Stanford MEDICINE | Department of Biomedical Data Science

COLLABORATION & CAREERS FORUM

Resume Book

January 23rd 2024 | Stanford University | Stanford Alumni Center

Stanford MEDICINE | Department of Biomedical Data Science

WELCOME TO DBDS

Greetings!

We are delighted to share this resume book of graduate students in our Department of Biomedical Data Science. The resumes of our MS and PhD students are included, as well as resumes from students in other departments advised by our faculty and involved in closely related research areas. There are several postdoctoral researchers from our department included, too. You are welcome to reach out directly to students to discuss internships and full-time opportunities. Some students may be available soon, while others are interested in connecting for future possible engagement. We hope to see you at the January 23rd Collaboration & Careers Forum where you will have a chance to meet these students and connect with faculty as well.

Best Regards,
Sylvia Plevritis, Chair of the Department of Biomedical Data Science
Karen Matthys, Executive Director, Department of Biomedical Data Science

Stanford MEDICINE | Department of Biomedical Data Science

CONTENTS

students actively searching for full-time positions — page 4

students actively searching for part-time or internship positions — page 25

students interested in forming collaborations — page 55

All resumes appear in alphabetical order by last name

Stanford MEDICINE | Department of Biomedical Data Science

FULL TIME APPLICANTS

This section contains resumes of MS students, PhD candidates, and postdocs who are actively searching for full-time employment opportunities

January 23rd 2024 | Stanford University | Stanford Alumni Center

Kate Callon

kcallon@stanford.edu | linkedin.com/in/kate-callon/ | https://github.com/kcallon | (408) 806-9401

EDUCATION

Stanford University 09/2019-06/2024

- MS in Computer Science (Artificial Intelligence track)
- BS in Computer Science (Computational Biology track)
- Relevant Coursework: Artificial Intelligence, Natural Language Processing, Deep Learning in Computer Vision, Deep Learning in Biomedicine, Deep Learning with Graphs, Deep Multi-task and Meta Learning.

EXPERIENCE

Denali Therapeutics – Modeling and Informatics Intern 06/2023-09/2023

- Developing cheminformatic computational tools with Django to support greater efficiency in the drug design and discovery process for neurodegenerative diseases.
- Implemented a full-stack matched molecular pair tool to be used weekly by the small molecule discovery team to allow for quicker comparison and analysis and promote new design ideas.

Covert Lab – Undergraduate Researcher 09/2022-present

- Contributing to the creation of a computational whole-cell model for E. Coli through Stanford's Covert lab.
- Researching optimization techniques for Vivarium, a tool used for systems biology model integration.

Genentech – Early Clinical Development Business Operations Intern 06/2022-09/2022

- Developing data-driven approaches to promote greater efficiency and visibility for clinical trial operations.
- Built an interactive webpage with Javascript and Python for teams throughout Early Clinical Development to automate resource management information and highlight key resource changes.

Hiesinger Lab – Undergraduate Researcher 01/2020-03/2022

- Supported research to investigate solutions for ischemic heart disease with Stanford's Hiesinger lab, including a machine learning project involving echocardiograms.

RECENT PROJECTS

De novo Design of Synthesizable Fluorescent Dyes for Bioimaging with Generative AI, CS273B Final Project

- Modified an existing generative model to design promising novel fluorophores for experimental validation.

WC-VAE: A Variational Autoencoder for Characterization of Whole-Cell Modeling Data, Senior Project

- Developed a variational autoencoder that meaningfully characterizes cell features in a low-dimensional space.

Gradient Vaccine and PALS for NLP Multi-task Learning, CS224N Final Project

- Explored the addition of gradient vaccine optimization and PALs layers to BERT for multi-task learning.

Patient Co-Pilot, Winner of the Health Universe August 2023 Hackathon

- Built a tool leveraging LLMs to guide patients through pre and post-operative care.

SKILLS

Programming: Python, Django, Pytorch, HTML, Javascript, Git, AWS.

Leadership: Stanford Biotech Group leader, Advanced Volunteer for Santa Clara COVID Vaccine Effort.

Languages: Mandarin (basic proficiency)

Additional Information: Lived overseas for 11 years in Singapore and Japan.

GE (CLAIRE) CHENG

(669) 224-8184 · gchengcareer@gmail.com · https://www.linkedin.com/in/gechengclaire/

Work EXPERIENCE

LINKEDIN INC. Sunnyvale, CA

Senior Software Engineer, LTS Hiring Solution 2018-Present

- Spearheaded the design and development of the orchestration layer for LinkedIn's breakthrough Recruiter AI Copilot within a two-month timeline. This project marked LinkedIn's first venture into applications built on a Python framework. My work involved seamlessly integrating the sophisticated logic of LangChain and the Large Language Model, pioneering deployment strategies on Ray Server and ensuring fluid communication via gRPC
- Successfully migrated 10+ entities from a decade-old system to LinkedIn's contemporary platform, achieving **~100% data consistency** and **zero production issue**
- Led the design and implementation of "Remote Work" filter **utilized in over 50M searches** on Recruiter Search
- Initiated Recruiter Search feature to promote female engineer recruitment. Aligned with multiple teams and coded across API, mid-tier, backend, and search layers. Benefited over **13K job seekers** and **40 companies**
- Created efficiency-improving process consolidating several tech stacks with **200K+ requests** per second to one tech stack weeks ahead of schedule, thereby providing better candidate tracking experience

YAHOO INC. Sunnyvale, CA

Software Development Engineer, Search Supply-Side Platform 2016-2018

Data Analyst Intern 2015

- Launched Google NFP service on Mozilla, **decreasing abandonment rate by 2%**
- Used statistical, optimization, NLP and ML methods to **grow user experience by 2%** while maintaining revenue
- Improved ads placement model with data mining and **increased user engagement metrics by 7%**

EDUCATION

STANFORD UNIVERSITY Stanford, CA

Master of Science in Biomedical Data Science, Specialization: Artificial Intelligence *2025 (Expected)*

Honors Cooperate Program (part-time) GPA 4.1/4.3

UNIVERSITY OF CALIFORNIA, DAVIS Davis, CA

Dual Bachelor of Science, Computational Statistics and Applied Mathematics 2012-2016

Honors: Phi Kappa Phi Honor Society; Golden Key International Honor Society

PROFESSIONAL HIGHLIGHTS

Research Exp: **De-artifact MRI Scans –** Formulated a deep generative model to eliminate machine-related artifacts, enhancing radiologist interpretations without compromising scan quality
MRI Sequencing Classification – Investigated 20+ computer vision models, designed a two-stage MRI sequence classifier, achieving 75% accuracy on a 5-category task
Probing Language Models for Alignment with Human Reasonings – Introduced a framework to examine the hidden states of pre-trained models, generating natural language justifications

Projects: **Diffusion Tensor Images Analysis** – Examined statistical models to estimate tensors
Iterative K-means on MapReduce – Compared distance metrics to reduce running time by 70%
Friend Recommendation System – Implemented "People You Might Know" algorithm
Investment Timing based on Stock Performance – Built and evaluated pair trading model

Skills: **Programming** (Python, Java, Spark, R, Hive, SQL, C++, C, Matlab, Scala, Hadoop, Pig, Oozie, Shell, MapReduce); **Libraries** (Lucene, TensorFlow, PyTorch, Keras, Scikit-learn, Pandas); **ML Techniques** (Generalized Linear Models, Neural Networks, CNN, GAN, Transformer, Clustering, Transfer Learning, Ensemble); **Tools and Platforms** (Eclipse, IntelliJ, Git, MS Office); **Data Visualization**

Daisy Yi Ding

735 Campus Drive, Stanford, CA, 94305 | Phone: 312-889-4165 | dingd@stanford.edu

EDUCATION

Stanford University 2019 - 2024
Ph.D. Candidate in Biomedical Data Science, M.S. in Computational and Mathematical Engineering. GPA: 4.03
Stanford Graduate Fellowship. Advisor: Rob Tibshirani.
Research Focus: Multi-modal Biomedical Data Integration, Computational Biology, High-dimensional Statistics.

The University of Chicago 2017
B.S. in Statistics and Economics. *Phi Beta Kappa.* Dean's List all quarters. GPA: 3.96

SELECTED PUBLICATIONS

Cooperative Learning for Multi-view Analysis
D. Ding, B. Narasimhan, R. Tibshirani, *Proceedings of the National Academy of Sciences (PNAS),* 2022.

Machine Learning-guided Lipid Nanoparticle Design for mRNA Delivery
D. Ding, Y. Zhang, Y. Jia, J. Sun. *ICML Workshop on Computational Biology*, 2023.

Handling Missing Data with Graph Representation Learning
X. Ma*, J. You*, **D. Ding***, M. Kochenderfer, J. Leskovec. *NeurIPS*, 2020. (*Equal contribution)

NGBoost: Natural Gradient Boosting for Probabilistic Prediction
T. Duan, A. Avati, **D. Ding**, S. Basu, A. Ng, A. Schuler. *ICML,* 2020.

The Effectiveness of Multitask Learning in Phenotyping with Electronic Health Records Data
D. Ding, C. Simpson, S. Pfohl, D. Kale, K. Jung, N. Shah. *Pacific Symposium on Biocomputing,* 2019, Spotlight presentation.

Deep Learning for Chest Radiograph Diagnosis: A Retrospective Comparison of the CheXNeXt Algorithm to Practicing Radiologists
P. Rajpurkar, …, **D. Ding**, …, A. Ng. *PLoS Medicine*, 2018. Covered by *MIT Technology Review* and *Stanford News*.

Learning to Summarize Radiology Findings
Y. Zhang, **D. Ding**, T. Qian, C. Manning, C. Langlotz. *International Workshop on Health Text Mining and Information Analysis, EMNLP,* 2018, Spotlight presentation.

EXPERIENCE

Stanford University, Graduate Student Researcher, Stanford, CA 2017 - Present
- Developed computational methods for multi-modal patient data encompassing genomics, transcriptomics, proteomics, imaging, and electronic health records, leveraging inter-modal biological relationships to amplify signals. The approach enabled novel insights into molecular variations of human health, thereby enhancing disease mechanism understanding and enabling more personalized treatment strategies.

Insitro, Machine Learning for Drug Discovery Intern, South San Francisco, CA Summer 2022
- Developed machine learning algorithms for modeling progressive neurodegenerative diseases by harnessing the aligned signals from in vivo patient data and iPSCs-based in vitro data across multiple omics modalities.

Goldman Sachs, Quantitative Strategist Intern, New York, NY Summer 2017
- Analyzed large-scale corporate bonds data and proposed an automatic trading strategy based on the converging behaviors; the strategy I proposed was successfully adopted following my presentation to the global team.

ADDITIONAL INFORMATION

Programming: Python, R.
Teaching Assistant: Representations and Algorithms for Computational Molecular Biology (Graduate-level Course).
Paper Reviewer: *NeurIPS, ICLR, ICML, KDD, ISMB (Computational Biology), Machine Learning for Healthcare.*
Organizing Committee: *NeurIPS AI for Science Workshop* 2022, 2023.

Tiffany Eulalio

Biomedical Data Scientist

Contact

Phone
(808) 421-8306

E-mail
eulalio@stanford.edu

Technical Profile

Programming languages:
R, Python, C++, JAVA, CUDA, Linux
Bioinformatics analysis:
Bioconductor, PLINK, SAMtools
Statistical analysis:
PCA, Logistic regression, Random Forest, Elastic Net, LASSO

Competencies

Proficient in high-performance computing clusters and SLURM schedulers.

Effective communicator with strong record in research publications and presentations.

Collaborative team player in interdisciplinary research environments.

Extensive research experience; from designing to publishing.

Proven leadership and mentorship abilities, especially with students from diverse backgrounds.

Dynamic Biomedical Data Scientist and Stanford University Ph.D. candidate specializing in genetic research, with a keen focus on DNA methylation, statistical modeling, and computational cell-type deconvolution.

Work History

2018-09 – 2024-06 (expected)

Graduate Student Researcher

Stanford University, School of Medicine, Stanford, CA

- Researched gene regulation and expression to pinpoint mechanisms defining human traits.
- Led Alzheimer's disease research via computational cell-type deconvolution and multi-omic QTL mapping.
- Enhanced DNA methylation data summarization using Principal Components Analysis for Alzheimer's disease insights; developed "regionalpcs" R package, now featured on Bioconductor.
- Developed data analysis pipelines and garnered biological insights.
- Presented research at national conferences: ASHG, AGBT, NLM, and at Stanford University.

2021-01 - 2021-06

Teaching Assistant

Stanford University, School of Medicine, Stanford, CA

- Delivered lectures on Translational Bioinformatics: omics data, public biomedical resources, machine learning, data mining, drug discovery, and mobile/digital health.
- Facilitated engaging class discussions to foster an interactive learning environment.

2015-01 - 2018-08

Undergraduate Student Researcher

University Of Hawaii at Manoa, Honolulu, HI

- Engineered advanced parallel algorithms for GPUs at the AlgoPARC team, University of Hawaii, using C/C++ and NVIDIA's CUDA platform.
- Led the development of a GPU-adaptable algorithm for a parallel priority queue using recursion, optimizing network path calculations.

Education

2018-05 – 2024-06 (expected)

Ph.D.: Biomedical Data Science

Stanford University - Stanford, CA

2015-01 - 2018-08

BS & BA: Computer Science, Biology

University of Hawaii at Manoa - Honolulu, HI

Publications

- Co-authored five peer-reviewed and seven consortium papers.
- Google Scholar: https://scholar.google.com/citations?hl=en&user=MCfaK84AAAAJ

Isaac Gibbs

igibbs@stanford.edu ⋄ 650-680-6341

EDUCATION

Ph.D. in Statistics, Stanford University. 2019 - Spring 2024
Advisor: Emmanuel Candès.
Thesis Topic: Development of novel statistical tools for quantifying the uncertainty underlying predictions made by black-box machine learning models (e.g. neural nets). Applications to trustworthy machine learning, fairness.

B.Sc. in Math and Computer Science, McGill University. 2015-2019
Graduated with first class honours.

PUBLICATIONS

Gibbs, I. and Candès, E. (2021). Adaptive conformal inference under distribution shift. *Advances in Neural Information Processing Systems 34* (oral presentation). https://arxiv.org/abs/2106.00170.

Gibbs, I. and Chen, L. (2020). Asymptotic properties of random Voronoi cells with arbitrary underlying density. *Advances in Applied Probability*, 52(2), 655-680.

Gibbs I., Leavey K., Benton S.J., Grynspan D., Bainbridge S.A., and Cox B.J. (2019). Placental transcriptional and histologic subtypes of normotensive fetal growth restriction are comparable to preeclampsia. *American Journal of Obstetrics and Gynecology,* 220(1):110.e1-110.e21.

PREPRINTS AND WORK UNDER REVIEW

Gibbs, I., Cherian, J., and Candès, E. (2023+). Conformal prediction with conditional guarantees. *arXiv preprint.* https://arxiv.org/abs/2305.12616. *Under review at Journal of the Royal Statistical Society: Series B.*

Gibbs, I. and Candès, E. (2023+). Conformal inference for online prediction with arbitrary distribution shifts. *arXiv preprint.* https://arxiv.org/abs/2208.08401. *Major revision at Journal of Machine Learning Research.*

WORK EXPERIENCE

Meta Internship in the Central Applied Sciences group. Summer 2022
Developed statistical methods for using A/B test data to identify subsets of the population that showed a strong positive response to the treatment.

AWARDS

Stanford Statistics Departmental Teaching Assistant Award. 2023
Received for outstanding contributions as a teaching assistant during Ph.D. at Stanford.

Dr. Feng Qian Convocation Prize. 2019
Awarded to top graduating students in computer science at McGill University.

NSERC Undergraduate Student Research Award. 2018, 2019
Received separately in 2018 and 2019 at McGill University.

Sir Edward Beatty Memorial Scholarship and Emily Ross Crawford Scholarship. 2018
For academic performance in B.Sc. at McGill University.

ADITI GOYAL

adigoyal@stanford.edu
805-574-9846
6 Comstock Circle, APT 118B, Stanford CA 94305
GitHub: aditi-goyal-523

Bioinformatics Scientist with multiple years of experience in: *R, Python, NGS Data Analysis, workflow design and evaluation, Variant Calling & Analysis, Data Visualization, Dimensionality Reduction, Linux/Command Line, HPC environments, version control (Git), markdown documentation, multi-omic data analysis, Genomic tookits), Public Databases, Machine Learning, Statistical modeling, High-Throughput Data Management, Interdisciplinary communication*

EDUCATION

Stanford University, School of Medicine **Expected: June 2024**
M.S. Biomedical Informatics

University of California, Davis **June 2022**
Double B.S. Genetics & Genomics and Statistics

- UC Regents Scholar, University Honors Student, University Distinguished Scholar

RECENT RESEARCH EXPERIENCE

Bioinformatics Researcher | Endy Lab at Stanford ***May 2023 - present***

- Lead bioinformatician in project aimed to characterize genome-wide transcriptional response of Staphylococcus epidermis to medically relevant glucose levels, actively contributing to a manuscript as first author.
- Developed custom NGS pipeline using bash scripting to process microbial whole genome sequences, allowing for streamlined downstream analyses.
- Identified differentially expressed genes across various treatment groups using DESeq2 and customized R analyses
- Discovered novel enrichment patterns in biological pathways using gene set enrichment analysis by employing FGSEA, GSEA, and REVIGO pathway analysis tools.
- Conducted dose response analysis using transcriptomic data and various exponential and nonlinear models, resulting in 38 candidates for genetic glucose senors.

Bioinformatics Researcher | Montgomery Lab at Stanford ***Jan 2023 - present***

- Leading efforts to conduct allele-specific expression (ASE) analysis to investigate the effects of SNV's on humans ability to respond to endurance exercise training regimens.
- Designed custom Snakemake pipelines in HPC computing environment to process raw sequence data and extract allele specific imbalance at an individual level.
- Integrated multi-omic data (ATAC-seq, RNA-seq, METHYL-seq, WGS) to gain a deep understanding of how expression levels change in response to acute bouts of exercise.
- Contributing to a nationwide consortium (Molecular Transducers of Physical Activity Consortium) aimed at understanding how the body responds to exercise.

OTHER RELEVANT EXPERIENCE

- *The Yelper Helper: Providing recommendations to Yelp users based on their prior review history*
- *Editor in Chief at The Aggie Transcript at UC Davis (Sep 2018 – June 2022)*

PUBLICATIONS

Rodriguez-Medina, J., Kim, H. G., Castro, J., Contreras, C. M., Glon, C. L., **Goyal, A.,** Guo, B. Y., Knowles, S., Lin, J. C., McGuiness, C. L., Sorkin, E., Stefani, J., Yegireddi, S. J., Chaganti, S., Cui, D., Deck, S. L., Deokule, Y., Douglas, H., Kenaston, M., O'Brien, A., … Furrow, R. E. (2020). *Draft Genome Sequences of 16 Halophilic Prokaryotes Isolated from Diverse Environments.* Microbiology resource announcements, 9(8), e01540-19. https://doi.org/10.1128/MRA.01540-19

Benjamin, K*., **Goyal, A***., Nair, R., Endy, D. *Characterizing the genome-wide transcriptional response of Staphylococcus epidermidis to medically relevant glucose levels (*First Authors), (Manuscript in progress, 2023)*

Haghani, V., Zhang, A., **Goyal, A.,** Korf, I., Sharifi, O., La Salle, J. *"Rocketchip: A Comprehensive Bioinformatics Workflow for ChIP-Seq Data Analysis" (Pending Publication, 2024)*

Goyal, A. (2022). *Identifying R loops with DNA/RNA ImmunoPrecipitation sequencing technology.* The Aggie Transcript at UC Davis

Goyal, A. (2021). *Genetic Algorithms: An Overview of how Biological Systems can be Represented with Optimization Functions.* The Aggie Transcript at UC Davis

Goyal, A. (2020). *Applications of Machine Learning in Precision Medicine.* The Aggie Transcript at UC Davis

Clarisse Hokia

chokia@stanford.edu | (562) 316-4514

EDUCATION

Stanford University, Stanford, CA
Master of Science, Biomedical Data Science — Expected June 2024
Bachelor of Science, Computer Science - Biocomputation Pathway — June 2023
Relevant Coursework: Data Science for Medicine, Biomedical Systems, Languages to Information, Algorithms, Artificial Intelligence, Biostatistics, Computational Biology, Bioethics, Biology Foundations, Machine Learning

TECHNICAL PROJECTS

covLLM: Large Language Models for COVID-19 Biomedical Literature — Spring 2023

- Developed a novel generative LLM to summarize and extract user-specified information from scientific literature.
- Trained baseline LLaMA 7B model on datasets containing abstracts from the COVID-19 Open Research Dataset.
- Generated model that performed comparably with ChatGPT and outperformed baseline LLaMA 7B model.

Fraction — Winter 2023

Technologies Used: JavaScript, MongoDB, Express, React, NodeJS, HTML

- Created Fraction, a web application that calculates per-person totals for group purchases.
- Collaborated with a team of 4 other students as part of Stanford's Software Project Experience course.
- Refined full-stack development skills, utilizing the software development cycle.

Analysis of Medicare Payments for In-Patient and Out-Patient Services — Fall 2022

Technologies Used: SQL, Google BigQuery, Python

- Queried and visualized variations in Medicare payments based on geographical location and procedure type.
- Utilized machine learning methods to demonstrate factors contributing significantly to trends over time.
- Explored Medicare datasets as part of Stanford's Data Management and Data Systems course.

