed Python scripts to expedite the data analysis and presentation, simplifying Excel data analysis methods and streamlining processes, and converting older code to be more broadly accessible for analysis

NSF REU Intern / Tierra Biosciences, June 2018 - August 2018, Berkeley, CA

- ❖ Worked at a biotech startup (then SynvitroBio) as an intern funded by the NSF REU Program
- ❖ Made use of cell-free systems in proof-of-concepts for screening various molecules of interest for drug discovery
- ❖ Tested and collected data on cell-free protein expression and presented findings to the company

Undergraduate Student Researcher / Carnegie Mellon University, August 2016 - December 2017, Pittsburgh, PA

- ❖ Made lipidoid nanoparticles in order to deliver siRNA to target cells, ran experiments to determine which lipids were most effective, and analyzed data from experiments to compare different lipids from a small lipidoid library
- ❖ Worked on oral drug delivery solutions for diabetes, <http://whitehead.cheme.cmu.edu>

Summer Intern / Radiant Genomics, June 2017 - August 2017, Emeryville, CA

- ❖ Identified sequences in the company's genetic library to be used for development of new antibiotics
- ❖ Initiated work on producing molecules from the identified sequences

Arthritis Foundation Summer Science Intern / UCSF, June 2016 - August 2016, San Francisco, CA

- ❖ Ran experiments to determine the resistivity of gut bacterial species to the rheumatoid arthritis drug methotrexate
- ❖ Analyzed the contents of the plates from the resistivity experiments to identify methotrexate-metabolizing bacteria
- ❖ Worked on a bacterial cloning project to identify a methotrexate metabolizing gene in a certain bacterial species
- ❖ Presented findings in a poster presentation at Stanford University

PUBLICATIONS

- ❖ Nayak RR, Stapleton-Gray K, O'Loughlin C, Fischbach M, Turnbaugh PJ. Methotrexate Is an Antibacterial Drug Metabolized By Human Gut Bacteria [abstract]. Arthritis Rheumatol. 2017; 69 (suppl 10). <http://acrabstracts.org/abstract/methotrexate-is-an-antibacterial-drug-metabolized-by-human-gut-bacteria-2/>.
- ❖ Nayak R, Alexander M, Stapleton-Gray K, Ubeda C, Scher J, Turnbaugh P. Perturbation of the Human Gut Microbiome by Methotrexate Contributes to the Resolution of Inflammation and Autoimmune Disease [abstract]. Arthritis Rheumatol. 2019; 71 (suppl 10). <https://acrabstracts.org/abstract/perturbation-of-the-human-gut-microbiome-by-methotrexate-contributes-to-the-resolution-of-inflammation-and-autoimmune-disease/>. Accessed January 3, 2020.
- ❖ Nayak RR, Alexander M, Deshpande I, Stapleton-Gray K, Rimal B, Patterson AD, Ubeda C, Scher JU, Turnbaugh PJ. Methotrexate impacts conserved pathways in diverse human gut bacteria leading to decreased host immune activation. Cell Host Microbe. 2021 Mar 10;29(3):362-377.e11. doi: 10.1016/j.chom.2020.12.008. Epub 2021 Jan 12. PMID: 33440172; PMCID: PMC7954989.

OTHER ACHIEVEMENTS

- ❖ Participated in the January 2024 Schrödinger "Hacking the Gender Stack" Hackathon and developed a cheminformatics app with my team that used Machine Learning and RDKit to predict molecular properties
- ❖ edX Certifications in: "Making Biologic Medicines for Patients: The Principles of Biopharmaceutical Manufacturing" and "Principles, Statistical and Computational Tools for Reproducible Data Science"
- ❖ IBM Data Science Professional Certificate and IBM Data Engineering Professional Certificate, which include training in subjects such as SQL, AWS, Machine Learning, AI, and Data Visualization
- ❖ Johns Hopkins Genomic Data Science Specialization through Coursera
- ❖ Golden Gate Section Society of Women Engineers (GGS-SWE) Scholarship