WORK EXPERIENCE

Deloitte, Hong Kong SAR — July 2023 – September 2023

Digital Consulting Intern

- Supported consulting team's data lead with modernizing database for a leading private hospital in Hong Kong
- Proposed database improvements based on client's existing database and HL7 FHIR standards
- Modeled iterations of database designs for client and infra team using entity relationship and UML diagrams

Stanford University Department of Cardiovascular Medicine, Stanford, CA — June 2021 – September 2022

Undergraduate Researcher

Advisors: Paul Wang (Principal Investigator), Meghedi Babakhanian

- Developed a novel treatment for cardiac arrhythmias utilizing high frequency ultrasound.
- Determined optimal designs for ultrasound transducers through Matlab simulations of ablation strength.
- Presented on full-time summer research through Bio-X (2022) and Bioengineering Department (2021) Programs.

LANGUAGES

English (fluent), Chinese (conversational), Tagalog (conversational), Spanish (elementary)

J. Weston Hughes

jwhughes@stanford.edu ❖ (781)-775-4166 ❖ San Francisco, CA ❖ github.com/weston100

GOALS

I'm a PhD student expecting to graduate in May, 2024, and looking for career opportunities starting in the Fall. I love building AI to improve medical care and broaden our understanding of human disease, and I'm experienced in designing tools to understand electrocardiograms and medical imaging modalities.

EDUCATION

Stanford University **Expected May, 2024**
PhD Candidate, Computer Science Advisors: James Zou & Euan Ashley
NSF Graduate Research Fellow

University of California, Berkeley **December, 2018**
Bachelor of Arts, Computer Science
GPA: 3.86

RESEARCH PROJECTS

Detecting 1,800 Diseases from The Electrocardiogram
- Training an ECG foundation model to detect 1,800 different diseases phenotyped from patient health records.
- Using results to guide development of novel ECG-based risk scores for varied diseases.

Simple ECG Models for Detecting Left Ventricular Systolic Dysfunction
- Demonstrated that small linear models and decision trees achieve accuracy close to published neural networks.
- Utilized as few as six parameters, versus millions in previous work.
- Currently in submission.

Predicting Cardiovascular Risk from Electrocardiograms
- Developed the SEER deep learning model to predict cardiovascular mortality and disease from ECGs.
- Designed secondary analyses to understand model generalization to single-lead setting and different sites and understand which features are important for model predictions.
- Outperforms and augments the ASCVD Pooled Cohort Equations, the current standard of care.

(See Google Scholar for publications)

WORK EXPERIENCE

Google **June 2022-September 2023**
Student Researcher
- Designed novel AI models to detect cardiac disease from Fitbit PPG and ECG signals, resulting in a patent.
- Worked with a team of researchers to improve atrial fibrillation deep learning algorithms to outperform feature-based methods.

Microsoft **May 2018-August 2018**
Research Intern

Grail **May 2017-August 2017**
Machine Learning Intern

Stanford University & University of California, Berkeley **August 2016-December 2022**
Course Assistant

SKILLS

Python, deep learning, artificial intelligence, PyTorch, TensorFlow, Pandas, SQL, AWS, Google Cloud, electronic health records, Electrocardiogram (ECG) data, PPG data, medical imaging data, teaching, leading research projects, writing research publications, creative writing, improv comedy, SCUBA diving instructor & Divemaster (NAUI).

Alex Loia

alex@loia.com
(406) 208-1478
Bay Area & Austin, TX
US & EU Citizen
linkedin.com/in/alexdloia
github.com/alexdloia
References upon Request

Languages & Frameworks

Python, C++, Java, C
PyTorch, TensorFlow
React, Angular/NgRx, Node
SQL, BigQuery, Firebase
GNU/Linux CLI

Projects

***BrainWave*, CS 210 First Place**
AI-powered meeting productivity tool that takes the hassle out of taking meeting notes and helps you stay focused and productive (**Node**, **Python**, **Firebase**, **OpenAI**)

***Representation Learning for Fast RL*, CS 230 Outstanding Project**
Reinforcement learning agents that learn from images can be improved by using representation learning to find important features (**PyTorch**, AWS)

***Smart Chunk Reader* (Question Answering)**
Dynamic chunk reader and a candidate model achieve high accuracy scores on the SQuAD 2.0 dataset (**PyTorch**, Azure)

SIMP-LLM
New AI model combining large language models and graph neural networks can predict drug-disease relationships with high accuracy. (**PyTorch**, BERT)

Hobbies

Snowboard, scuba, karate
Private pilot

My passions: Keep building, keep learning! Working with great people driven to build great things that ultimately bring us together and make our world a better place.

Professional

Software Engineering Intern, Google Cloud (Sunnyvale, CA: 6-9'23)
Implemented highly requested customer features for AI support agent design platform, Dialogflow CX (**Angular/NgRx**, Cloud AI & Industry Solutions)

- Agent-Level Route Groups: Enable users to reuse common conversational patterns as requested by Focus Accounts like PayPal and Wells Fargo, shaving days off the process of maintaining complex bots
- Workspace User Settings: Transform our app into a development environment for AI agents, empowering users to seamlessly pick up where they left off
- Saved users significant time and effort, reduced the risk of errors, and partnered with backend, UX, and PM teams to drive OKRs on unblocking Focus Accounts, reducing the cost and time to build bots, and promoting product excellence.
- Received spot bonus for designing and implementing long-awaited features

Software Engineering Intern, Google Cloud (Sunnyvale, CA: 6-9'22)
Full-stack improvements and optimizations of mission-critical long-haul optical transport design tool, Hauler. (**Python**, **Angular**, **SQL**, Google Global Networking)

- Accelerated frequent API call load times by 10x through cache implementation. Built utility to keep critical app data in sync with other internal network infrastructure teams. Mitigated common network design errors by enforcing business logic. Improved critical query performances by 80% via precomputation and optimization.
- Team commendation for impactful contributions after a short ramp-up time.

Cofounder/CTO, Hollr Technologies Inc. (Remote, 7'20-11'21)
Built browser extension platform that integrates with popular sites to streamline the sharing process and place relevant recommendations from trusted friends directly where decisions are made (**React**, Firebase, Google/Facebook APIs, Affiliate Marketing)

- Winner, 2021 Pear VC Student Competition. Spr 2021 Cohort, Cardinal Ventures, Stanford Univ.

Software Developer Intern, Macromoltek (Austin, TX: 8'18-9'19)
Rebuilt public website and dedicated application for antibody design and modeling at Y Combinator-backed startup, used by major pharmaceutical and biotech companies (Django/**Python**, HTML/JS, C#, PostgreSQL).

- Achieved 20% wetlab efficiency increase and 30% load time improvement with focus on accessibility.

Academics

Stanford University: BS Computer Science with Distinction (9'19-6'23), MS Biomedical Informatics (3'22-1'24 - expected)
BS Biocomputation Track, Tau Beta Pi Engineering Honor Society (GPA: 4.04). MS GPA: 3.98.

- President, Chi Alpha Christian Fellowship ('22-'23) & JKA Shotokan Karate ('21-). Resident Assistant, Cedro ('21-'22).
- **CS 210** Senior Project First Place: BrainWave (AI meeting productivity, partnership with BMW)
- **CS 124/224N**: NLP w/ Deep Learning (**PyTorch**); **CS 221/229/230**: AI & Deep Learning (**PyTorch/TensorFlow**).
- Projects: CS 230: Universal Representation Learning for Faster Reinforcement Learning - *Fall '21 Outstanding Project Award*. CS 224N: Smart Chunk Reader. BMI 212: Semantic Integration of Medical Publications and Large Language Models for Drug-Disease Link Prediction.

Gautam Machiraju

6044 College Ave, Apt C | Oakland, CA 94618 | 614-599-6288
gmachi@stanford.edu ◇ gmachiraju.github.io ◇ @gmachiraju

Education

Stanford University 2018 – 2024
Ph.D. Biomedical Informatics — Department of Biomedical Data Science (BDS)
Advisors: Profs. Parag Mallick (Radiology), Christopher Ré (Computer Science)
Committee: Profs. Sylvia Plevritis (BDS, Radiology), James Zou (BDS)
Funding: NIH NLM Training Grant, Stanford Data Science Scholarship, Canary-ACED Fellowship
Thesis: AI-driven data copilots for scientific discovery

University of California, Berkeley 2012 – 2016
B.A. Applied Mathematics (emphasis in Mathematical Biology), Minor in Bioengineering

Recent Work Experience

Graduate Student Researcher @ Mallick & Ré Labs, Stanford University 2018 – Present
(1) **Published** techniques to foster explainability capabilities in foundation models; applications identify class-specific regions in high-dimensional data and few-shot settings with only class labels: e.g. segmentation of tumors in gigapixel pathology & binding pockets in protein structures; (2) **published** explainability evaluation frameworks; (3) applying methods to identify novel prognostic biomarkers of cancer progression in multiplexed histopathology

Graduate Research Consultant @ Geldsetzer, Thieme, Baiocchi Labs, Stanford University 2021 – present
Applied computer vision for global health monitoring: (1) developing deep regression models to predict maternal & child health outcomes of remote villages in low- & middle-income countries using satellite imagery, government surveys, and other remote sensing data; (2) **published** app-based mpox detection with mobile-friendly vision models trained on social media corpus of skin lesion images; (3) developing few-shot detector of human trafficking compounds in the Brazilian Amazon

Student Researcher @ IBM Research Summer – Fall 2023
Submitted work on fact-checking vision-language foundation models for improved image tagging & retrieval

Graduate Teaching Assistant @ Department of BDS, Stanford University 2019 – 2020
TA for BIOMEDIN 214: Representations & Algorithms for Computational Molecular Biology (with Russ Altman); developed coding assignments & exams, led weekly office hours, delivered lectures on CS fundamentals in Python & deep learning for bioinformatics; received **teaching award** based on student evaluations

Bioinformatics Research Assistant @ Mallick Lab, Stanford University 2016 – 2018
Published mathematical model (ODEs) for biomarker shedding kinetics using tumor microenvironment data; simulation of serum biomarker viability; other projects include (1) **submitted** work on NLP of PubMed articles to construct biomarker database & (2) anomaly detection for multivariate time-series (multi-omics + wearables) to identify drivers of viral disease onset

Bioengineering SDE Intern @ Strateos (formerly Transcriptic) Summer 2016
Software engineering & scientific computing to refine robotic platform for wet-lab automation; projects in numerical optimization of robotic gantry movement and queueing & search of platform's job requests

Relevant Skills

Programming languages	Python (Pytorch, Tensorflow), shell [extensive] R, MATLAB [intermediate] SQL, CUDA, C++, Java, Spark, Scala, Javascript, HTML/CSS [basic]
Workflow	Cluster (Slurm), Cloud (GCP, Kubernetes), integrations (VSCode, W&B, Jupyter, rmate)
Design	LaTeX, vector graphics (BioRender, Adobe suite)
Research mentorship	Mentored 15 URM high school and college students on independent research projects

Keywords: Interpretability & explainability, foundation models, deep learning, weak supervision, AI4science
Further inquiry: Details on academic service and coursework (in BDS, CS, EE, Stats, Maths, etc.) can be found on Linkedin; preprints and publications, as well as my full CV, can be found on my personal website

Claire Martin

clma6962@colorado.edu

Current Address:
2245 Larkin St
San Francisco, California

Permanent Address:
3717 E 7th Avenue Parkway
Denver, Colorado

Education

Stanford University, Medicine - Palo Alto, California *September 2022 ~ Present*
Pursuing Master's (MS) in Biomedical Informatics **GPA : 4.00 / 4.00**

University of Colorado, College of Engineering - Boulder, Colorado *August 2016 ~ May 2020*
Bachelor of Science in Engineering **GPA : 3.80 / 4.00**
Major: Computer Science [Major GPA: 3.86] **Minor:** Financial Analytics

- **Regent Scholarship** : One Regent scholarship is awarded each year per Colorado high school based on high academic achievement as measured by GPA, rank in class, level of course work, and test scores
- **University of Colorado Esteemed Scholars Award (Joseph A. Sewall)** : Awarded the most prestigious academic scholarship offered by University of Colorado to in-state students due to proven academic excellence and standardized test scores
- **Engineering Merit Scholarship:** Awarded in conjunction of the Esteemed Scholars Award due to promising engineering scholars, who continuously prove to be exceptional in academics and contribute to the college as a whole

Professional Work Experience

Apple - San Francisco, California
Data Engineer & Scientist on ACWellness *July 2022 ~ Present*

- Lead development in platform to aid providers in next steps for patient care by extracting key ontologies in patient/provider discussions.
- Research methods to improve patient care through the intersect of their wearable health data and medical claims. Currently developing proactive interventions for patients expressing aberrational behavior from their "baseline health" through a collaborative filtering & anomaly detection approach.
- Own data delivery for team, develop pipelines for our providers to understand our patients' health trends in a secure and privacy oriented method (AWS infra / Airflow). Own data delivery to incorporate this data into clinic operations to enhance patient experience.

Komodo Health - San Francisco, California
Data Engineer & Scientist *December 2021 ~ July 2022*

- Won company wide competition pitching a network representation of patient pathways through their prescriptions, procedures and diagnosis. The work utilized Dijkstra's & Markov Chains to determine most frequent pathways and predictive behaviors from events. Bayesian belief network was later used to further refine the relationships amongst the events in patient journeys.
- Led a team to further develop and implement the above product. First iteration corroborated the research finding Mono increases the likelihood, & may be a necessary precondition of MS by exploring patient pathways with and without the diagnosis.
- Developed customized product deliveries through scalable data pipelines for our customers using Snowflake, DBT, & Pyspark

Apple - San Francisco, California
Data Engineer Siri Reliability *August 2020 ~ December 2021*

- Drove a task force to research and analyze user behavior associated with different types of user interactions. Analysis was used to aid in the development of an adaptive model to predict type of requests. Current model is a boosted tree classifier, created through Turi Create and tested/deployed in Bolt.
- DRI for Siri Reliability team in latest Siri OS release. Work entailed root cause analysis in regressing aspects of the product, and working with cross functional teams to ensure fixes are implemented prior to release.
- Designed, created and maintained Python/Spark data pipelines for parsing and storing Siri logs (billions of requests daily). ETL is scheduled through Airflow. Derived fields from logs are stored in Iceberg tables that live in AWS.
- Transitioned team's infrastructure from Scala pipeline that fed into Hive tables in MS05 (scheduled through Oozie). New design allowed for higher scalability and flexibility with schema evolution.

Software Engineer Siri Search *May 2019 ~ August 2019*

- Developed full stack Flask web application with subscription service; which compiled test analysis reports for Siri software releases.
- Subscription included recommendations for tests analysis based off of test suite relevancy features and personal user background.
- Systematic query analysis and reports created through application reduced review of 200,000 test reports weekly.
- Won company wide internship competition for best new idea for Apple product or feature.

Other

Favorite Project: Kanye's Growth and Development *May 2018 ~ June 2018*

- Scraped and statistically analyzed data on Kanye's musical, social media, and news presence over the course of his career
- Developed a classification system attempting to separate "Old" versus "New" Kanye," based off of sentiment value of his lyrics and press releases
- Created a network system to analyze Kanye's professional collaboration and influences over his career to find correlations with sentiment changes
- Defined transition of statistical significance in Kanye's lyricism and public appearance in 2016, with cause supported by the infamous trip to Paris

Technical/Personal Skills

- *Languages/Tools:* C++, C, Python, Scala, Airflow, Spark, Oozie, Presto, Git, VBA, Flask, Javascript, HTML, CSS, SQL, Unix, MatLab
- Coursework: Discrete Structures, Theory of Computation, Software Development, Artificial Intelligence, Data Science: Statistical Modeling, Big Data Analysis, Data Mining, Machine Learning, Medical Ontologies, Genetic Phenotyping
- Skills: Data Wrangling / Analysis, Data Communication, Public Speaking, Personable, Full Stack Development,

Samson Mataraso

San Francisco Bay Area
(925) 822-4948 / samson920@gmail.com

EDUCATION

2020-Present *PhD, Biomedical Data Science*, Stanford University, Stanford, CA
GPA: 4.2/4.0; NSF Graduate Research Fellow

2015-2018 *B.S. EECS*, *Minor in Bioengineering*, University of California, Berkeley, Berkeley, CA

EXPERIENCE

2022-Pres. AI/ML Fellow, Longitude Capital, Menlo Park, CA
- Conduct comprehensive landscape analysis and designed a proprietary evaluation framework for AI/ML-based biotech companies
- Diligence companies utilizing AI/ML in biotech or healthcare

2022-Pres. Entrepreneurial Playbook Fellow, Nucleate
- Interviewed dozens of academic spinout founders and VCs
- Writing a playbook on "Operationalizing a Biotech Company out of Academia"
- This playbook will be used as a resource for all Nucleate Activator startups

2020-Pres. Graduate Student, Prof. Nima Aghaeepour, Department of Anesthesiology, Stanford University, Stanford, CA
- Develop novel machine learning methods for the analysis of multi-modal datasets
- Lead data analysis and interpretation for interdisciplinary, collaborative projects
- Design in vitro and in vivo proof of concept studies for therapeutic I co-developed
- Teaching assistant for "Algorithms for Computational Molecular Biology"

2015-2020 Data Scientist; Head of Data Science, Dascena, Inc., Oakland, CA
- Designed and coordinated one of the first randomized clinical trials of a EHR, machine learning-based diagnostic, showing ~58% reduction in ICU mortality
- Data science and product lead as the company raised $50M
- Led team which generated data resulting in an FDA Emergency Use Authorization for an algorithm predicting complications due to COVID-19
- Managed and led partnerships with Fortune 500 companies like Danaher and JnJ

COMMUNITY SERVICE

2020-Pres. DEI Committee, Department of Biomedical Data Science, Stanford University
I am part of a student run DEI committee that contributes to various efforts to improve diversity, equity, and inclusion in biomedical data science at Stanford

2021 Coding Bootcamp Instructor, Stanford Summer Research Program
I developed and taught an R class to undergraduates from historically underrepresented groups who came to Stanford to perform bioscience research.

2021 Mentor, Inclusive Mentoring in Data Science
Mentor for the pilot offering of this program that paired students from colleges that do not have lots of research opportunities with PhD students at Stanford

Ali Mottaghi

AI in Healthcare | PhD at Stanford

@ mottaghi@stanford.edu +1 (650) 665-3661 in ali-mottaghi alimottaghi @samottaghi

Experience

AI Engineer
Hippocratic AI
Jun 2023 - Aug 2023
Palo Alto, CA

- Orchestrated the development and optimization of fine-tuning algorithms for Hippocratic AI's expansive healthcare-centric language models, achieving unprecedented levels of clinical accuracy and compliance.
- Introduced and seamlessly integrated Retrieval Augmented Generation (RAG) algorithms, dramatically enhancing real-time, context-sensitive medical information retrieval capabilities.
- Formulated and executed specialized methodologies for data labeling and curation, explicitly tailored for large language models in healthcare, thereby ensuring industry-specific accuracy and adherence to ethical standards.

Fine-Tuning LLMs | Retrieval Augmented Generation (RAG)

AI Scientist
Intuitive Surgical
Jun 2021 - Sept 2022
Sunnyvale, CA

- Engineered state-of-the-art surgical activity recognition algorithms, transforming data capture and analytics in operating rooms.
- Innovated a cutting-edge domain adaptation algorithm, achieving a 5% performance boost across multiple operational environments, including different operating rooms and hospitals.

Publication: Adaptation of Surgical Activity Recognition Models Across Operating Rooms. **Ali Mottaghi**, et al. MICCAI 2022. **Paper**

Domain Adaptation | Video Understating | Data Efficiency

ML Scientist
Curai Health
Jun 2020 - Sept 2020
Palo Alto, CA

- Revamped the Curai Health chatbot's medical symptom recognition algorithms, achieving an exceptional 20% performance improvement.
- Conceived a novel active learning technique specifically for handling long-tailed dataset distributions, providing a robust solution for class imbalance.

Publication: Medical symptom recognition from patient text: An active learning approach. **Ali Mottaghi**, et al. ML4H 2020. **Paper**

LLM Chatbots | Data Engineering | Hugging Face | PyTorch

Graduate Research Assistant
Medical AI and Computer Vision Lab
Sept 2018 - Current
Stanford, CA

- Crafted an innovative active learning framework utilizing deep generative models, resulting in a 10% performance improvement while utilizing 30% less training data.
- Collaboratively authored a comprehensive survey on cutting-edge developments in medical computer vision, providing invaluable insights for professionals in the field.

Publication: Deep learning-enabled medical computer vision. Andre Esteva,... **Ali Mottaghi**, et al. npj Digital Medicine 2021. **Paper**

Generative AI | Medical Computer Vision | Technical Writing

Education

Ph.D. in Electrical Engineering
Ph.D. Minor in Computer Science
Stanford University
Sept 2018 - Current Stanford, CA

Advised by Serena Yeung and affiliated with the Stanford AI Lab.

Academic Services: Reviewer for JBHI 2020, AISTATS 2021, ML4H 2021, IPCAI 2022, NeurIPS Datasets Track 2021, 2022

M.S. in Electrical Engineering
Stanford University
Sept 2018 - Jun 2020 Stanford, CA

Classes: AI-Assisted Healthcare (CS337)
Building for Digital Health (CS342)
Biodesign Fundamentals (MED275b)
Startup Garage (Auditing STRAMGT377)

Teaching: AI in Healthcare (CS271)
Linear Dynamical Systems (EE263)
Probabilistic Systems Analysis (EE178)

GPA: 3.99/4.0 with 150 credits

B.Sc. in Electrical Engineering
Sharif University of Technology
Sept 2014 – Jul 2018 Tehran, Iran

Ranked 1st with GPA 19.79/20

Skills

Python
C and C++
R
Tensorflow and Keras
Pytorch

Awards

- **Stanford Graduate Fellowship (SGF) in Science and Engineering**
  highest award offered to incoming Stanford graduate students
- **Intuitive Surgical Fellowship**
  supporting PhD students
- **Ranked 1st in the class of 2018**
  in Sharif University of Technology EE Department (nearly 200 students)
- **ETH Zurich Student Summer Research Fellowship**
  summer 2017

Minh Nguyen

(415) 341-2096 minh084@stanford.edu https://www.linkedin.com/in/minh084/

EDUCATION

2018 – current Ph.D student in Biomedical Informatics, Stanford University
Ph.D minor in Management Science and Engineering
2013 – 2016 M.A, Biostatistics, University of California, Berkeley
2007 – 2009 B.S, Nursing, San Francisco State University

ACADEMIC EXPERIENCE

09/2019 – 12/2020 ***Teaching Assistant***, Stanford University
Assisted Professor Shah in teaching the Data Science for Medicine course.

05/2015 – 07/2016 ***Graduate Student Researcher***, University of California, Berkeley
Collaborated with UCSF/SFGH surgeons to examine data from a multi-site, randomized clinical trial of severely injured patients from 12 Level I trauma centers in North America.

01/2016 – 05/2016 ***Statistical Consultant***, University of California, Berkeley
Provided consulting as an advanced graduate student under Professor Huang's supervision to the UC Berkeley community and general public on their research projects and work-related issues.

08/2014 – 05/2016 ***Graduate Student Instructor***, University of California, Berkeley
Led sections for undergraduate statistics courses, reviewing and teaching concepts of statistics.

PROFESSIONAL EXPERIENCE

08/2010 – current ***Clinical Nurse***, University of California, San Francisco Health

PUBLICATIONS

Nguyen, M., Morrison, T., Owen, A., Baiocchi, M. Tie-breaker Designs for a Pragmatic Clinical Trial. *American Causal Inference Conference (ACIC)* 2023. Honorable Mention Poster Award

Nguyen, M., Eulalio, T., Marafino, B. J., Rose, C., Chen, J.H, Baiocchi, M. Thick Data Analytics for Iterative Model Development. *American Medical Informatics Association. AMIA Informatics Summit proceedings* 2023. *Under revision at the American Statisticians*

Nguyen, M. et al. Developing Machine Learning Models to Personalize Care Levels among Emergency Room Patients for Hospital Admission. *Journal of the American Medical Informatics Association*; 2021;28(11):2423-2432.

Nguyen, M. et al. Machine Learning for Initial Insulin Estimation in Hospitalized Patients. *Journal of the American Medical Informatics Association.* 2021;28(10):2212-2219.
Noshad, M. et al. Context is Key: Using the Audit Log to Capture Contextual Factors Affecting Stroke Care Processes. *American Medical Informatics Association. Annual Symposium proceedings. AMIA Symposium* 2020; 953–62.

Nguyen, M. et al. Dynamic Impact of Transfusion Ratios on Outcomes in Severely Injured Patients: Targeted Machine Learning Analysis of the PROPPR Randomized Clinical Trial. *The Journal of Trauma and Acute Care Surgery*. 2020;89(3):505-513

AWARDS & ACTIVITIES

2019 – 2020 Stanford BioDesign Innovation Award, BioDesign Next extension funding
2021 – 2022 Student representative, Stanford's Biomedical Informatics graduate program

Eric Pan (eric.pan@stanford.edu) LinkedIn: /in/ericpan64
Engineering for healthier living GitHub: /ericpan64

Education

- **Stanford University** Stanford, CA / Remote (HCP)
  M.S. Biomedical Data Science, 3.75 *Jan 2022 – Present*
  - **Coursework**: Data Science for Medicine, Stanford Health Care Consulting Group, Human-Centered NLP, Data Fusion in Biomedicine, Directed Research (HealthRex Lab, Shah Lab), Designing Healthy Behaviors (d.school)
- **Georgia Institute of Technology** Remote (OMSCS)
  M.S. Computer Science, 3.8 *Jan 2019 – Dec 2020*
  - **Coursework**: Big Data for Health (link), Education Technology (link), Health Informatics, Information Security, Artificial Intelligence, Graduate Algorithms, High-Performance Architectures, Human-Computer Interaction
- **University of California, Los Angeles** Los Angeles, CA
  B.S. Bioengineering, Bioinformatics Minor, 3.65 *Sept 2014 – Aug 2018*
  - **Coursework**: Medical Decision Making, Bioinformatics Algorithms, Systems Biology, Biocompatibility, Neuroengineering, Biomedical Transducers, Cell Culture Lab, Thermodynamics, Machine Learning

Work Experience

- **Stanford Medicine & Stanford Health Care** Stanford, CA
  Research Assistant *Apr 2023 - Aug 2023*
  - **ML on FHIR**: Developed an Epic API integration for mapping real-time data for use in machine learning models between Stanford Health Care and the Shah Lab.
- **Georgia Institute of Technology** Remote
  Instructional Assistant (CS6440: Intro to Health Informatics, Part-Time) *Jan 2021 - Aug 2023*
  - **Graduate TA**: Mentored students on various full-stack healthtech applications ranging from medical records, applied machine learning, nutrition, wearables, price transparency, and other healthcare-related topics.
  - **Course Content**: Developed graduate-level course material, Gradescope autograders, and mentoring student projects. Course content includes EHR workflows, medical ontologies, SMARTonFHIR, and healthtech projects.
- **Canvas Medical** Remote
  Senior Software Engineer *Nov 2021 - Feb 2023*
  - **API Refactor Architect**: Co-led refactor of our FHIR API using modern Python and functional programming concepts with a heavy focus on performance and developer usability (leading to clinician usability, link here).
  - **Python Library Developer**: Designed and developed Pydian (link) and advocated for open-sourcing.
  - **FHIR API Development**: Developing core FHIR features using tools like FastAPI, Mirth, and GraphQL.
  - **Diff Tool Utility**: Designed and developed API request replay tool leveraging subprocesses and PostgreSQL data to ensure secure data migration and prevent several critical bugs missed from initial team QA.
  - **ONC Certification**: Designed and developed solutions focusing on core components of ONC EHR certification.
- **1upHealth** Cambridge, MA
  Data Software Engineer *Mar 2021 - Nov 2021*
  - **ETL Infrastructure**: Designed, reviewed, and implemented multiple FHIR mappings and data flows for multiple healthcare organizations using Apache NiFi. Led design reviews on architecture improvements.
  - **Technical Domain Expert**: Led training sessions on the entire ETL pipeline (AWS Setup, Data Ingestion, FHIR Mapping, Apache NiFi) to sync existing and onboard new team members for urgent July 1 deadline. Debugged and improved AWS deployments via the admin console and Terraform scripts (EC2, S3, RDS, VPC, Lambda).
  - **Internal Projects**: Scoped and developed several end-to-end projects including a SQL DDL generator, a PDF parser, and a FHIR spec web scraper. Primarily used Python, Docker, and Jupyter Notebooks.
- **Epic Systems** Madison, WI
  Electronic Data Exchange (EDI) Engineer *Sept 2018 - Mar 2020*
  - **Primary Technical Support**: Resolved 100+ support tickets across functional areas as the primary integration expert. Organized weekly calls and managed new integration projects across customers.
  - **Clinical Foundation System Coordinator**: Designed fixes and triaged 75+ development requests for clinical interfaces to improve implementation time with Epic's Foundation System.
  - **Beaker Pathology Data Conversion**: Implemented and debugged lab results interfaces to convert 10+ years of pathology report data into Epic for a large academic health system.

Rex Shen

Office: 390 Jane Stanford Way, Stanford, CA 94305, USA
rshen0@stanford.edu • +1 (952) 905-0268

EDUCATION

Stanford University, Stanford, California, USA

- Statistics Ph.D. Student — Sep 2022 – Jun 2025 (Expected)
  - Interests: Synthetic Data Generation/Perturbation, Generative AI
- M.S in Statistics — 2020 – Jun 2022
  - GPA: 4/4
- B.S.H in Mathematical and Computational Science with Honors, Graduated with Distinction — Sep 2018 – Jun 2022
  - Honors Thesis: "A Seasonality-Adjusted Sequential Test for Vaccine Safety Surveillance"
  - Advisors: Professors Bradley Efron, Lu Tian
  - GPA: 4.02/4

INTERNSHIP EXPERIENCE

Biostatistics Intern at Daiichi Sankyo — May 2021 – August 2022

- Project 1: Basket Clinical Trial Design
  - Design Bayesian Statistical Models in R for pooling patients with similar and different cancer types for clinical study
- Project 2: Algorithm Development and Data Visualization using R Shiny
  - Developed simulated patient data from Kaplan-Meier Plots in R
  - Built interactive web applications using R Shiny
  - Ran statistical models on clinical trial data for a lung cancer drug using R

Data Scientist and Software Engineer at UnitedHealth Group, Optum — June 2019 – August 2020

- Project: Automated HealthCare Fraud Detection
  - Developed ML Algorithms in Python and a Graph Convolutional Network using Graph Databases (i.e. TigerGraph) for real-time fraud detection
  - Presented findings in a "Shark Tank" format to Senior Management, including VP
  - Implemented Autoencoder NN in Python to further improve prediction for real-time fraud detection

PAPERS

PUBLICATIONS

[1] R. Shen, K. Moll, Y. Lu, and L. Tian. A Seasonality-Adjusted Sequential Test for Vaccine Safety Surveillance. *Biometrics*. Jan 2023.

[2] X. Shen, X. Bi, R. Shen. Data Flush. *Harvard Data Science Review, 4(2)*. May 2022.

[3] R. Shen, L. Luo, and H. Jiang. Identification of Gene Pairs Through Penalized Regression Subject to Constraints. *BMC Bioinformatics, 18(1), 466*, Nov 2017.

[4] R. Shen. Gene Pair Analysis for Brain Tumors. *Siemens Foundation and Discovery Educ.*, Oct 2017.

SUBMITTED MANUSCRIPTS/MANUSCRIPTS IN PROGRESS

[1] Y. Liu, R. Shen, and X. Shen. Perturbation-Assisted Sample Synthesis: A Novel Approach for Uncertainty Quantification. *Undergoing Revision in IEEE Transactions on Pattern Analysis and Machine Intelligence*, May 2023. https://arxiv.org/abs/2305.18671

[2] X. Shen, Y. Liu, and R. Shen. Boosting Data Analytics with Synthetic Volume Expansion. *Undergoing Revision in The Annals of Applied Statistics*, October 2023. https://arxiv.org/abs/2310.17848

[3] Haiming Zhou, Rex Shen, Sutan Wu, and Philip He. A Bayesian Basket Trial Design Using Local Power Prior. *Submitted to Biometrical Journal*, December 2023. https://arxiv.org/abs/2312.15352

[4] J. Kazdan, R. Shen, and H. Sun. Collage Operators in Diffusion for Dimension Changes. *Manuscript in Progress*, Dec 2023.

[5] R. Shen, L. Tian, L.Zhu, and V. Devanarayan, et al. Nonparametric ANCOVA for Longitudinal Outcomes in a Randomized Clinical Trial. *Manuscript in Progress*, Dec 2023.

PATENTS/AWARDS

- US Patent 054642/536653: Graph Convolutional Anomaly Detection — Aug 2019
- The Firestone Medal (Top 10% Stanford Undergrad Honors Theses) — May 2022
- Phi Beta Kappa Society (Top 10% Stanford Undergrads Selected by Faculty Committee) — May 2022
- Siemens National Competition Semi-Finalist — Oct 2017

COMPUTER SKILLS

COMPUTER LANGUAGES (IN ORDER OF PROFICIENCY)
Python, R, C++, Java, C, Swift

RELEVANT COURSEWORK

Machine Learning, Generative Modeling, Data-Mining and Analysis, Statistical Methodology/Theory, Probability Theory, Real Analysis, Linear Algebra & Multivariate Calculus,

TEACHING EXPERIENCES

2022 - 2023: Teaching Assistant for **STATS 202: Data Mining and Analysis**, **STATS 248: Causal Inference in Clinical Trials and Observational Study (II)**, **STATS 101: Data Science 101**
2023 - 2024: Teaching Assistant for **STATS 300A: Theory of Statistics I** (First Year Statistics Ph.D. Course), **STATS 216: Statistical Learning**

Funmi Solano

(817) 726-1554 | fas22@stanford.edu | www.linkedin.com/in/funmisolano

EDUCATION

Stanford University, Stanford, CA

- BS, Biomedical Computation — Sep. 2018 – Sep. 2023
- Prospective MS, Biomedical Informatics — Jan. 2022 – Jun. 2024

DATA SCIENCE EXPERIENCE

Cardinal Free Clinics, Redwood City, CA

Computation + Data + Tech (CDT) Chair — Feb. 2022 – Feb. 2023

Epic Task Force Member — Sep. 2023 – present

- Pulled and analyzed patient data from clinic EMR to create 8 reports for clinic operations and SSRFC conference submissions; gained better understanding of SQL logic for database queries
- Created a guide on how to navigate the EMR and create reports to help train future CDT chairs
- Facilitate clinic's transition from old EMR to Epic; create written guides, video walkthroughs, and checklists to help train clinic volunteers on how to use Epic

Edwards Lifesciences, Irvine, CA

Corporate Quality intern — Jun. 2022 – Aug. 2022

- Took ownership of project to migrate data for ~1700 project risk assessments, a type of quality process, from a legacy system to a new electronic form database; worked independently with minimal supervision
- Gained more familiarity with Microsoft Excel, Power Automate, and Visual Basic for Applications
- Detailed instructions on how to navigate legacy system and recommendations for next steps; facilitated continuation of this project via final presentation to team leadership

Alife Health, San Francisco, CA (remote)

Data Science intern — Jun. 2021 – Aug. 2021

- Worked on the Follicle Forecaster project of the Ovarian Stimulation product, which uses follicle sizes and estradiol levels to predict the best trigger day for an *in vitro* fertilization (IVF) cycle
- Gained greater command of the Numpy, Pandas, and Pytorch libraries in Python; gained in-depth knowledge of the IVF landscape and medical device regulation
- Contributed to the following publication:
  Fanton, M., Nutting, V., **Solano, F.** et al. (2022). An interpretable machine learning model for predicting the optimal day of trigger during ovarian stimulation. Fertility and sterility, 118(1), 101–108. https://doi.org/10.1016/j.fertnstert.2022.04.003

MENTORSHIP EXPERIENCE

Inclusive Mentoring in Data Science, Stanford, CA

Teaching assistant — Jan. 2023 – Mar. 2023

- Served as primary point of contact for data science mentorship program with 20 Stanford graduate mentors and 20 non-Stanford mentees from underrepresented backgrounds
- Hosted weekly 1-hour mentoring sessions with mentee over topics such as common programming languages, differences between data job roles, and introduction to machine learning
- Recruited panelists for a virtual round table of data science professionals at end of program

Sophomore College, Stanford, CA

Sophomore College Assistant, Needs Finding in Healthcare — May 2021 – Sep. 2021

- Co-assisted a class of 12 students and 6 faculty for a 3-week summer course sponsored by the Stanford Byers Center for Biodesign; quickly adapted to the evolving needs of a brand-new course offering
- Organized transportation and supplies for 3 on-campus and 3 off-campus social events using a $2,000 budget
- Supported a student with mobility accommodations, such as by learning how to drive a golf cart

RELEVANT SKILLS

- **Technical:** Microsoft Office, Google Workspace, Python, R, SQL, Tableau

Gwanggyu Sun

Bioengineering Ph.D. candidate, Stanford University | Phone: 650-391-4414 | E-mail: ggsun@stanford.edu

KEY SKILLS

Computational Modeling, Systems Biology, Bioinformatics, Data Science, Data Visualization, Machine Learning, Python

EDUCATION

Jun 2024 (exp.)	Ph.D. in Bioengineering, Stanford University, CA
2020	M.S. in Bioengineering, Stanford University, CA
2017	B.S. in Chemical and Biological Engineering (summa cum laude), Seoul National University, S. Korea
	□ Double major in Biological Sciences, minor in Computer Science & Engineering

RESEARCH EXPERIENCES

Covert Lab, Graduate Research Assistant, Stanford University, Stanford, CA 2018-present

Work under the supervision of Dr. Markus Covert (principal investigator)

- □ Built a whole-cell computational model of *E. coli* that can simulate the growth of *E. coli* cells under multiple environments and conditions [4, 5]
- □ Integrated various types of genome-scale data into the whole-cell model to expand the model's capabilities, cross-evaluate heterogeneous datasets, and gain novel insights on *E. coli* physiology [1, 2]
- □ Developed data exchange pipelines between the whole-cell model and the EcoCyc database [3]
- □ Built LSTM networks that can emulate the whole-cell model and massively speed up discovery
- □ Managed continuous integration, version control, testing, and overall maintenance of the codebase over high-performance cloud computing environments (Stanford's HPC cluster and Google Cloud)

BridgeBio Pharma, Quantitative Biology Intern, Palo Alto, CA 2021-2022

- □ Developed a Python codebase that can be used to efficiently build arbitrary ODE models and fit their parameters, with easy accessibility for less technical users
- □ Built *in silico* human cell models to simulate and evaluate the performance of candidate drugs for the treatment of erythropoietic protoporphyria (EPP) and KRAS-positive cancer

Nucleic Acid Bioengineering Lab, Research Assistant, Rice University, Houston, TX 2016

Work under the supervision of Dr. David Zhang (principal investigator)

- □ Developed a computational/experimental protocol to enrich specific loci from genomic DNA/RNA samples with rationally designed hybridization probes to increase sequencing efficiency and sensitivity

Molecular Biotechnology & Biomaterials Lab, Undergraduate Research Assistant, Seoul, S. Korea 2015-2017

Work under the supervision of Dr. Byung-Gee Kim (principal investigator)

- □ Developed an algorithm that uses genome-scale metabolic models of microbes to predict which regulatory genes to over-/under-express to maximize chemical yields in engineered microbes [6]

PUBLICATION HIGHLIGHTS

1. **G. Sun**, A. Zhang, M.W. Covert. "A whole-cell computational model highlights evolutionary constraints that shape the features of *E. coli*'s rRNA operons", manuscript in preparation
2. **G. Sun***, M.M. DeFelice* et al., "Cross-evaluation of *E. coli*'s operon structures via a whole-cell model suggests alternative cellular benefits for low- versus high-expressing operons.", in review
3. P.D. Karp et al., "The EcoCyc Database (2023)", *EcoSal Plus* (2023)
4. **G. Sun***, T.A. Ahn-Horst*, M.W. Covert. "The *E. coli* whole-cell modeling project." *EcoSal Plus* (2021)
5. D.N. Macklin et al., "Simultaneous cross-evaluation of heterogeneous E. coli datasets via mechanistic simulation." *Science* (2020)
6. M. Kim, **G. Sun**, D.Y. Lee, and B.G. Kim. "BeReTa: a systematic method for identifying target transcriptional regulators to enhance microbial production of chemicals." *Bioinformatics* (2016) (*Equal contributions)

HONORS & AWARDS

Siebel Scholars Award, Class of 2024	2023-2024
Kwanjeong Educational Foundation Overseas Scholarship	2017-2022
Star Mentor Award, Stanford Bio-X Undergraduate Summer Research Program	2022, 2023

Nikolai G. Vetr

Contact Information
Phone: (602) 578-9196
Email: nikgvetr@stanford.edu
LinkedIn: linkedin.com/in/nikolai-vetr
GitHub: github.com/NikVetr/

Education

Postdoc, Montgomery Lab, Stanford University — *Current*
Pathology + Genetics + Biomedical Data Science

PhD, University of California, Davis — *2020*
Dissertation: *Exploring and Extending Multivariate Brownian Diffusion Models of Phenotypic Evolution for Bayesian Phylogenetic Inference*
Anthropology + Population Biology + Data Science & Informatics

BA, Vanderbilt University — *2013*
Earth & Environmental Sciences + Ecology, Evolution & Organismal Biology
Departmental Honors, *summa cum laude*

Recent Work

Vetr, N., Gay, N., and Montgomery, S. 2023. *The impact of exercise on gene regulation in association with complex trait genetics.* Conditionally accepted to Nature Communications.

Abell, N., **Vetr, N.***, Montgomery, S., et al. 2023. *A Survey of High Depth Allele-Specific Expression Across Normal Tissues and Ovarian Cancers.* In Prep.

MoTrPAC Study Group†. 2023. *Temporal dynamics of the multi-omic response to endurance exercise training across tissues.* Accepted to Nature.

**dual first authorship*, †*Author Group: 2 (of 8)*

Leadership

Founder, Applied Bayesian Statistics Research Cluster — *2019 - 2020*
President, Board of Directors, *Wild Animal Initiative* — *2020-Present*
President, Board of Directors, *Rethink Priorities* — *2023-Present*

Languages

Programming: R, Stan, BASH, Python, C++, CSS, HTML, JS
Natural: Russian, English, Spanish

Teaching

Associate Instructor, University of California, Davis — *2015 - 2020*
Human Evolution + Primate Evolution + Human Evolutionary Biology
Carpentries Instructor, Data & Software Carpentries — *2019*
Course Coordinator, Workshop in Applied Phylogenetics — *2019*

Selected Grants & Awards

NIH T15 — *2021*
Excellence in Data Science Community Training and Outreach — *2019, 2020*
Outstanding Graduate Student Teaching Award Nominee — *2016, 2019, 2020*
1st Place Picnic Day Exhibit Award in "Secrets of Nature" Category — *2017*
NSF Graduate Research Fellowship — *2015*

Service

Journal Review: *Evolution* (2017), *Science Communications* (2018), *Cell Reports* (2021), *Human Genetics and Genomics Advances* (2022)
Grant Review: *WAI Grants* (2021, 2022, 2023)

Skills & Interests

- Generalized Linear Models
- Multiomic Data Integration
- Time Series Modeling
- Bayesian Methods
- Monte Carlo Methods
- Causal Inference
- Computer Vision
- Artifical Neural Networks
- Data Visualization
- Science Communication

Juan Manuel Zambrano Chaves

jmz@stanford.edu (650) 334-8729 jmzam.github.io
JMZambranoC JMZAM

EDUCATION

Stanford University
PhD Biomedical Data Science 06/2020 - 06/2024
MS Biomedical Informatics 2018 - 2020
Universidad de los Andes (Bogota)
MD (Summa Cum Laude) 2010 - 2017
BS Biomedical Engineering (Summa Cum Laude) 2009 - 2013

PROFESSIONAL EXPERIENCE

Google Research (Mountain View, CA) - Research Intern 10/2023 - Present
- Main contributor to a project leveraging large language models in the medical realm, at a team at the intersection of Health AI and Google DeepMind.

Microsoft Research (Redmond, WA) – Research Intern 06/2023 - 09/2023
- Led a project developing large multimodal models as a research intern in Health Futures. Publication in progress.

GRADUATE RESEARCH EXPERIENCE

Stanford University – Department of Biomedical Data Science
- Research foundation models in medicine: develop methods for training, evaluation, and applications.
- Developed methods to benchmark language models, including large language models, in the radiology domain.
- Developed models that predict cardiovascular risk using features from 8,000+ abdominal tomography images and electronic health record data, outperforming best existing models by 58% F1 score.
- Developed tissue saliency, a feature attribution aggregation method that enables dataset-wide explainability of computer vision models offering quantitative explainability across tissues and samples.
- Created population health applications with our computer vision models by measuring biomarkers in already-acquired images of 17,000+ patients; discovered phenome-wide associations of skeletal muscle biomarkers with hundreds of disease phenotypes, and identified opportunity for 1,000-fold increase in low muscle mass diagnosis (SABI 2022, RSNA 2022; manuscripts under review).
- Co-developed methods to perform simultaneous prediction of diseases with shared pathophysiology as a multi-task problem, demonstrating benefit in label efficiency with this approach (MICCAI 2022).

SELECTED PUBLICATIONS

[**Zc**BADRLLC]. **RaLEs: A Benchmark for Radiology Language Evaluations.** NeurIPS 2023. [view paper]

[**Zc**CWDBBMRSJRP]. **Opportunistic Assessment of Ischemic Heart Disease Risk Using Abdominopelvic Computed Tomography and Medical Record Data: a Multimodal Explainable Artificial Intelligence Approach.** Scientific Reports, 2023. [view paper]

[CBDVsP**Zc**APLC]. **RoentGen: Vision-Language Foundation Model for Chest X-ray Generation.** arXiv 2022. [view paper]

[BG**Zc**MSRRPWBC] **Opportunistic Incidence Prediction of Multiple Chronic Diseases from Abdominal CT Imaging Using Multi-task Learning.** MICCAI 2022. [view paper]

[DSVECD**Z**CL]. **ViLMedic: a framework for research at the intersection of vision and language in medical AI.** ACL 2022. [view paper]

SKILLS

Foundation Models	Medical Imaging	Python/Pytorch/JAX
Natural Language Processing	Tabular EHR Data	Disease Prediction
Computer Vision	Deep Learning	Opportunistic Imaging

Stanford MEDICINE | Department of Biomedical Data Science

PART TIME APPLICANTS

This section contains resumes of MS students, PhD candidates, and postdocs who are actively searching for part-time employment or internship opportunities

January 23rd 2024 | Stanford University | Stanford Alumni Center

Matthew Aguirre

Stanford University
Biomedical Informatics Ph.D Candidate

Email: magu@stanford.edu
Website: maguirre1.github.io

Education

Stanford University
Ph.D, Biomedical Informatics, 2024 (planned).
M.S., Statistics, 2023 (concurrent with Ph.D).

Harvard College
A.B. *cum laude*, Applied Mathematics, 2017.

Experience

Department of Biomedical Data Science, Stanford University
Biomedical Informatics Ph.D Candidate, Pritchard lab (2019–)
- Thesis: Network models for the genetic architecture of complex traits.
  Focus areas: Transcriptional networks, dynamical systems, graphical models.
- Internship: Guardant Health (Summer 2023).
  Focus areas: Cancer biology, cell-free DNA, methylomics, survival analysis.
- Additional projects in functional genomics and population genetics.
  Focus areas: Meta-analysis, machine learning, simulation-based inference.

Deptartment of Pediatrics, Stanford University
Research Data Analyst, Priest and Rivas Labs (2017–19)
- Studies and methods development for genetic epidemiology at biobank scale.
  Focus areas: Association studies, structural variation, risk modeling.

Teaching and Service

Department of Biomedical Data Science, Stanford University
- DBDS Student Representative, Fall 2022–23.
- Teaching Assistant, BMI 217: Translational Bioinformatics, Winter 2021–22.
- Grader, BMI 214: Representations and Algorithms for Computational Molecular Biology, Winter 2021–22.
- Mentor, BIODS 360: Inclusive Mentoring in Data Science, Winter 2021–23.

Awards

- Teaching award, Stanford Department of Biomedical Data Science (2022).
- Microsoft Research PhD Fellowship (2021).
- Reviewers' choice abstract (top 10% of posters), American Society of Human Genetics (ASHG) Annual Meeting (2020).

Selected Manuscripts

Matthew Aguirre, Jan Sokol, Guhan Venkataraman, Alexander Ioannidis, "A deep learning classifier for local ancestry inference." *arXiv*, 2020.

Matthew Aguirre, Yosuke Tanigawa, Guhan Venkataraman, Rob Tibshirani, Trevor Hastie, Manuel Rivas, "Polygenic risk modeling with latent trait-related genetic components." *EJHG*, 2021.

Matthew Aguirre, Manuel Rivas, James Priest, "Phenome-wide burden of copy-number variation in the UK biobank." *AJHG*, 2019.

Skills

Programming: Python, R, Unix, bash, SLURM/HPC, MATLAB, LaTeX.
Tools: Jupyter, git, conda, torch, tensorflow, networkx, GATK.

Susanna Avagyan

Research Assistant and Graduate Student in Stanford University

I am currently pursuing my MS in Biomedical Data Science at Stanford University, while also working as a Research Assistant in the Newman Lab at Stanford's Institute for Stem Cell Biology and Regenerative Medicine. Through research and industry experiences, I have developed a strong passion for complex disease research from a multi-omics perspective with the use of ML/DL methods and translational application to precision medicine.

savagyan@stanford.edu

6505469317

Stanford, United States

linkedin.com/in/susanna-avagyan-051b2a181

EDUCATION

MS in Biomedical Informatics
Stanford University

09/2023 - Present *Stanford, USA*

BS in Data Science / Bioinformatics Track
American University of Armenia

08/2018 - 06/2022 *Yerevan, Armenia*

WORK EXPERIENCE

Research Assistant
Newman Lab, Institute of Stem Cell Biology and Regenerative Medicine

06/2023 - Present *Stanford, USA*

Responsibilities

- Contributing to ML/DL-driven research on tumors and tumor microenvironments.
- Assisting in the development of computational methods for studying genomic and transcriptomic data.
- Collaborating with the team on processing and analyzing bulk, single-cell, and spatial genomic data.

Contact : Dr. Aaron Newman / Principal Investigator, Newman Lab - amnewman@stanford.edu

Data Scientist
Vivan Therapeutics

01/2020 - 09/2023 *London, UK (remote, contracted locally by Armenian Bioinformatics Institute)*

Responsibilities

- Developing an AI tool for personalized cancer therapeutics by analyzing patient genomic data.
- Assisting the lab team in explorative and statistical analysis of experiments.
- Participating in strategic planning, goal-setting, and mentoring new hires within the data science team.

Contact : Dr. Nahuel Villegas / Chief Scientist Officer, Vivan Therapeutics - nahuel@mypersonaltherapeutics.com

SKILLS

Python/R | SQL (Oracle, SQL Server) | Leadership | Statistical Analysis | Command Line Tools for Genomics | Computational Biology Algorithms | Public Speaking | Organizational and Communication Skills

RESEARCH PROJECTS

Methodology to predict absolute developmental potential from scRNA-seq data (06/2023 - Present)
- **Research Assistant**

Single-cell analysis of human healthy and diseased skin focused on stromal cells (10/2022 - Present)
- **Junior Group Research Lead**

Temporal changes of gene expression in health and mental disorders (05/2022 - 08/2023)
- **Second Author / Submitted to Nature Schizophrenia**

Subtyping of Cancer using ML/DL methods based on Multi-Omics and Clinical Outcome (01/2022 - 09/2023)
- **Primary Researcher / Published Editorial**

ORGANIZATIONS

Armenian Bioinformatics Institute (09/2021 - Present)
Junior Group Research Lead / Mentor

AWARDS

Best Bachelor's Degree Student Award (09/2021)
An award given to the best students in the field of IT at the State Educational Awards

LANGUAGES

English
Full Professional Proficiency

Armenian
Native or Bilingual Proficiency

INTERESTS

Precision Medicine | Bio GenAI | E-health

Joseph Boen

EDUCATION

Stanford University, Ph.D. Biomedical Data Science 2023 – present

University of Oxford, M.Sc. by Research (with Distinction) in Oncology 2022 – 2023

Johns Hopkins University, B.S. (Honors) Applied Mathematics & Statistics, B.S. (Honors) Biomedical Engineering 2018 - 2022

SELECTED HONORS & FELLOWHIPS

2022, Oxford University Clarendon Scholarship *full tuition & fees, 1 award/course*
2022, The Queen's College Oxford Graduate Fellowship *research funding, 2 awards/year*
2022, National Science Foundation Graduate Research Fellowship (declined for Oxford) *~$150k award, 12 % acceptance rate*
2022, Dept. of Energy Computational Science Graduate Fellowship Finalist (withdrawn for Oxford) *~$150k award, ~20 awards/year*
2018, Jennifer and Seymour Baron Scholarship, Johns Hopkins University *merit scholarship, 1 award/class*

SELECTED RESEARCH EXPERIENCE

Stanford University, Department of Biomedical Data Science *Prof. Barbara Engelhardt* September 2023 – present

- Scaling gaussian processes for analyzing high-resolution spatial transcriptomics.

University of Oxford, Mathematical Institute *Prof. Helen Byrne* 2022 – 2023

- Optimal transport methods for integrating and analyzing highly multiplexed spatial proteomics. Distinguished thesis.

Johns Hopkins University, School of Medicine *Prof. Robert Stevens* 2021 – 2022

- Machine learning for event forecasting in high frequency ICU data. Presentation at ICCAI '23, under review at IEEE JBHI.

Johns Hopkins University, Department of Biomedical Engineering *Prof. Jeremias Sulam* 2020, 2022

- Kernel methods for continuous and early time series classification. Publication in Advanced Science.
- Memory efficient computer vision models for representation learning of high-resolution digital pathology data.

SELECTED INDUSTRY & GOVERNMENT EXPERIENCE

United States Department of Defense, *Applied Research Mathematician* Summer 2022

- Classified technical research. Authored internal publications, contributed to programs, and briefed leadership and partners.

Novartis Institutes for Biomedical Research, *Bioinformatics Researcher* Summer 2021

- Computational analysis of single cell omics data for drug target identification. Authored internal publication.

Los Alamos National Lab, *Computational Scientist* Summer 2021 – Spring 2022

- Quantum chemistry simulations on quantum computers. Conference presentation at APS '22, publication in JCTC.

Johns Hopkins University Applied Physics Lab, *Research Software Engineer* Summer 2019 – Summer 2020

- Error mitigation and compiler optimization for quantum computers. Conference presentation at APS '21.

SELECTED PUBLICATIONS

P. Schleich, **J. Boen** et al. "Partitioning Quantum Chemistry Simulations with Clifford Circuits" *J. of Chem. Theory and Comp. 2023*

X. Jiang, **J. Boen** et. al. "Accurate Prediction of Antimicrobial Susceptibility for Point-of-Care Testing of Urine in Less than 90 Minutes via iPRISM Cassettes" *Advanced Science 2023*

S. Nair, **J. Boen** et. al. "A Real-Time Deep Learning Approach for Inferring Intracranial Pressure from Routinely Measured Extracranial Waveforms in the Intensive Care Unit" Under Review, *IEEE J. of Biomedical and Health Informatics 2023.*

SELECTED COURSEWORK & SKILLS

Computer	Python, C, R, Linux, version control, high performance computing
Machine Learning	CV, NLP, generative models with *PyTorch,* data science with *sklearn/pandas/numpy*
Bioinformatics	Omics analysis with *scanpy*, *seurat*, protein visualization and modeling with *PyRosetta, PyMOL*
Courses (*grad level)	Machine Learning*, Matrix Analysis*, Bioinformatics*, Statistics, Probability, Optimization

Laura Bravo Sánchez

lmbravo@stanford.edu

Google Scholar

Experience

2023, 2024 Graduate Teaching Assistant - Stanford University. Computational Methods for Biomedical Image Analysis.

2019-2021 Research Assistant - Universidad de los Andes. Leader of the surgical scene segmentation and activity recognition project.

2019 Research Scientist - Tecnología y Gerencia S.A.S. Part of the experimentation team in charge of applying Machine Learning techniques to financial data.

2017-2018 Graduate Teaching Assistant - Universidad de los Andes. Biomedical Image Analysis.

2016-2017 Undergraduate Teaching Assistant - Universidad de los Andes. Scientific Programming and Biomedical Image Analysis.

Other research

MSc. Research
Creation of the FLC dataset for fine-grained object localization tasks.
Totæ Lacrimæ: automatic recognition of human emotions based on tear crystal micrographs.

Additional Info

2020 - Volunteer at Visible Hands Corporation mentoring women in the Innovation Girls 4.0 initiative.
2018 - Mathematics teacher for adults at ColombiaCrece
2017 - Member of the sports climbing team at Universidad de los Andes
2015 - Volunteer at Visible Hands Corporation.
2012- 2014 - Volunteer at Techo Colombia building houses and raising funds.
Other - I enjoy climbing and embroidery

Education

2021 - **PhD. Candidate, Biomedical Data Science** — Stanford University
Advisor: Serena Yeung

2017 - 2019 **MSc., Biomedical Engineering** — Universidad de los Andes
Advisor: Pablo Arbeláez. Thesis: Language-Guided Instrument Segmentation for Robot-Assisted Surgery

2013 - 2017 **BSc., Biomedical Engineering** — Universidad de los Andes
Minor in French Culture and Language
Thesis: Development of an electrical stimulation device to reduce immobilization muscle atrophy.
Advisors: Mario Valderrama and Juan Cruz.

Research

2022 - Present **Current Research** — Stanford Medical AI and Computer Vision Lab
Human Mesh Recovery and Pose Estimation in video for understanding child development.

2017 - 2021 **Past Research** — Biomedical Computer Vision Group
Scene Parsing and Action Recognition in surgical videos. Improving efficiency in COVID-19 test usage with machine learning.

Publications and Awards

Z. Weng, **L. Bravo-Sánchez**, et al. HARMONI: Artificial Intelligence-Powered 3D Analysis of Video-Based Caregiver–Child Interactions. In preparation.

Z. Weng, **L. Bravo-Sánchez**, S. Yeung-Lévy, Diffusion-HPC: Generating Synthetic Images with Realistic Humans. Accepted at International Conference on 3D Vision 2024.

M. Escobar, G. Jeanneret, **L. Bravo-Sánchez**, et al., 'Smart pooling: AI-powered COVID-19 informative group testing'. Sci Rep 12, 6519, 2022.

C. González*, **L. Bravo-Sánchez***, and Pablo Arbelaez. Surgical instrument grounding for robot-assisted interventions, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 10:3, 299-307, 2022

Fulbright Colombia Minciencias Scholarship recipient 2021 Cohort.

A. Huaulmé, et al., 'MIcro-surgical anastomose workflow recognition challenge report', Computer Methods and Programs in Biomedicine, τ. 212, σ. 106452, 2021.

C. González*, **L. Bravo-Sánchez***, and Pablo Arbelaez, 'ISINet: An Instance-Based Approach for Surgical Instrument Segmentation'. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Lecture Notes in Computer Science(), vol 12263. Springer, Cham.

Leader of Team Uniandes. We won first place in the Activity Recognition task at MISAW challenge (MICCAI 2020) and 5 awards at the Robust Endoscopic Instrument Segmentation Challenge (MICCAI 2019).

T. Ross et al., "Comparative validation of multi-instance instrument segmentation in endoscopy: Results of the ROBUST-MIS 2019 challenge," Medical Image Analysis, vol. 70. Elsevier BV, p. 101920, May 2021.

L. Bravo*, A. Pardo*, G. Perez*, P. Arbeláez. Finding Four-Leaf Clovers: A Benchmark for Fine-Grained Object Localization in the Sixth Workshop on Fine-Grained Visual Categorization (FGVC6), CVPR 2019.

Bryan J Bunning

bbunning@stanford.edu; 847-638-8245; Sunnyvale, CA

EDUCATION

Stanford Medicine, Department of Biomedical Data Science **Stanford, CA**
Ph.D. Candidate, Biomedical Data Science *Sep '21 – Present (exp. Grad '25)*

- GPA: (3.87/4)
- Dissertation (in progress): Developing methods, products, and tools in translational science. Goal to understand and improve operations of human subjects' research at Stanford, focused primarily on Electronic Health Record integration and data capture. Evaluation of AI models within Stanford Hospital.
- Topic Areas: Clinical Trials, FDA regulation in AI and clinical trials, evaluation of health AI, biostatistics, Real World Data, wearables, Electronic Health Record integration and health data interoperability
- Teaching Assistant of Data Science for Medicine, a graduate course led by Stanford Health's Chief Data Scientist on types of health data and how to utilize it for machine learning (2022 & 2023)

Columbia University **New York, NY**

- Master of Science in Biostatistics: GPA: (3.88/4) *Class of '21*
- **GRE**: Analytical Writing (5.5/6, 98%), Quantitative (169/170, 94%), Verbal (159/170, 82%)
- Thesis covering adaptive clinical trial design of the COVID-19 vaccines and other therapies

University of Chicago **Chicago, IL**

- Bachelor of Science in Biological Sciences, Immunology Specialization *Class of '16*

EXPERIENCE

Stanford University Quantitative Sciences Unit **Stanford, CA**
Biostatistician/Data Scientist *Jun '20 – Jul '21*

- Designed a clinical trial, wrote its protocol, received required approvals, and performed statistical analyses
- Corresponded with the US FDA to address statistical comments on a novel COVID-19 clinical trial design
- Assisted in writing a pre-Emergency Use Authorization (EUA) for a COVID-19 diagnostic device
- Worked with doctors, nurses, and IT professionals to create and maintain a clinical trial database

TLPInvest **Libertyville, IL**
Investment Committee Member | External Consultant *Jun '19 – Present*

- Member of family office investment committee and serve as an outside consultant to numerous early and growth stage portfolio companies and affiliated investors
- Evaluate decks, clinical trial outcomes, and regulatory risk across healthcare, medical device, & biotech

Sean N. Parker Center for Allergy Research at Stanford University **Stanford, CA**
Life Science Research Professional | Clinical Research Coordinator *Jun '16 – Jul '19*

- Broad responsibilities including immunology wet lab experimentation (CyTOF, flow cytometry, LC-MS), dry lab computational analyses and modeling, and patient facing clinical visits
- Created recruitment strategy for pediatric twins doubling cohort size, resulting in high impact publication
- Coinventor on patent "Microfluidic device and diagnostic methods for allergy testing …"

SKILLS

- Data science, data visualization, machine learning, AI, statistical analyses
- R (tidyverse, statistical/ML packages), SQL, EHRs (OMOP-CDM, FHIR, EPIC), MS Office
- Study design, clinical operations, clinical decision making, diagnostic statistical properties
- Analyzing trial results, various -omics, cytometry, oncology animal models, budgets

EXTRACURRICULARS

Swimming, cycling, hiking, triathlons, competitive video games, food allergy patient advocacy
University of Chicago Varsity Swim Team (2012-2016)

- NCAA All-American, previous 6x school and 1x conference record holder

August Burton

13074 Summerharvest Drive, Draper, UT 84020 | 801-696-9586 | aburton6@stanford.edu

EDUCATION

Stanford University **Stanford, CA**
Biomedical Computation, Simulations Track *June 2024*
GPA: 4.027
MCAT: 523

RESEARCH EXPERIENCE

Cremer Lab of Quantitative Microbial Physiology **Stanford, CA**
Research Assistant and Stanford BioX Undergraduate Fellow *2022-Present*

- Mathematically model toxic peptide digestion by the gut microbiome along the large intestine
- Quantitatively investigate carcinogenic hydrogen sulfide production by the gut microbiome
- Develop a bioinformatics workflow to predict the extracellular protein digestion capability of gut bacteria

LEADERSHIP AND SERVICE

Stanford Pickleball Club **Stanford, CA**
President and Founder *2022-Present*

- Lead the Stanford Pickleball Club with over 450 current members
- Provide weekly open play sessions for all levels, with an emphasis on accessibility for beginners
- Run training sessions for the club's competitive team

Stanford Health Care Volunteer **Stanford, CA**
Inpatient Unit Volunteer *2023-Present*

- Volunteer four hours weekly in an inpatient unit at Stanford Hospital
- Attend to patient needs and assist at the nurse's station

Heart and Mind Thought Group **Stanford, CA**
Co-chair and Founder *2023-Present*

- Lead biweekly discussions and writing workshops for LDS Stanford students, taking an intellectually rigorous approach to modern spirituality, thought, and culture within the LDS faith tradition
- Collaborate with leading LDS scholars to infuse our discussions and Stanford LDS culture with responsibly researched religious insights and critiques

Church of Jesus Christ of Latter-Day Saints **McAllen, TX**
Full-Time Representative *December 2018 – July 2020*

- Collaborated with the BYU Record Linking Lab to bring a free, cutting-edge genealogy service to residents of South Texas
- Designed the database that collects, curates, and distributes all data for the BYU collaboration
- Provided community service 12 hours a day

EMPLOYMENT

Sophomore College Assistant **Stanford, CA**
Teaching Assistant and Resident Assistant *Summer 2022*

- Assisted staff in planning and implementing the experiential curriculum for a three-week immersive *Needs Finding in Healthcare* Stanford Biodesign course
- Guided students through early stages of the Stanford Biodesign process

Jacob (Jake) Chang

jachang4@stanford.com | 808-381-1535

EDUCATION

Stanford University, Palo Alto, CA **September 2022 – Present**
Ph.D. Student
Biomedical Informatics

University of Notre Dame, Notre Dame, IN **August 2016 – May 2020**
Bachelor of Arts
Majors: Statistics, Sociology
Minor: Glynn Family Honors Program

RESEARCH INTERESTS

I am interested in developing computational and statistical methods geared toward understanding the spatial and temporal processes of the tumor microenvironment. I am particularly interested in spatial statistics, high-dimensional statistics, time series, and probabilistic approaches to machine learning.

RESEARCH / EMPLOYMENT

Plevritis Lab, Palo Alto, CA **April 2023 – Present**
Ph.D. Student

- Build computational methods to understand the spatial and temporal dynamics of the tumor microenvironment
- Developed a statistical framework to detect differential spatial cell-type colocalization patterns between primary tumors and metastatic disease

Genentech, South San Francisco, CA **May 2022 – August 2022**
Molecular Oncology / Oncology Bioinformatics Intern

- Analyzed bulk-RNA seq data to understand transcript changes in colorectal cancer cell lines after downregulating the Wnt pathway
- Utilized single-cell reference data and bulk-RNA deconvolution methodologies to estimate specific cell types within bulk samples

84.51˚, Chicago, IL **June 2020 – May 2022**
Data Scientist, Insights & Statistical Learning

- Consulted and advised on issues regarding algorithmic fairness and statistical bias to promote best data practices
- Generated insights on pharmaceutical operations, healthcare access, and nutritional literacy for Kroger customers

Institute for Pure and Applied Mathematics (IPAM), Los Angeles, CA **Summer 2019**
UCLA Undergraduate Researcher

- Leveraged support vector regression and neural network ensembles within genetic algorithms to minimize dilution of precision in Walker constellations while working on summer project for The Aerospace Corporation
- Responsible for experimental design, data manipulation, and analysis to improve experimental efficiency and obtain robust statistical results

Lab for Big Data Methodology, Notre Dame, IN **January 2019 – May 2020**
Undergraduate Researcher

- Performed text-mining, data-cleansing, structural topic modeling, and sentiment analysis using R
- Built random forest classifier using text comments from ratemyprofessor.com to predict professor ratings

LEADERSHIP & OUTREACH

GEM Fellowship

- GEM fellow sponsored by Stanford University and Genentech.

Biomedical Data Science Peer-to-Peer Mentor, Palo Alto, CA

- Participant and organizer of a mentorship program to advise and assist prospective Biomedical Data Science students from underrepresented backgrounds with their graduate school application

SKILLS & PROFICIENCIES

Computer: Python, R, SQL, Spark
Language: Spanish (Full Professional Capacity)

RELEVANT COURSEWORK

Artificial Neural Networks, Data Mining, Mathematical Statistics, Probability, Abstract Mathematics, Real Analysis, Spatio-Temporal Statistics, Statistical Computing, Statistical Methods, Algorithms in Computational Biology, Data Science in Medicine, Machine Learning for Neural Data Analysis, Modern Statistical Methods, Biostatistics, Probabilistic Models

Wenyuan Sandy Chen

cwyuan1010@gmail.com Linkedin

EDUCATION

Master in Biomedical Data Science
Stanford University
09/2023 – 06/2025

- Research Interest: Primary focus lies in the application of machinelearning and AI models to advance the precision medicine.
- Key areas of interests: Multimodal AI and Large Language Models, Computer Vision Algorithms, Predictive Modeling for Patient Outcomes, etc.

BS in Data Science and Human Biology
University of California San Diego
09/2018 – 03/2023

- Overall GPA: 3.97 (Magna Cum Laude)

ORGANIZATIONS

Heal in Pocket (Cali Registered Non-Profit Org)
SDE & Fundraising Chair
05/2023 – present | San Diego, United States
https://healinpocketusa.wixsite.com/heal-in-pocket

- **Start-Up "Heal in Pocket"**: Initiated based on volunteer experience at a homeless medical outreach program. The aim is to deliver a free EHR and telehealth mobile app to make healthcare accessible for underserved communities.
- **Strategic Partnerships:** Open Collective Foundation for fiscal, local non-profit, Street Corner Care
- **Technical Contribution:** Crafted UI/UX using React Native. Constructed the API and backend server through MongoDB, Express, and Node.js.
- **Leadership Role:** Directed team meetings, coordinated project efforts, conducted code reviews, executed testing protocols, and facilitated the onboarding of new team members.

LANGUAGES

Mandarin	● ● ● ● ●
Cantonese	● ● ● ● ●
English	● ● ● ● ○
French	● ● ○ ○ ○
Python	● ● ● ○ ○

SKILLS

- Python, Java, Javascript, Solidity, SQL
- PyTorch and TensorFlow
- Machine Learning Model: Decision Tree, Regression, KNN, PCA, RNN, CNN, LSTM, Transformer, etc.
- Large-Scale Data Processing: Apache Spark and Amazon Web Services (AWS)

PROFESSIONAL EXPERIENCE

Research Assistant
Prof. Rose Yu's Lab in Department of Computer Science and Engineering & Halıcıoğlu Data Science Institute of UCSD
07/2022 – present | San Diego, United States

- **Collaboration with Abiomed**: Developed a data-driven simulator for the Impella device.
- **Modeling Mean Arterial Pressure (MAP)**: Predicted MAP based on motor speed using device data.
- **Sim2Real Integration**: Utilized Domain-Adversarial Training of Neural Networks (DANN) with a probabilistic neural process forecaster in PyTorch.
- **Simulation Verification**: Successfully replicated and confirmed results for Acute Myocardial Infarctions (AMICGS) and High Risk Percutaneous Coronary Interventions (HRPCI) cohorts.
- **Model Performance Comparison**: Evaluated differences between the conditional LMU encoder and DANN model by adjusting motor speed.

Research Assistant
Prof. Enfu Hui's Lab in Department of Cell and Developmental Biology of UCSD
02/2019 – 05/2020 | San Diego, United States

- **CD80 Glycosylation Study**: Examined the impact of N and O glycosylation on CD80 binding levels with CTLA-4. Investigated CTLA-4's function as an immune checkpoint and its suppressive regulation of T cells.
- **Mutant Design with Software:** Utilized MegAlign for DNA sequencing alignment and A plasmid Editor (ApE) to design CD80 mutants with diverse glycosylation sites.
- **Cell Culturing**: Grew desired mutants across various cell types and introduced the protein PD-L1 to identify glycosylation sites influenced by the PD-L1/CD80 interaction.

AWARDS

Outstanding Capstone Project Award
Halıcıoğlu Data Science Institute of UCSD
09/06/2023

2023 University BlockChain Contest 2nd Place
Franklin Templeton
02/06/2023
Received a prize of 7k

Ya-Chi **Chu**

ycchu97@stanford.edu | yachichu | Ya-Chi Chu

"Ya-Chi is a PhD candidate at Stanford University studying applied math and looking for opportunities to apply her skills in machine learning (ML) research roles."

Experience

Stanford University Management Science & Engineering Department (Prof. Madeleine Udell) — *Stanford, CA, U.S.*

Research Assistant — *Jan. 2023 - PRESENT*

- Designed faster and accurate randomized algorithm for eigenvalue problem with applications in large-scale ML algorithms.
- Accelerated the linear system solver inside the interior-point methods with applications in ML models.

Stanford University Mathematics Department — *Stanford, CA, U.S.*

Administrative Teaching Assistant – Linear Algebra and Differential Calculus of Several Variables — *Spring 2023*

- Assisted with running the course, including office hours, exam grading duties, and other administrative responsibilities.
- Communicated with part-time students and students who need academic accommodation.
- Supervised the graders (e.g., weekly keeping track of the grading progress, re-distributed incomplete tasks) and reported status to instructors.

Stanford University Mathematics Department — *Stanford, CA, U.S.*

Course Assistant – Applied Matrix Theory — *Spring 2022, Winter 2022*

- Hosted office hours and content review sessions, and graded homework and exams.

Skills

Highlighted Domain fields: Machine Learning (ML), Reinforcement Learning (RL), Optimization (theory/algorithms), Statistics, Mathematical Modelling, Computational Mathematics

Programming Languages: Python (PyTorch, Scikit-Learn etc.), C++, R, Matlab, Julia, LaTeX

Education

Stanford University — *Stanford, CA, U.S.*

Ph.D. in Mathematics - GPA: 4.025/4.00 — *Sep. 2021 - PRESENT*

- Supervisor: Prof. Madeleine Udell and Prof. Lexing Ying
- **Research direction:** Accelerate optimization algorithms via randomization techniques with applications to ML on large-scale datasets
- **Selective Courseworks:** Machine Learning (ML), Reinforcement Learning (RL), NLP, Convex Optimization, Data Analysis

National Cheng Kung University — *Tainan, Taiwan*

M.S. in Applied Mathematics - GPA: 4.30/4.30 — *Sep. 2019 - Jun. 2021*

- Supervisor: Prof. Ruey-Lin Sheu
- **Thesis:** "On Separation Properties of Quadratic Level/Sublevel Sets and Its Applications"

National Taiwan University — *Taipei, Taiwan*

B.S. in Mathematics - GPA: 4.16/4.30 — *Sep. 2015 - Jun. 2019*

- Graduated with Dean's Award

Publications

- Nguyen, H. Q., **Chu, Y. C.**, & Sheu, R. L. (2021). On the convexity for the range set of two quadratic functions. *Journal of Industrial and Management Optimization*, 18(1), 575-592. https://doi.org/10.3934/jimo.2020169
- Nguyen, H. Q., **Chu, Y. C.**, & Sheu, R. L. (2023). Separating disconnected quadratic level sets by other quadratic level sets. *Journal of Global Optimization*, 1-27. https://doi.org/10.1007/s10898-023-01330-8

Honors & Awards

2021 **The Sunseri Fellowship,** Stanford University — *Stanford, CA, U.S.*

2014 **1st Place (with US$ 20,000 scholarship),** Shing-Tung Yau High School Mathematics Award — *Taipei, Taiwan*

SASKIA COMESS

(360) 318-3588 | saskiaco@stanford.edu | github.com/saskialynn

EDUCATION

***PhD Environment & Resources,* Stanford University** 9/2020 - present

- **PhD Minor in Statistics**: Completed, December 2022
- Qualifying Exams: Passed, November 2022
- Advisors: Susan Holmes (Statistics) and Gary Shaw (Epidemiology)
- Fields of Concentration: Statistics and Epidemiology

***M.A. Statistics,* Yale University Graduate School of Arts and Sciences** 8/2019-5/2020

***M.P.H. Environmental Health,* Yale School of Public Health** 8/2017-5/2019

***B.A. Science, Technology & Society; Minor: Mathematics,* Vassar College** 9/2013-5/2017

- Cum Lauda, Phi Beta Kappa Award, Departmental Honors, General Honors

AWARDS, FELLOWSHIPS, GRANTS

Stanford Data Science Scholar, Stanford University 9/2022-6/2024

Stanford Impact Labs Summer Collaborative Research Fellowship, Stanford University 6/2022

Stanford Regulation, Evaluation, and Governance Summer Institute Fellowship, Stanford Law School 6/2022

HB and JS Nicholas Fellowship for Graduate Study, Vassar College 5/2018, 5/2019

Stolwijk Fellowship Award; Climate Change & Health Initiative Grant, Yale University 4/2018

Solomon Fellow in Health Law and Policy, Yale Law School 6/2019-5/2020

Phi Beta Kappa Prize, Vassar College 5/2017

- Awarded to the student with "the most distinguished academic record of the graduating class."

Stamps Foundation Scholarship [awarded but not accepted], University of Washington 4/2013

PUBLICATIONS & PRESENTATIONS

1. **Comess, S.;** Chang, H.; and Warren, J. "A Bayesian framework for incorporating exposure uncertainty into health analyses with application to air pollution and stillbirth." *Biostatistics.* doi.org/10.1093/biostatistics/kxac034
2. **Comess, S.;** Wang, H.; Holmes, S.; and Donnat, C. "Statistical Modeling for Practical Pooled Testing During the COVID-19 Pandemic." *Statistical Science.* 37 (2) 229 - 250, May 2022. doi.org/10.1214/22-STS857
    a. Interactive R-Shiny Application: homecovidtests.shinyapps.io/Group-testing/
3. **Comess, S**.; Donovan, G.; Gatziolis, D.; and Deziel, N. "Exposure to Atmospheric Metals Using Moss Bio-Indicators and Neonatal Health Outcomes in Portland, Oregon." *Environmental Pollution.* Vol. 284, Sept. 2021. doi.org/10.1016/j.envpol.2021.117343
4. **Comess, S**.; Akbay, A.; Vasiliou, M.; Hines, R.; Joppa, L.; Vasiliou, V.; Kleinstreuer, N. "Bringing Big Data to Bear in Environmental Public Health: Challenges and Recommendations." *Frontiers in Artificial Intelligence.* Vol. 3, May 2020. doi.org/10.3389/frai.2020.00031
5. Donovan, G.; Gatziolis, D.; Jakstis, K., and **Comess, S**. "The natural environment and birth outcomes: comparing 3D exposure metrics derived from LiDAR to 2D metrics based on the normalized difference vegetation index." *Health and Place.* Vol. 57, May 2019, Pg. 305-312 doi.org/10.1016/j.healthplace.2019.05.011
6. *Invited Speaker*: Stanford Women in Data Science, Health and Environment Data Panel. 26 April 2023.
7. *Poster Presentations*
    a. "A Bayesian framework for incorporating exposure uncertainty into health analyses with application to air pollution and stillbirth." *Society for Epidemiology Research*, June 2022, Chicago.
    b. "Exposure to Atmospheric Metals Using Moss Bio-Indicators and Neonatal Health Outcomes," *International Society of Environmental Epidemiology*, Aug 2019, Utrecht, The Netherlands.
    c. "Exposure to Atmospheric Metals Using Moss Bio-Indicators and Neonatal Health Outcomes," *Planetary Health Alliance Annual Meeting*, Sept 2019, Stanford University.

OANA M. ENACHE
650-799-5918 | oenache@stanford.edu | www.oanaenache.com | www.linkedin.com/in/oana-enache

EDUCATION

STANFORD UNIVERSITY, Department of Biomedical Data Science, School of Medicine, Stanford, CA
Doctor of Philosophy, expected May 2025. GPA: 3.96/4.0.

- **Relevant technical coursework:** Deep Learning, Decision Making Under Uncertainty, Data Science for Medicine, Deploying & Evaluating Fair AI in Healthcare, Design and Conduct of Clinical Trials, Social Epidemiology, Causal Inference

DUKE UNIVERSITY, Department of Biostatistics & Bioinformatics, School of Medicine, Durham, NC
Master of Biostatistics, May 2021. GPA: 3.9/4.0. Tuition scholarship.

- **Thesis:** Predicting premature ventricular contractions and understanding provider knowledge using digital health tools

UNIVERSITY OF CALIFORNIA, BERKELEY, Department of Mathematics, College of Letters and Science, Berkeley, CA
***Bachelor of Arts*, Applied Mathematics, Concentration: Computational Biology**, December 2014.

EXPERIENCE

STANFORD UNIVERSITY, Stanford, CA 2022-present
***PhD Candidate,* Health Policy Data Science Lab (Advisor: Dr. Sherri Rose)**

- Developing novel method detecting fraudulent insurer gaming of payment models for millions of Medicare beneficiaries
- Developing new end-to-end auditing framework to help federal agencies better evaluate payment-related artificial intelligence and machine learning models impacting Medicare Advantage beneficiaries
- Designed an original intersectional reporting approach for more transparent reporting of patient identities in clinical research

DUKE UNIVERSITY, Durham, NC 2020-2021
***Masters Student Researcher,* Big Ideas Lab (Advisor: Dr. Jessilyn Dunn)**

- Developed machine learning model to predict premature ventricular contractions using smart watch and electrocardiogram data
- Designed survey to assess cardiologists' understanding of wearable device features relevant to patient care and device regulation

BROAD INSTITUTE OF MIT AND HARVARD, Boston, MA 2015-2019
***Associate Computational Biologist II (2018-2019),* Golub Lab (Advisor: Dr. Todd Golub)**

- Discovered and published novel off-target effect of gene editing Cas9 protein by using gene expression data of 165 cancer cell lines
- Consulted with diverse teams from several pharmaceutical companies on experiment planning and analysis for drug discovery projects

***Associate Computational Biologist I (2015-2017),* Golub Lab**

- Developed, packaged, and released an open-source Python package providing tools for processing and analysis of GCT and GCTx-formatted data (cmapPy, available through pip and Bioconda) and led team of 3 developers in making similar R, Java, Matlab packages
- Developed a quantitative metric to identify genotype-specific effects in isogenic cell lines in collaboration with several oncologists

PATENTS

- Ben-David U, Golub T, Beroukhim R, **Enache OM**, Rendo V. 2019. DNA damage response signature guided rational design of CRISPR-based systems and therapies. U.S. Provisional Patent Application #62/909,131. *Provisional Patent*, filed October 2, 2019.
- Griffith O, **Enache OM**, Pepin F, Spellman P, Gray J. 2013. Gene expression panel for breast cancer prognosis. CA:2869313:A1. *Patent*, filed April 5, 2013, and issued October 10, 2013.

SELECT FIRST AUTHOR PUBLICATIONS

A link to my complete publication list can be found on my website or my LinkedIn profile.

- (In review) **Enache OM**, Goldman Rosas L, Rose S. "Current Clinical Research Reporting Paradigms May Misrepresent Participant Identities."
- **Enache OM**, Rendo V, Abdusamad M, Lam D, Davison D, Pal S, et al. "Cas9 activates the p53 pathway and selects for p53-inactivating mutations". Nature Genetics. 2020; 52: 662–668.
- **Enache OM**, Lahr DL, Natoli TE, Litichevskiy L, Wadden D, Flynn C, et al. "The GCTx format and cmap{Py, R, M, J} packages: resources for optimized storage and integrated traversal of annotated dense matrices". Bioinformatics. 2018. doi:10.1093

ADDITIONAL INFORMATION

Teaching

- Teaching Assistant for Data Science for Medicine, Stanford University School of Medicine: Fall 2022
- Teaching Assistant for Introduction to Statistical Programming, Duke University School of Medicine: Fall 2020

Justice, Equity, Diversity, and Inclusion (JEDI) work

- Student representative on admissions committee, Department of Biomedical Data Science, Stanford University: 2023-present
- Department-wide student JEDI representative, Department of Biomedical Data Science, Stanford University: 2022-2023
- Analyst, Inaugural report on compensation, promotion and hiring practices for institute employees by gender and race/ethnicity, Broad Institute of MIT and Harvard: 2018
- Founding member and officer, Women in Mathematics, UC Berkeley: 2013-2015

LESLIE ANASU ESPINOZA CAMPOMANES

459 Lagunita Drive, Stanford, CA 94305
(650) 714-7440 | leslieec@stanford.edu

Innovating bio/med technology and education with a planetary perspective for social good and development

EDUCATION

STANFORD UNIVERSITY
B.S Biomedical Engineering and Bioengineering Stanford, CA 2019 - 2023
B.A Human Rights

STANFORD SCHOOL OF MEDICINE
Candidate for Master of Biomedical Informatics Stanford, CA 2022 - Present

- GPA: 3.85/4.0
- Vice Provost Undergraduate Education STEM fellow for outstanding academic performance and healthcare equity
- Recipient of BioDesign fellowship evaluating innovation gaps in urologic diseases focused in prostate cancer at Stanford Hospital
- Basser fellow designing eco-friendly and sustainable ventilation systems in high-altitude Andean villages
- Student Advisory Board for Stanford Thinking Matters Program evaluating required introductory programs impact
- Research Assistantship advised by Prof. Camarillo, working on energy absorption efficacy for concussion prevention in sports
- Stanford Seeds of Change Leader advising STEM opportunities for minority high school girls in East Palo Alto and Redwood City
- Coursework: ML-AI, Bio-design, Bioprinting, Optimization Methods, Wet lab, Electrical, Mechanical & Biological Models
- Scientific Papers co-authorship:
  - Wang DC, Wu Y, Santos-Valencia F, Espinoza-Campomanes LA, Franks KM, Luo L. "Developmental and experience-dependent separation of hypothalamic circuits regulating hunger and thirst." Manuscript in preparation.
  - Yuzhe Liu; Ramanand V. Vegesna; Xianghao Zhan; Weiguang Yang; Leslie Anasu Espinoza-Campomanes; Gerald A. Grant; David B. Camarillo "An adaptive and wearable hydraulic shock absorber with fully efficient energy dissipation"

EXPERIENCE

Stanford School of Engineering – Bioengineering and Biomedical Engineering Department Stanford, CA
Teaching Assistant Sept. to Present

1. ***BioDesign Senior Capstone (BIOE 141A/B)***
2. ***Biomedical System Prototyping Lab (BIOE 123)***

- Mentor 54 Stanford senior students majoring in Bioengineering throughout the development of their major capstone projects
- Co-developed the first capstone course syllabus with generative AI considerations at Stanford Bioengineering
- Accompany and troubleshoot 56 junior/senior students building blood centrifuges (electronics, CAD, laser cut, programming)
- Facilitate networks within and outside the Bioengineering department at Stanford across faculty, companies and entrepreneurs
- Evaluate, provide feedback, and support students' and team's performance and their collaborative work
- Run logistics of the course, items ordering, assignments development and grading, office hours, and interactive events

Tangible Science Redwood City, CA
BioDesign Intern Jul. to Sept. 2023

- BioDesign process and research for a new market area in the optics and eye contact care field
- Strategized sales and marketing campaigns current and incoming products of the company

LivaNova San Francisco, CA
Business Development Intern Jun. to Aug. 2022

- Critical research, identification, and selection of new markets in cerebrovascular diseases experience around the United States
- Implement commercial and research-driven strategies to boost startup evaluations for the Neuromodulation divisio

Google Mountain View, CA
Google Latinx Scholar at the Student Leadership Summit Mar. to May 2022

- Google Student Leadership Summit 2022 gathers top students in technology from across the United States and Puerto Rico for a month of conferences. Selected over 900+ applicants to Google informative sessions on the use of AI in health and education
- Conducted a final speech for the 2022 conference cohort about our learnings and goals as a Latinx community in tech spaces

Hospitales Madrid and CINAC Neuroscience Center ***(Pioneer Parkinson's' Disease research center in Europe)*** Madrid, Spain
Clinical Intern Mar. to Jun 2022

- Selected as a clinical rotation student in CINAC, one of the top neuroscience research centers for Parkinson's Disease in Europe
- Accompanied and supported Parkinson's patients' consultations and non-surgical High-Intensity Focused Ultrasound procedure
- Participated in and conducted clinical studies and bioengineering research on Parkinson's non-invasive procedures

Neurobiology Luo Lab Stanford University Stanford, CA
Neuroscience Research intern Jan. 2020 to Mar. 2022

- Researched prenatal and neonatal neuronal circuits and evaluated brain activation through gene expression
- Developed bio-computational programs that segment neuronal activation signaling in nuclei
- Used DeepLabCut and Ilastik to analyze neonatal mice behavior and neuronal axonal projections respectively

ADDITIONAL INFORMATION

- **Technical Skills:** R, Python, C++, C, Matlab, SQL, UX, Arduino, CAD
- **Languages:** Spanish (Native Speaker), English (Fluent), Quechua – Peruvian Native Language (Intermediate)
- Leader-Member of the Peruvian National Science Club developing heat-warning sensor backpack anti pickpocketing
- Microsoft Peru speaker advocating for digital tools in Peruvian rural villages which turned into a project by Ministry of Education
- Fascinated about personalized medicine, Top Fighter (Peruvian Martial Arts Fitness program) trainer, District Volleyball player

Paula Gablenz

paula.gablenz@gmail.com

Education

Ph.D. Student in Statistics, September 2020 - June 2025 (expected)
Stanford University

M.S. in Statistics, September 2018 - June 2020
Stanford University

B.Sc. in Economics, October 2013 - May 2016
Ludwig-Maximilians-University (LMU) Munich

Research Experience

Research Assistant, September 2018 - September 2019
Graduate School of Business, Stanford University

Pre-Doctoral Research Fellow, September 2016 - August 2018
Stanford Institute for Economic Policy Research (SIEPR),
Stanford University

Research Assistant, July 2017 - January 2018
Psychophysiology Laboratory, Stanford University

Research Assistant, March 2016 - August 2016
The World Bank Group, Development Impact Evaluation Unit

Student Research Assistant, September 2014 - January 2016
Chair of Economic Theory, Ludwig-Maximilians-University Munich

Teaching Experience

Data Science 101 / Stats 101, Instructor, Summer 2022 and Spring 2023
Data Science 101 / Stats 101, TA, Spring 2022
Introduction to Applied Statistics (Stats 191), TA, Winter 2022
Data Mining and Analysis (Stats 202), TA, Fall 2020 and 2021
Introduction to Regression Models (Stats 203), TA, Winter 2020
Machine Learning and Causal Inference, Grader, Spring 2018

Professional Experience

Data Science Intern, June 2023 - September 2023
QuantCo

Data Analyst, September 2013 - August 2014
Munich Re, Agriculture and Weather Risk Unit

Summer Intern, July 2013 - September 2013
Munich Re, Casualty Reinsurance

Scholarships

Ric Weiland Graduate Fellowship, Stanford University: 2023 - 2025
German National Academic Foundation: 2015 - 2020
Honors Program, Ludwig-Maximilians-University Munich: 2014 - 2016

EBRU HOSGUR

ehosgur@stanford.edu linkedin.com/in/havinhosgur/ (650) 832-3333

Education

STANFORD UNIVERSITY **Stanford, CA**
Academic Master of Science in Biomedical Data Science *June 2025*

- **Artificial Intelligence:**
  Artificial Intelligence: Principles and Techniques, Applied Machine Learning, Natural Language Processing with Deep Learning, Computer Vision: Foundations and Applications, Data Driven Medicine
- **Computer Science:**
  Data Management and Data Systems, Mining Massive Datasets, Design and Analysis of Algorithms, Probability for Computer Scientists, Programming Abstractions

STANFORD UNIVERSITY **Stanford, CA**
Master of Science in Epidemiology and Clinical Research *June 2024*

UNIVERSITY OF CALIFORNIA, LOS ANGELES **Los Angeles, CA**
Bachelor of Science in Molecular, Cell, and Developmental Biology

Professional Experience

Computer Science Department, Stanford University **Stanford, CA**
Course Assistant *June 2023 - present*

- Directed and coordinated sections and office hours for two core Computer Science classes, Design and Analysis of Algorithms and Probability for Computer Scientists, impacting 900+ students across multiple quarters
- Designed assignments, lecture materials, and exams, fostering algorithmic problem-solving skills

Biomedical Data Science Department, Stanford University **Palo Alto, CA**
Graduate Research Assistant *July 2023 - present*

- Conduct AI-based analysis of clinical data for Urinary Tract Infection (UTI) detection
- Created a labeled dataset using a manually curated electronic phenotyping rules for gradient-boosted tree models
- Ongoing project: Implement NLP techniques to mine Electronic Health Records for UTI symptoms

Inspirit AI **San Carlos, CA**
Instructor *August 2021 - June 2023*

- Facilitated a dynamic learning environment that fostered a deep understanding of AI fundamentals in an AI-intensive high school program for 100+ students
- Mentored and empowered students to harness AI technology for social change, assisting them in translating their knowledge into practical, real-world applications that benefit communities and society

Projects

Predicting Airbnb Listing Prices

- Developed a tool using linear regression, multi-class logistic regression, and neural network models to help Airbnb hosts list properties at a competitive price

Querying, Visualizing, and Predicting Healthcare Pricing

- Trained a logistic regression prediction model to predict billing provider costs for cataract surgeries to recommend a more equitable surgery cost
- Analyzed public CMS Medicare data to better understand the characteristics of surgery pricing

Exploring Popular GitHub Repositories

- Wrote SQL queries to create plots to answer questions about the quality of highly watched GitHub repositories in regards to languages used, average file size, and average commit message length

Technical Skills

Languages: Python, SQL, R, C++
Libraries: Pandas, NumPy, PyTorch, TensorFlow
Tools: VS Code, LaTeX, Linux, Vim

JIANG YIXING

Phone: 470 4023247 Email: jyx@stanford.edu

EDUCATION

Stanford University **Jun 2022 – Jun 2027**

- PhD student under Biomedical Data Science program, advised by Prof. Andrew Ng (Stanford Machine Learning Group) and Prof. Jonathan Chen (HealthRex Lab)
- Relevant coursework: Machine Learning, Artificial Intelligence, Deep Learning, Meta Learning, Natural Language Understanding, Natural Language Processing, Clinical Informatics, Computational Biology, Statistical Inference, Data Mining, Time Series Analysis

National University of Singapore (NUS) **Aug 2017 – Jun 2021**

- Bachelor of Biomedical Engineering (Research-focused Pathway) with Minor in Computer Science

AWARDS

- ***PhD Fellowship***: National Science Scholarship (5-year funding from Singapore) **Fall 2022 – Spring 2027**
- Undergraduate Honours (Highest Distinction), **Valedictorian**

WORKING EXPERIENCE

Visiting PhD Student, Rajpurkar Lab, Harvard Medical School **Jul 2023 – Sep 2023**

- Developed multi-agent framework for evaluating large language models in dermatology

Research Engineer (AI), Institute of High Performance Computing, A*STAR **Jul 2021 – Jun 2022**

- Developed generative models for OCT penetration depth enhancement
- Developed geometric deep learning algorithms on retinal OCT volumes for referral triaging

PUBLICATION AND SUBMISSION

- **Jiang, Y.**, Irvin, J., Ng, A. Y., & Zou, J. (2023). VetLLM: Large Language Model for Predicting Diagnosis from Veterinary Notes. In *Proceedings of the Pacific Symposium on Biocomputing (PSB) 2024* [In Press]
- **Jiang, Y.**, Lee, A., Ni, X., Corbin, C. K., Irvin, J., Ng, A. Y., & Chen, J. C. (2023). Probabilistic Prediction of Laboratory Test Information Yield. In *Proceedings of the AMIA 2023 Annual Symposium*
- Soh, Z. D.[*], **Jiang, Y.**[*], Rajamaren, S., Nongiur, M., Majithia, S., Tham, Y. C., Rim, T. H., Qian, C., Victor, K., Aung, T., Wong, T. Y., Xu, X., Liu, Y., & Cheng, C. (2023). From 2 dimensions to 3rd dimension: Quantitative prediction of anterior chamber depth from anterior segment photographs via deep-learning. In *PLOS Digital Health*
- **Jiang, Y.**, Alford, K., Ketchum, F., Tong, L., & Wang, D. M. (2020, September). TLSurv: Integrating Multi-Omics Data by Multi-Stage Transfer Learning for Cancer Survival Prediction. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB' 20)*.
- Qian, C., **Jiang, Y**., Soh, Z.D., Ganesan, S. S., Xiao, S., Tham, Y. C., Xu, X., Liu, Y., Zhong, H., & Cheng, C. (2022). Smartphone-acquired anterior segment images for deep learning prediction of anterior chamber depth: a proof-of-concept study. In *Frontiers in Medicine*

TEACHING AND LEADERSHIP EXPERIENCE

Student Lead, Stanford Machine Learning Group **Jan 2023 – Present**

- Mentored eight master's and undergraduate students and led four projects (one ongoing) in the AI for healthcare bootcamp, resulting in two top-tier publications
- Established extensive collaborations with faculty members and technical IT staff at the School of Medicine

Teaching Assistant, Stanford University **Sep 2023 – Present**

- Teaching Data Science for Medicine (BIOMEDIN 215)

Teaching Assistant, School of Computing NUS **Jan 2019 – May 2021**

- Taught database systems, data structures and algorithms, programming methodology and software engineering

SKILLS AND CERTIFICATIONS

- **Programming**: Python, Java, MATLAB and C programming. More than 13 years of programming experiences
- **Machine learning** and **Data Science**: PyTorch, scikit-learn, SQL, pandas, OpenCV
- **Open-source contributions:** deployr-dev, GPT4-Automation, ngboost, PyTorch

Rohit Khurana

22208 N 36^{th} St · Phoenix, AZ 85050
480-294-9245 rkhurana@stanford.edu

Education

Stanford University **June 2025**
Master of Science, Biomedical Data Science *Stanford, CA*

Vanderbilt University **May 2023**
Bachelor of Science *Nashville, TN*

- *Majors*: Computer Science (Honors), Molecular & Cellular Biology (Highest Honors), Mathematics | *GPA*: 3.993/4.00
- *Thesis*: Leveraging algorithmic toolkits and high-throughput, multiplexed imaging techniques to characterize the tumor microenvironment.

Experience

Mayo Clinic **November 2022 - Present**
Biostats Intern - AI & Informatics *Rochester, MN*

- Currently developing a natural language processing pipeline for sentiment analysis and topic modeling of reviews from leading patient portals (e.g., MyChart).

Undergraduate Researcher (Computational Biology) **August 2022 - May 2023**
The Zhou Lab *Nashville, TN*

- Contributed to the development of a computational filter to refine structural variant calling results in long-read sequencing data.

Undergraduate Researcher (Single-Cell Biology) **January 2020 - May 2023**
The Ihrie Lab *Nashville, TN*

- Utilized cytometric approaches and high-dimensional analyses to computationally characterize subpopulations of cells in glioblastoma (funded by the Data Science Institute Summer Research Fellowship).
- Revised & fine-tuned a computational workflow to facilitate downstream cyclic immunohistochemistry analysis [2].
- Developed a computer vision model to identify human-defined diagnostic cells in whole slide images of cortical tubers, a neuropathological manifestation of tuberous sclerosis complex. Experimented with achieving end-to-end integration with laser-capture microdissection equipment to automate identification, dissection, and deep-sequencing of these cells [1].
- Quantified Fluorogold injection sites & tested an OpenSCAD pipeline to generate 3D-printed mouse head molds [3].

MathWorks **May 2022 - August 2022**
Software Engineering Intern *Natick, MA*

- Developed a real-time resource monitor for MATLAB Online in C++ to display to users metrics concerning their compute session. Completed all associated unit, integration, and system-level tests.
- Prototyped adding additional functionality to worker containers to directly monitor GPU utilization status.

Research Intern (Cancer Biology) **January 2017 - May 2019**
The Mehta Lab *Phoenix, AZ*

- Investigated the HDAC family for druggable, surrogate targets of OLIG2, a protein critical for gliomagenesis.
- Determined that human glioma stem cells (GSCs) are adversely affected by the knockdown of HDAC1 and radiation by facilitating downstream expression of acetylated p53 and cleaved caspase 3.
- Observed how GSCs adjust their molecular behavior in response to ionizing radiation [3].

(Selected) Honors & Awards

Early Career Researcher Symposium Speaker **October 2021**
International TSC & LAM Research Conference *Remote*

- Project: Computational identification of balloon cells in whole slide images of cortical tubers.

(Selected) Posters, Presentations, & Manuscripts

[1] **Khurana, R.**, Brockman, A. A., Mobley, B. C., Ess, K. C., Ihrie, R. A. Computational identification of balloon cells in whole slide images of cortical tubers [abstract-selected talk & poster]. In: International TSC & LAM Research Conference; 2021 October 28-30.

[2] Brockman, A. A., **Khurana, R.**, Bartkowiak, T., Thomas, P. L., Sivagnanam, S., Betts, C. B., Coussens, L. M., Lovly, C. M., Irish, J. M., Ihrie, R. A. Alignment, Segmentation and Neighborhood Analysis in Cyclic Immunohistochemistry Data Using CASSATT. Cytometry Part B: Clinical Cytometry, 104(5), 344–355. Cover article.

[3] Chervonski, E., Brockman, A. A., **Khurana, R.**, Chen, Y., Greenberg, S., Hay, M. S., Luo, Y., Miller, J., Patelis, D., Whitney, S. K., Walker, M., Ihrie, R. A. Creation and validation of 3D-printed head molds for stereotaxic injections of neonatal mouse brains. J. Neurosci. Methods 360, 109255 (2021).

Joshua Emilio Lazaro

E:jelazaro@stanford.edu **P**:(832)-561-1956

Education

Stanford University School of Medicine — **Stanford, CA**
PhD. Student in Biomedical Data Science — September 2023 – Present

The University of Texas at San Antonio (UTSA) — **San Antonio, TX**
Bachelor of Science | Statistics and Data Science — August 2019-May 2023
Concentration: Biology — GPA: 3.95

Research Experience

Medicaid and Medicare Office of Enterprise and Data Analytics — **Washington, DC**
Coding it Forward Fellow — June 2023 – September 2023
Research mentor: Social Science Research Analyst James DelAguila

- Major Project: Drivers of Medicare Advantage Adoption: Modeling Market and Beneficiary Characteristics
- Spearheaded in-depth quantitative analyses to identify key factors influencing Medicare Advantage adoption, contributing to data-driven decision-making
- Employed programming languages, such as SAS and python to manipulate, analyze, and implement Generalized Linear Models with notable predictive accuracy, further laying the groundwork for enhanced forecasting techniques in future initiative.

Undergraduate Researcher at The New Haven Free Clinic — **New Haven, CT**
Yale School of Medicine and The New Haven Free Clinic — June 2022 – September 2022
Research mentor: Yazhini Ramesh, Yale MPH Candidate

- Major Project: Linear models highlight decreased costs and challenges of inpatient care for undocumented patients in New Haven.
- Utilized SAS and R to process and analyze 3k patient observations spanning 4-5 years, underscoring the significant "value-added" contributions of the HAVEN Free Clinic to the Yale New Haven Hospital
- Formulated linear and logistic regression analyses complemented by medical reports and Electronic Health Records comparing patient blood pressure trends over various treatments and appointment

Stanford Summer Research Program Genetics Scholar — **Stanford, CA**
Stanford Summer Research Program School of Medicine — June 2021 – August 2021
Laboratory of Nima Aghaeepour, Ph.D., mentored by Ph.D. candidate, Camilo Espinosa

- Major Project: Differential disease trajectories between children in low- and middle-income countries through multivariate machine learning models
- Analyzed over 7k proteomic, metabolomic, and clinical features through R programming to study patient outcomes between urban and rural children in low middle income countries.
- Independently harmonized, integrated, and cleaned collinearities in heterogeneous data to adapt, gather, and examine, for machine learning algorithms
- Constructed multivariate machine learning algorithms such as XGBoost to calculate probabilities of survival and found biological differences in urban and rural environments

Leadership Experience

Director of Outreach — **San Antonio, TX**
Code Quantum - "The First Minority Gender Hackathon in San Antonio" — March 2021 – March 2023

- Recruited over 108 students to participate in San Antonio's first marginalized gender annual hackathon.
- Established key sponsorships with industry giants such as Google, JPMorgan, and Valero, ensuring robust funding for workshops and a comprehensive career fair.
- Cultivated partnerships between universities in San Antonio to promote institutional diversity
- Actively mentored participants, providing them with continuous guidance during the 24-hour coding challenge.

President — **San Antonio, TX**
Association for Computing and Machinery Womens/Minority Chapter (ACM-WM) — March 2021 - April 2022

- Empowered over 300 members (117 non-CS majors) in the field of Computer Science by conducting bi-weekly meetings with underrepresented speakers in various sectors of the field
- Recruited sponsors such as MatrixAI, HEB and Paycom to sponsor events and workshops
- Founded an introductory Python workshop for members with no experience in CS
- Collaborated with organizations around UTSA to empower students outside of CS via workshops

Sophia Katherine Longo

sklongo@stanford.edu · +1-(650)-704-2141

Stanford University, expected graduation: BS in June 2024, MS in June 2025 · GPA: 3.804
B.S. Honors Candidate Human Biology, concentration: Cancer Biology and Therapeutic Intervention
Co-Terminal M.S. Candidate Biomedical Data Science
Full Resume: https://www.linkedin.com/in/sophia-longo/

Personal Statement: As an aspiring physician-scientist, I am deeply inspired by people and biology. My mid- and long-term vision is to harness my passion for science to contribute to patient care, developing new approaches for disease prevention and treatment, and healthcare accessibility. My goal is to gain experience working directly with people facing health issues and to make a scientific contribution to our understanding of human disease. I have a strong wet lab research skill set and am aiming to further develop my computational biology skill set.

RESEARCH EXPERIENCE

Researcher, Khavari Epithelial Biology Lab — **Stanford, CA**
Genome regulation in stem cell differentiation and cancer, PI: Dr. Paul Khavari — *June 2017 - Present*

- Current project (honors thesis): establishing which disease-associated, non-protein-coding single-nucleotide polymorphisms (SNPs) identified through Genome Wide Association Studies (GWAS) dysregulate Major Histocompatibility Complex (MHC) surface expression to drive pathogenesis of cancer and autoimmune diseases; wet lab techniques: Massively-Parallel Reporter Assay (MPRA), CRISPR Interference screen, Next Gen RNA-sequencing, polymerase chain reaction (PCR), flow cytometry; dry lab: bedtools, motifbreakR; scraping data through GTEx and GWAS APIs; co-lead with Dr. Jordan Meyers (post-doc)
- First-author publication in *Nature Reviews Genetics* titled "Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics" (Longo, Guo, Ji, & Khavari, 2021)
- Partnered with lab member to build interactive "Human Skin Browser" (using Shiny - RStudio) of single-cell RNA-seq and spatial transcriptomic data of human skin (Thrane et al., 2023)

Researcher, Stanford Skin Innovation & Interventional Research Group (SIIRG) — **Stanford, CA**
Total body skin imaging for atypical nevus syndrome, Lead: Dr. Albert Chiou — *August 2021- Present*

- Data collection from a novel total body skin imaging device developed at Stanford for tracking of moles for atypical nevus syndrome patients with goal of optimizing clinical implementation of device to optimize removal of potentially cancerous lesions while minimizing excessive biopsies
- Created data analysis pipeline synthesizing skin imaging device data, EHR data, patient survey data to determine optimal clinical implementation of device; computational analysis using Python and R; project is currently in phase of data validation

PUBLICATIONS

Longo, S. K., Guo, M. G., Ji, A. L., & Khavari, P. A. (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nature Reviews Genetics*, *22*(October). https://doi.org/10.1038/s41576-021-00370-8

Thrane, K., Winge, M. C. G., Wang, H., Chen, L., Guo, M. G., Andersson, A., … Ji, A. L. (2023). Single-Cell and Spatial Transcriptomic Analysis of Human Skin Delineates Intercellular Communication and Pathogenic Cells. *Journal of Investigative Dermatology*, *143*(11), 2177-2192.e13. https://doi.org/10.1016/j.jid.2023.02.040

CLINICAL EXPERIENCE & COMMUNITY ENGAGEMENT

COVID-19 Case Investigator & Contact Tracer — **Santa Clara County Dept. of Public Health**
Supervisor: Linda Momota Gentry (408)-242-8153 — *September 2020 – June 2021*

- Case Investigation: cold call people who have tested positive for COVID-19 to check-in health status, provide self-isolation guidance, elicit contacts exposed during their contagious period; file forms in the event of worksite, school, or travel spread
- Contact Tracing: call all contacts elicited from COVID-19 cases investigation to provide quarantine guidance and resources and to help connect them with proper testing
- Made calls in English and Spanish (without translator); 1/3 calls in Spanish

Oncology Clinical Practicum, HM Hospitales — **Madrid, Spain**
Doctor: Dra. Irene Moreno Candilejo (principal investigator) — *October & November 2022*

Founding Co-Director, Menlo-Atherton High School Service Learning Center — **Atherton, CA**
Faculty Advisor: Andrew Stuart (astuart@seq.org) — *August 2017 – June 2019*

Stanford Club Tennis: Co-Vice President (2022-2023), Practice Coordinator (2021-2022)

Conor Messer

Palo Alto, CA
conorsmesser@gmail.com | linkedin.com/in/conormesser

Education

Master of Science in Biomedical Informatics 2023 – 2025
Stanford University Palo Alto, CA

Bachelor of Science in Bioengineering, with honors, *summa cum laude* 2014 – 2019
Northeastern University Boston, MA
Minors: Computer Science and Vocal Performance

Fellowships

Knight-Hennessy Scholar 2023 – 2025
Admitted to multidisciplinary and multicultural leadership program for top graduate students Stanford University

Fulbright Student Research Fellowship 2019 – 2020
Awarded by U.S. Department of State to perform research and promote cultural exchange Abu Dhabi, UAE

Distinguished Bioengineering Scholar Award Spring 2018
Awarded to top two juniors in the Bioengineering department based on academic achievementNortheastern University

National Science Foundation Research Experience for Undergraduates Summer 2015
Granted to perform research at University of Colorado Boulder, CO

Research Experience

Associate Computational Biologist II Sep 2020 – July 2023
Broad Institute of MIT and Harvard, PI: Gad Getz Cambridge, MA

- Utilize statistical methods to analyze genomic and transcriptomic data in multiple cancer drug resistance projects, in collaboration with and funded by IBM
- Develop, debug and maintain state-of-the-art computational tools in Python, MATLAB, R, and C++
- Produce visualizations to inform data exploration, assess quality control and prepare publication-ready figures
- Experience with various NGS technologies, including WES, WGS, and bulk RNASeq and exposure to single cell and proteomic datatypes

Fulbright Scholar August 2019 – June 2020
Khalifa University, PI: Federico Renda Abu Dhabi, United Arab Emirates

- Developed optimization and path planning algorithms in Python for design of concentric tube surgical robots
- Integrated multiple libraries (collision-detection, dynamic KD-tree, optimization) with a novel sampling-based planner and kinematic model to explore design possibilities

Research Engineer Co-op July 2018 – December 2018
Liberating Technologies Inc. Boston, MA

- Designed and tested prototype orthotic device using 3D printing, thermoform plastics, textiles, and hardware

Research Assistant January 2015 – June 2015
Northeastern University, PI: Jeffrey Ruberti Boston, MA

- Compared growth and alignment of 200 image pairs of corneal cells to discover effects of stress on growth
- Analyzed data in MATLAB to produce decisive results and visualizations for publication in *Tissue Engineering*

Selected Publications

Parry E, Leshchiner I, Guieze R, ... **Messer C**, et al (2023). Evolutionary history of transformation from chronic lymphocytic leukemia to Richter syndrome. *Nature Medicine*, 1-12.

Messer C, et al (2022). CTR DaPP: A Python Application for Design and Path Planning of Variable-strain Concentric Tube Robots. *2022 IEEE 5th International Conference on Soft Robotics*.

Renda F, **Messer C**, Rucker C, & Boyer F (2021). A Sliding-rod Variable-strain Model for Concentric Tube Robots. *IEEE Robotics and Automation Letters*, 6(2), 3451-3458.

Zareian R, Susilo M, Paten J, ... **Messer C**, et al. (2016). Human Corneal Fibroblast Pattern Evolution and Matrix Synthesis on Mechanically Biased Substrates. *Tissue engineering. Part A*, 22(19-20), 1204–1217.

Paulina Paiz (She/Her/Hers)

ppaiz@stanford.edu http://pau-paiz.com/ https://github.com/paupaiz +1 267 683 6989

Education

April 2023 - Present
Stanford University - M.S. Biomedical Data Science
Courses in Biostatistics, Mathematical Modeling, Computational Immunology, Machine Learning Approaches for Data Fusion, Linear Dynamics

Aug. 2016 – Aug. 2020
University of Pennsylvania - B.A., *magna cum laude,* Cognitive Science
Concentration: Computation. GPA: 3.77/4.00

Research Experience

UC San Francisco (UCSF)
07/2021 - 06/2023
Post-Baccalaureate Research Opportunity to Promote Equity in Learning (PROPEL) Scholar- Analytical trainee applying statistical and computational methods to develop ArchR, an R package for processing and analyzing single-cell ATAC-seq data. NIH-sponsored program supporting the development of underrepresented minority students to competitively apply and succeed in biomedical science graduate studies. Research Advisor: Ryan Corces, Ph.D.

DeepChem
04/22 - 09/2022
Google Summer of Code Contributor- Strengthened DeepChem's open source codebase and bioinformatics arm by adding utilities for training deep learning models on omics data, enabling loading and sharding. Research Advisor: Bharath Ramsundar, Ph.D.

Loyal Cellular Longevity
01/2021 - 05/2021
Research & Development Intern- Profiled the pharmacokinetic effects and mechanism of action of a drug candidate for the VP of Aging. Research Advisor: Michael LaCroix-Fralish, Ph.D.

Rockefeller University
09/2020 - 01/2021
Research Assistant- Devised image analysis pipeline based on Watershed thresholding to quantify and track SARS-CoV-2 infection patterns on lung and brain organoids.
Research Advisor: M. Zeeshan Ozair, MD, Ph.D.

University of Pennsylvania
04/2019 - 03/2020
Undergraduate Research Assistant- Designed and deployed agent-based model to explore topological measures for identifying and predicting the spread of complex contagions on social networks. Funded by Facebook for in-house content moderation efforts. Research Advisor: Douglas Guilbeault, Ph.D.

Industry Experience

Genentech
06/2023 - 08/2023
Computational Biologist Intern- Developed algorithms to decode cell-cell communication and ligand-receptor activity in the context of breast cancer. Research Advisor: Lyndsay Murrow, Ph.D.

Octant Bio
02/2021 - 05/2021
Data Science Contractor- Optimized luciferase screening pipeline in R for functional characterization of G-Protein Coupled Receptors with deep mutational scanning

Google
05/2019 - 09/2019
BOLD Intern- Built internal tool and conducted data analysis to improve structured interviewing and performance review processes for 1K+ software engineers. Mapped Alphabet's bets and projects at interface of medicine and engineering

Leadership & Awards

07/2022
Helmholtz Zentrum München- Poster presentation at scientific conference on Bioengineering Solutions for Biology and Medicine

06/2022
Lindau Nobel Laureate Meeting- Bayer Pharmaceutics invited speaker for panel on "AI in the Life Sciences" alongside Nobel Prize-winning scientists

09/2021 - 06/2022
Brilliant.org, Author- Pitched and published interactive course to help biologists acquire statistical tools useful for experimental design and hypothesis testing

Technical Skills

Programming: Python, R, Julia, Java, MATLAB, Bash
Tools & Frameworks: PyTorch, Mathematica (Wolfram), Hugging Face, scikit-learn, Git, Docker, Slurm, High-Performance Computing, R Shiny

Adyant Shankar

450 Jane Stanford Way, Stanford, 94305 | 603-320-3578 | adyant@stanford.edu | LinkedIn Profile: https://www.linkedin.com/in/adyant-shankar-b800131b0/

EDUCATION

Stanford University — **Stanford, CA**
Biomedical Informatics (MS), Master of Science — *2025*
Biomedical Computation (BS), Bachelor of Science — *2024*

Relevant Coursework

- CS 161: Design and Analysis of Algorithms
- CS 279: Computational Biology: Structure and Organization of Biomolecules and Cells
- CS 274: Representations and Algorithms for Computational Molecular Biology
- BIOMEDIN 215: Data Science for Medicine
- CS 107: Computer Organization and Systems
- Bio 83: Biochemistry and Molecular Biology
- Bio 82: Genetics
- Bio 86: Cell Biology

RESEARCH EXPERIENCE

Genetics Research at Snyder Lab — **Stanford, CA**
5/2023-Present

- Investigated the biological mechanisms through which diseases exhibit cardiovascular symptoms
- Utilizing proteomics, genomics, metabolomics, as well as wearable data to accomplish research goals
- Developing novel ML computational approaches to diagnose Long-COVID in patients

Cardiovascular Health Research at Broad Institute of MIT and Harvard — **Cambridge, MA**
1/2020-Present

- Spent 100+ hours researching and developing predictive models and tools to understand patient risk of cardiovascular disease
- Engaged in 4 projects working on wearable and genetics data from extensive databases
- Leveraged knowledge of statistical algorithms to extract clinical data and develop models and visualizations

HONORS/ACTIVITIES

Society for Improving Diagnostic Medicine — **Stanford, CA**
10/2023

- Selected to present project on improving efficiencies in diagnosing rare diseases
- Spoke to more than 300 researchers and healthcare providers at international medical conference
- Selected as the youngest speaker at the conference to present research

Biomedical Computation Honors Program — **Stanford, CA**
6/2023

- Selected to write honors thesis for the biomedical computation major
- Proposed a thesis on a novel ML algorithm to predict future onset of Long-COVID for patients
- Leading work to investigate the immune system response to COVID-19 and Chronic Fatigue Syndrome

Skills

- Proficient in Python, R, C++, Unix/Linux systems
- Experience with UKBiobank and Apple Watch data
- Proficient with ML algorithms and visualization
- Experience with Electronic Health Record data
- Data mining and wrangling
- Project Management and Problem-Solving

Jacob (Jake) Silberg

18 6th Avenue, Brooklyn, NY, 11217
415-889-9368 | JSilberg@Stanford.edu | GitHub and projects at JakeSilberg.com

ACADEMIC

Stanford University

Ph.D. student in Biomedical Data Science 2023-present

M.S. student in Data Science, Department of Statistics. GPA: 3.97 2020-2023

Research:

- PhD research rotations with Prof Roxana Daneshjou on explainability in dermatological AI models, and Prof Anshul Kundaje on long-range context models for DNA
- Funded Research Assistant in Prof. Andrew Ng's Stanford ML Group on multi-modal ALS predictions
- Led Highway Removal work in Prof Stefano Ermon's SustainLab, using diffusion models to inpaint satellite imagery
- Graduate Fellow at the Human-Centered Artificial Intelligence Institute (Stanford HAI)

Teaching: Course Assistant for CS 229 (Machine Learning). Held weekly office hours, lead problem-solving discussion sections and lectured in course on evaluation metrics for ML models, Convolutional Neural Nets, and Transformers

Coursework awards: CS 236 - Deep Generative Models, awarded Best Social Impact Final Project

Harvard University 2011-2015

A.B. Magna Cum Laude in Social Studies. Secondary in Global Health and Health Policy. GPA: 3.88

- Senior Thesis: *Too Cool for School: Peer Effects and Calendar Effects on Attendance in a Large Urban School District.* Received Highest Honors; Hoopes Prize (for outstanding undergraduate research)
- Undergraduate English Orator at 2015 Commencement
- 2015 Frederick Sheldon Traveling Fellow: Awarded, based on academic achievement, for politics research in Madrid

WORK EXPERIENCE

InvivoScribe 2022-2023

Machine Learning Engineer

- Led development of deep learning models for cell-level classification based on Set Transformer architecture
- Led proprietary automated gating approach

McKinsey & Company 2014-2020

Chief of Staff to Chairman of the McKinsey Global Institute, James Manyika (2018-2019)

- As first Chief of Staff, co-developed role's responsibilities and structure
- Contributed to MGI and external publications on artificial intelligence, including the 2018 AI Index Report; co-authored MGI paper on AI and bias, adapted into HBR article
- Developed materials and led targeted research for Chairman's events, including with heads of state and heads of major international organizations

Senior Business Analyst (2018-present), Business Analyst (2016-2018), Summer Business Analyst (2014)

- Projects have included advising VPs of leading technology company on comprehensive risk review and CEO of leading technology company on growth strategy, as well as scaling a new product at a leading telecommunications company

U.S. Department of Commerce June-Oct 2016

Press Assistant, Office of Public Affairs

- Wrote speeches and op-eds on behalf of U.S. Secretary of Commerce Penny Pritzker
- Staffed Secretary at U.S.-Africa Business Forum, attended by heads of state including President Obama

Valet: Mobile fundraising startup for charities, schools and politicians 2012-2013

Co-founder / Public Engagement

- Led engagement with clients, including Cory Booker for Senate, facilitating over $2,000,000 in total donations
- Managed implementation at events with development directors

LEADERSHIP

Immediate Gratification Players: Harvard's long-form improvisational comedy troupe 2011-2015

President and performer

- Directed and led rehearsals and performances on campus and around the country
- Managed $15,000 annual budget and relations with donors and alumni

LANGUAGE

Spanish: Fluent (C1 certification from Spanish government's DELE program, 2016)

Timothy Sudijono

390 Jane Stanford Way | 845-275-7649 | timothysudijono@gmail.com

EDUCATION

Stanford University **Palo Alto, CA**
PhD Statistics. GPA: 4.13/4.0 *September 2021 -*
- NSF Graduate Fellow

Research Areas: Causal Inference, Statistical Mechanics

Brown University **Providence, RI**
Sc. B. Applied Mathematics with Honors. GPA: 4.0/4.0 *May 2019*
- Rohn Truell Premium Prize (Top Applied Math Graduate)
- Phi Beta Kappa, Magna Cum Laude, Top 150 Putnam Exam (2015)

SELECTED RESEARCH EXPERIENCE

Inference for Synthetic Controls via Refined Placebo Tests.
Draft Available upon Request (2023). Joint with Lihua Lei.

Fluctuation Bounds in the Restricted Solid-on-Solid Model of Surface Growth.
Submitted (2023). ArXiv:2304.07160.

A Topological Data Analytic Approach for Discovering Biophysical Signatures in Protein Dynamics.
PLOS Computational Biology (2022). Joint work with W. S. Tang et al.

A statistical pipeline for identifying physical features that differentiate classes of 3D shapes.
Annals of Applied Statistics (2021). Joint work with B. Wang et al.

PROFESSIONAL EXPERIENCE

Stanford University **Palo Alto, CA**
Teaching Assistant & Instructor *September 2021 - Present*
- As Instructor: Probability Theory Qualifying Exam Workshop (2023)
- As TA: Theory of Probability I, II, Time Series Analysis.

AQR Capital Management **Greenwich, CT**
Portfolio Implementation Analyst *July 2019 - April 2021*
- Implemented and monitored long-short equity and style premia portfolios. Responsible for launching ESG component of two client portfolios.
- Designed and implemented new portfolio rebalancing scheduler in Python; reduced codebase size by 20% and minimized manual intervention in the system.
- Conducted quantitative research to improve investment process. Topics included denoising factor risk models, transaction cost modeling, and shortfall analysis.

SKILLS
- **Programming Languages:** Python, R

MIN WOO SUN

minwoos@stanford.edu | https://minwoosun.github.io/

Interest / Skills: *High-dimensional Statistics (regularization, curse of dimensionality, multiple testing), Machine Learning (supervised and unsupervised learning), Multi-modal Data Fusion, Game Theory, Programming (Python, R, bash, git, LaTeX)*

EDUCATION

Stanford University *Sept. 2020 - Present*
Biomedical Data Science PhD
Teaching Assistant: BIOMEDIN215 (2022 Data Science for Medicine)
Teaching Assistant: BIOMEDIN217 (2022 Translational Bioinformatics)

University of California, Los Angeles (UCLA) *Sept. 2013 - Dec. 2017*
Statistics B.S. and Economics B.A.

RESEARCH EXPERIENCE

Tibshirani Lab, Stanford University *Sept. 2020 - Present*
PhD, advised by Rob Tibshirani

- Design multimodal convex optimization methods using large-scale molecular data (RNA, DNA methylation, etc.) for bulk cell-type deconvolution
- Develop a distributionally robust unsupervised method using reconstruction loss for identifying rare cell-type from single cell transcriptomics data (collaboration: Jerby Lab)
- Created a nested cross-validation algorithm for computing valid confidence intervals from cross-validation estimates for the Cox model test error (published in Statistics in Medicine)

Wall Lab, Stanford University *Jun. 2017 - Sept. 2020*
ML Researcher, advised by Dennis Wall and Stefano Moretti

- Created game theoretic algorithms using Shapley value as a variant-disease association method and implemented on whole genome sequencing data to identify novel candidate genes for autism.
- Published 3 first author papers in peer-reviewed journals (BMC Bioinformatics, Pacific Symposium on Biocomputing, Biomedical Informatics Insights) on the game theory work

WORK EXPERIENCE

Guardant Health *Jun. 2023 - Sept. 2023*
R&D Bioinformatics Intern

- Implemented statistical models for detecting cell-type using DNA methylation data

Invitae *Jan. 2018 - Aug. 2020*
Data Scientist (full-time)

- Designed and deployed machine learning tools to optimize lab workflows and monitor lab errors.

RECENT PUBLICATIONS

Full list of publications available on Google Scholar

* **Sun, M. W.**, Troxell, D., and Tibshirani, R. Public health factors help explain cross country heterogeneity in excess death during the COVID19 pandemic. Nature Sci Rep 13, 16196 (2023). https://doi.org/10.1038/s41598-023-43407-0

* **Sun, M. W.** and Tibshirani, R. Confidence intervals for the Cox model test error from cross-validation. Statistics in Medicine.

* **Sun, M. W**., Moretti, S., & et al. (2020). Game theoretic centrality: A novel approach to prioritize disease candidate genes by combining biological networks with the Shapley value. BMC Bioinformatics, 21(1), 356.

Gloria Vergara Neyra

gloriavn@stanford.edu | 678-330-3562 |Stanford, CA and Atlanta, GA

EDUCATION

Stanford University | Stanford, CA — B.S. Bioengineering '24, Co-terminal M.S. Biomedical Data Science '25
September 2020 – June 2025 (Expected) | Current GPA: 3.838
Relevant Courses:
Bioinformatics: Representations and Algorithms for Computational Molecular Biology
Computer Science: Intro to MATLAB, Intro to Scientific Python, Programming Methodology (Python), Programming Abstractions (C++)
Bioengineering: Systems Biology, Computational Biology, Physical Biology, Systems Physiology and Design
Math: Linear Algebra, Multivariable Calc & Modern Applications; Multivariable Integral Calc; Differential Equations & Fourier Methods

The Gwinnett School of Math Science and Technology | Lawrenceville, GA — High School Degree
August 2016 – May 2020 | GPA: 3.96 / 4.00 | Honors graduate
QuestBridge National College Match Award to Stanford University | December 2019
President of National German Honor Society | Aug 2019-May 2020
Veteran member of the Science Olympiad Team | Aug 2016-May 2020

EXPERIENCE

Representations & Algorithms for Computational Molecular Biology | Stanford University — *Course Projects*
September 2023 – December 2023
Biomedical informatics course consisting of a series of programming projects to develop an understanding of fundamental algorithms of computational biology. Implemented the following algorithms for the following using Python and its relevant packages, such as sci-kit learn, pandas, and NumPy: DNA sequence alignment, K-Nearest Neighbors classification, Gene Set Enrichment Analysis, Protein Folding.

Stanford Bio-X Undergraduate Summer Research Program | Stanford University — *Research*
June 2023 – September 2023
Selected as one of 70 Bio-X Fellows. Conducted research in the lab of Dr. Rogelio Hernández-López and investigated HER2 expression and tumor heterogeneity in breast cancer cell lines. Engineered T cells to express various SynNotch-CAR genetic circuits to target cancer cells with varying HER2 expression levels. Learned to use Flow Cytometry and an Incucyte. Presented my findings at the Bio-X Symposium to Fellows, faculty, and other Stanford affiliates.

Fundamentals for Engineering Biology Lab | Stanford University — *Course Project*
January 2023 – March 2023
Collaborated with two peers to engineer *E. coli* to express a fluorescent single chain variable fragment (scFv) to potentially target an ovarian cancer antigen. Our team designed the plasmid used for bacterial transformation, induced expression in transformed bacteria, and used an SDS-Page and Western blotting to verify production of the scFv.

Bioengineering Systems Prototyping Lab | Stanford University — *Course Project*
January 2023 – March 2023
Worked with two peers to design, construct, test and improve three versions of a centrifuge prototype. Used CAD to design and 3D print parts. Collaborated with my group to implement a rudimentary PID controller to control centrifuge speed. Enhanced the user interface to allow for user input of speed and run time, and coded sounds to indicate centrifuge run status. Showcased our final prototype to peers, course staff, and other Stanford affiliates.

The Tech Desk at Lathrop Learning Hub | Stanford University — *Tech Desk Consultant, Team Lead*
September 2021 – Present
Help students access free-to-rent technology to aid their learning at Stanford. Manage, troubleshoot, and distribute thousands of pieces of technology to students and staff. Promoted to Team Lead in June 2022. Responsible for hiring, training, and supporting other consultants to create an inclusive work environment. Serve as a go-to person for questions and responsibilities outside the scope of a consultant's knowledge.

LAB SKILLS

Pipetting, microscopy
Gel electrophoresis, PCR
Bacterial cultures, DNA miniprepping
Flow cytometry
Incucyte Live-Cell Analysis

TECHNICAL SKILLS

Python, Jupyter Notebook
MATLAB
C++
FlowJo
SnapGene

LANGUAGES

Spanish – native speaker
German – intermediate

John N Wang

503-800-1607 | johnwang.bio | jwang003@stanford.edu | jwang307 | john-n-wang

Education

Stanford University Sept. 2021 - June 2025

M.S. Biomedical Data Science, B.S. Computer Science

- 4.02 GPA
- Coursework (**Graduate Level**): Linear Algebra, Multivariate Calculus, Computer Systems, Algorithmic Design, **Computational Biomolecular Structure, Algorithms for Molecular Biology, NLP with Deep Learning, Inference and Regression, Deep Learning for Genomics, Graph ML, Systems for ML**

Experience

Data Science Intern | *Python, Docker, Survival Analysis* June 2023 – Present

Virtualitics, Inc.

- Led project on building a generalized predictive maintenance (PMX) module from ground up. Finished module weeks ahead of schedule and implemented novel survival ML pipelines with API and SDK development
- Built statistical methods to impute incomplete data, sample time-series logs, and sample / denoise sensors readings. Improved use cases of module over **10x**
- Reduced time needed to build predictive maintenance solutions from the scale of **months to a few hours**

Drug Discovery Research Intern | *Python, Java, Bash, Linux* June 2022 – June 2023

Dror Lab, Stanford Artificial Intelligance Lab

- Modified computationally predicted protein binding pockets for higher efficacy in drug discovery through binding site refinement with templates identified from local structure search. Adjacent work published in *eLife*, 2023
- Created **Java, Python, and Bash** scripts to automate workflow for searching **50,000** protein database. Improved docking performance in over 50% of protein structures in dataset with new method
- Reduced search and analysis time by over **20** times using parallelization and automated analysis

Bioinformatics Research Intern | *Java, Geneious, BLAST* July 2020 – August 2021

Portland State University

- Built programs to identify unknown genetic markers in 900+ cruciviruses with **Hidden Markov Model Search**
- Developed Java scripts for phylogenetic visualization and workflow integration with **BLAST** and **HMMER**. Discovered and annotated **442** new cruciviral genomes. Automated detection of genes for viral identification
- Research presented at 2022 International Symposium on ssDNA Viruses in Sete, France

Activities and Projects

Pre-training for Few Shot Molecular Property Prediction | *PyTorch, PyG, RDkit* October 2023 – Present

- Exploring pre-training methods for generating molecular representations for general downstream molecular tasks
- Testing denoising 3D structures to learn all-atom embeddings for tasks beyond QM9 (FS-Mol, MoleculeNet, etc.)

LLMs for Computational Biology Analysis | *Python, LLMs, Prompt Engineering* April 2023 – June 2023

- Explored the ability of RLHF-tuned LLMs for learning bioinformatics software and automating ML pipelines
- Prompt engineered automatic workflow that successfully replicated molecular prediction paper results using LLMs
- Developed novel method using LLMs to evaluate long form analysis output from chat-bots using QA eval method

Drug Repurposing with Language Model Embeddings | *Python, PyTorch, AWS* Jan. 2023 – March 2023

- Trained biomedical language models to embed knowledge graph of drug, disease interactions. Predicted drug repurposing targets with graph embeddings. Improved training speed with fine-tuning and compressed model
- Achieved state-of-the-art performance with **94%** prediction accuracy on held-out sections of knowledge graph

Multi-label Subgenre Classification of Rock Music | *Python, PyTorch, Sci-kit Learn* Oct. 2022 – Dec. 2022

- Tested feature reduction techniques to improve multi-label classification in a low data environment

Technical Skills and Achievements

Languages: Java, Python, C/C++, R, Bash, x86-64 Assembly, CUDA
Tools and Frameworks: Git, Linux/Unix, Docker, AWS, PyTorch, Pandas, Google Cloud Platform
Bioinformatics Tools: PyMOL, Geneious, Schrodinger, Biopython, RDkit, Autodock Vina, Protein Design
Achievements: 2 x AIME Qualifier, 2 x USA Biology Olympiad Semifinalist, Top 5 National Science Bowl
Interests: Stanford Climbing, Table Tennis Teams. Adventure Program Specialist. Stanford Alpine Club member

Maggie **Wang**

121 Campus Dr · Stanford, CA 94305

(906)231-3068 | mwang102@stanford.edu | wangmagg | wangmagg

Education

Stanford University — Stanford, CA

PhD Biomedical Data Science — *2021 - Expected 2026*

- GPA: 4.14/4.30
- Relevant coursework: Statistical Inference, Statistical Learning, Applied Statistics, Causal Inference, Data Science for Medicine

Johns Hopkins University — Baltimore, MD

B.S. Biomedical Engineering, B.S. Computer Science — *2017 - 2021*

- GPA: 4.00/4.00
- Relevant coursework: Machine Learning, Deep Learning, Algorithms, Data Structures, Databases, Computational Genomics

Skills

Programming Python, R, C++, C, SQL, Java, PyTorch, Git

Technical causal inference, statistics, experimental design, machine learning

Research

Stanford University — Stanford, CA

Graduate Research Assistant, Advisor: Mike Baiocchi — *2022 -*

- Developing causal inference methods for evaluating behavioral interventions that account for social dynamics and context-dependence [P1].
  Skills: Python, R, experimental design, network modeling, semi-synthetic simulation, stochastic optimization

Stanford University — Stanford, CA

Graduate Research Assistant, Rotation — *2021 - 2022*

- Designed a user study showing that bottleneck models can be edited by humans to improve performance on out-of-distribution data [J2].
  Skills: Python, Streamlit, Hugging Face, PyTorch, interpretable machine learning
- Analyzed national health claims data to reveal gaps in nephrology care for high-risk patients with chronic kidney disease [J3].
  Skills: SQL, Python, descriptive analysis, survival analysis

Johns Hopkins University — Baltimore, MD

Undergraduate Research Assistant — *2018 - 2021*

- Adapted the "surface registration via currents" algorithm to extract hippocampus thickness maps from brain MRI scans.
  Skills: Python, diffeomorphic mapping, optimization
- Applied regression methods to model the relationship between brain atrophy and genetic indicators in Huntington's disease [J1].
  Skills: Python, image quality control, linear regression

Canary Center for Early Cancer Detection — Stanford, CA

Summer Research Intern — *2018*

- Used supervised machine learning methods to enable earlier identification of aggressive cancer from longitudinal blood biomarker trajectories [P2].
  Skills: Python, k-nearest neighbors, auto-regressive modeling

Publications

Journal & Conference Papers

[J1] Chin-Fu Liu, Laurent Younes, Xiao Tong, Jared T. Hinkle, **Maggie Wang**, et al. "Longitudinal Imaging Highlights Preferential Basal Ganglia Circuit Atrophy in Huntington's Disease". *Brain Communications* (2023). DOI: 10.1093/braincomms/fcad214.

[J2] Mert Yuksekgonul, **Maggie Wang**, and James Zou. "Post-hoc Concept Bottleneck Models". *ICLR* (2023). arXiv: 2205.15480.

[J3] **Maggie Wang***, Samson Peter*, Chi Chu, Delphine Tuot, and Jonathan Chen. "Underutilization of Nephrology Referral at High Kidney Failure Risk Levels". *JAMA Network Open* (2022). DOI: 10.1001/jamanetworkopen.2022.25797.

Posters & Invited Talks

[P1] **Maggie Wang** and Michael Baoicchi. "Designing Randomized Experiments for Behavioral Interventions Under Interference and Context-Dependence". *Stanford Causal Science Conference* (2023).

[P2] **Maggie Wang**, Sam Gambhir, and Sharon Hori. "Early Detection of Aggressive Cancer Using Longitudinal Biomarker Measurements". *Early Detection of Cancer Conference* (2018).

Betty Xiong
CV

Contact Details

Telephone: +1 (650) 334-8515
Email: xiongb@stanford.edu

Education Background

2021 – 2026 PhD of Biomedical Informatics, Master in Biomedical Informatics (2023)
Stanford University, School of Medicine, expected graduation 2026
Quad Fellow 2023, HAI Fellow 2022, Fulbright Future Scholar 2021

2014 – 2019: Bachelor of Engineering (Research and Development) (Honours), Bachelor of Science
The Australian National University, High Distinction Average, University Medal

Language skills: English (first language); Fluent in Mandarin, Cantonese and French, spoken and written
Programing languages: Python, C++, R

Work Experience and Employment

Nov 2019 – Jul 2021: Biomedical Engineer, [WearOptimo, Brisbane, Australia]
Position: Graduate engineer
Duties: Research and development into medical devices for skin sensing capabilities – working in a dynamic, multi-disciplinary team environment – developing ML algorithms on data streams for early detection of adverse cardiovascular events

Feb 2015 – Dec 2016: Parliamentary Assistant, [Employer – Department of the House of Representatives, Parliament House of Australia, Canberra]
Position: Messengerial Attendant
Duties: Assistance in the House of Representatives Chamber - Collection and delivery of documents - Assistance in the Serjeant-At-Arms' Office

Jul 2016 – Nov 2016: Intern at the Office of the Chief Scientist [Dr. Alan Finkel, Dr. Will Howard, Dr. Amber Beavis
Position: Intern
Duties: Review of Commonwealth Science Council agenda papers - Research on briefing papers and speeches - Collaboration and discussion of datasheets and data analysis - Assistance with communications outputs

Jul 2015 – Aug 2015: Amgen Scholar Research Internship on the effect of the size of injected nanomedicines on the accumulation and distribution in murine placenta and foetus [Ass. Prof. Horacio Cabral, Kazunori Kataoka Lab, Department of Bioengineering, The University of Tokyo, Japan]
Position: Research Scholar
Duties: Laboratory work involving collaboration with members - Preparation of medicines and animal work

Publications

Betty Xiong, James Zou, Waqar Ali, Roxana Daneshjou, Jonathan Williams (2023). Diagnosis and management of hidradenitis suppurativa: Analysis of US insurance claims data. *JAAD International*, Volume 14

Rishi Bommasani, Kevin Klyman, Shayne Longpre, Sayash Kapoor, Nestor Maslej, **Betty Xiong**, Daniel Zhang, Percy Liang (2023). The Foundation Model Transparency Index. *arXiv*:2310.12941

Daniel N. Sosa, Rogier Hintzen, **Betty Xiong**, Alex de Giorgio, Julien Fauqueur, Mark Davies, Jake Lever, Russ B. Altman (2023). Associating biological context with protein-protein interactions through text mining at PubMed scale. *Journal of Biomedical Informatics*, Volume 145

Wang Wei Lee, Yu Jun Tan, Haicheng Yao, Si Li, Hian Hian See, Matthew Hon, Kian Ann Ng, **Betty Xiong**, John S. Ho and Benjamin C.-K. Tee (2019). A neuro-inspired artificial peripheral nervous system for scalable electronic skins. *Science Robotics, (Selected as Cover)*

Ivy Zhang

ivyzhang@stanford.edu | 862 Cambridge Ave, Menlo Park, CA, 9402 | (646) 436-4263

EDUCATION

Stanford University, Stanford, CA *Expected 2027*
PhD Candidate in Statistics

Yale University, New Haven, CT December 2021
MS in Statistics and Data Science

University of Pennsylvania—The Wharton School, Philadelphia, PA Class of 2018
BS in Economics; Concentration in Statistics; Minor in Mathematics and French

WORK EXPERIENCE

WHOOP, Boston, MA May 2021 – August 2021; January 2022 – August 2022
Data Scientist II, Health Data Science

- Built a production algorithm in Python for disease detection using sensor data, in pipeline for FDA review
- Developed proof of concept analyses to drive go/no-go decisions for R&D's latest product prototypes
- Guided other data scientists/researchers on their projects' statistical methodology

Yale School of Medicine, New Haven, CT September 2020 – August 2022
Graduate Research Assistant

- Researched long-term patient outcome prediction for pediatric, adolescent, and young adult cancer survivors
- Built survival and classification models to predict patient survival and second tumor development

ZS Associates, New York, NY July 2018 – August 2020
Decision Analytics Associate Consultant

- Led teams across multiple offices globally to design and execute analyses to address client business problems
- Developed statistical models in R using patient level data to guide biotech sales and marketing decisions
- Facilitated training sessions on statistical methods and tools for colleagues in the US and India offices

J.P. Morgan, New York, NY Summers 2014 – 2017
Investment Banking Summer Analyst

- Conducted IPO valuation analyses and interfaced with client during roadshow as sole J.P. Morgan representative for a biotech IPO execution
- Built financial models for valuation insights of various financing needs such as mergers, joint ventures, private placement and sell-side for clients in the healthcare and financial institutions sectors

TEACHING EXPERIENCE

Stanford University, Department of Statistics January 2024 – Present
Lecturer: Introduction to R

Middlebury College, Department of Mathematics January 2022 – June 2022
Lecturer: Introduction to Data Science

PUBLICATIONS AND TALKS

- **Zhang, I.**, Hart, G. R., Qin, B., & Deng, J. (2023). Long-term survival and second malignant tumor prediction in pediatric, adolescent, and young adult cancer survivors using Random Survival Forests: a SEER analysis. *Scientific Reports*, *13*(1), 1911. www.nature.com/articles/s41598-023-29167-x
- **Zhang, I.** & Chan, C. (2022, October 07) "Revisiting the glass ceiling: A study of gender gap in statistics academia" [Conference presentation]. Women in Statistics and Data Science 2022

SKILLS AND ACTIVITIES

Tools: Production-level Python, R, and basic SQL; Pycharm, Git, and cluster computing
Languages: Fluent in Cantonese; Proficient in French; Conversational in Mandarin
Activities: Competitive Latin and ballroom dancing; co-founded the first Women in Statistics group at Stanford

Stanford MEDICINE | Department of Biomedical Data Science

COLLABORATION STUDENTS

This section contains resumes of MS students, PhD candidates, and postdocs who are interested in networking, forming industry connections, and potentially forming a collaboration with new research partners

January 23rd 2024 | Stanford University | Stanford Alumni Center

STEPHANIE ARTEAGA

(559) 799-8690 • sarteaga@stanford.edu

EDUCATION

Stanford University **Jul. 2023 – Present**
Ph.D. Student in Biomedical Data Science
University of California, San Diego **Sept. 2013 – Dec. 2017**
Bachelor of Science in Bioengineering: Bioinformatics; GPA 3.808

RESEARCH EXPERIENCE

UNIVERSITY OF CALIFORNIA, LOS ANGELES – LABORATORY OF DR. DANIEL H. GESCHWIND

PROGRAMMER ANALYST II **Jul. 2021 – Jul. 2023**

Position independently analyzing human genetic data for brain development, neurodevelopmental and neurodegenerative disorders.

- Established next-generation sequencing pipelines on local and cloud computational servers.
- Wrote bioinformatic methods and results sections and created figures for manuscript.
- Prepared manuscript for submission to journals.
- Trained current (10+) and new (5+) lab members on performing bioinformatic, computing, and analytic tasks.
- Supervised incoming Computational Staff Research Associate in pipeline modification, scripting, and database maintenance.

STAFF RESEARCH ASSOCIATE II **Jun. 2019 – Jul. 2021**

Position assisting with analysis and processing of neuropsychiatric and neurodegenerative disorder data.

- Conducted statistical analysis of data related to autism spectrum disorder (ASD).
- Performed functional genomic variant annotation.
- Managed and optimized storage and sharing of large genomic datasets.
- Coordinated the transfer and backup of whole-genome and whole-exome sequencing data between servers and cloud services.

STAFF RESEARCH ASSOCIATE I **Jun. 2018 – Jun. 2019**

Position assisting with processing autism spectrum disorder data.

- Developed and maintained bioinformatic pipelines for the analysis of next-generation sequencing data.
- Processed high-throughput whole-genome and whole-exome sequencing data.
- Generated sample databases and performed sample inventory tasks.

TEACHING AND MENTORING EXPERIENCE

UNIVERSITY OF CALIFORNIA, SAN DIEGO – ENGINEERING LEARNING COMMUNITIES PEER EDUCATOR **Sept. 2016 – Dec. 2016**

- Facilitated group learning among underrepresented engineering students by prompting participants to engage actively with course material alongside their peers.
- Fostered a sense of community among participants by instilling motivation and encouragement while eliminating the large classroom environment.
- Assisted students by providing short lectures, exam preparation, and problem-solving methods.

IDEA SCHOLARS BIG PROGRAM **Sept. 2014 – Jun. 2016**

- Served as a mentor, confidant, guide, and friend to incoming IDEA Scholars to aid them in their transition from high school to college during their first year at the university.

PUBLICATIONS

Cirnigliaro, M.*, Chang, T. S.*, **Arteaga, S. A.***, Pérez-Cano, L., Ruzzo, E. K., Gordon, A., Bicks, L. K., Jung, J.-Y., Lowe, J. K., Wall, D. P., & Geschwind, D. H. (2023). The contributions of rare inherited and polygenic risk to ASD in multiplex families. *Proceedings of the National Academy of Sciences*, *120*(31), e2215632120. https://doi.org/10.1073/pnas.2215632120
* Authors contributed equally to this work

Butler-Laporte, G., Povysil, G., Kosmicki, J. A., Cirulli, E. T., Drivas, T., Furini, S., Saad, C., Schmidt, A., Olszewski, P., Korotko, U., Quinodoz, M., Çelik, E., Kundu, K., Walter, K., Jung, J., Stockwell, A. D., Sloofman, L. G., Jordan, D. M., Thompson, R. C., … **Arteaga, S. A.**, … Richards, J. B. (2022). Exome-wide association study to identify rare variants influencing COVID-19 outcomes: Results from the Host Genetics Initiative. *PLOS Genetics*, *18*(11), e1010367. https://doi.org/10.1371/journal.pgen.1010367

HONORS

ADVANCE Scholar, Stanford University **2023**
Cum Laude Honors, University of California, San Diego **2017**
Provost Honors (Awarded 11 of 13 Quarters), University of California, San Diego **2013 - 2017**
IDEA Scholar, University of California, San Diego **2013 - 2017**

Suhana Bedi

Stanford University,
291 Campus Drive, Stanford, CA 94305

+1 (469) 931-4069
Email
Linkedin

Summary

A self-starter and passionate first-year PhD student with experience in knowledge extraction, knowledge graphs, schema development, ontology, tandem coding and hyperscale AI. My interests lie in the spheres of explainable AI in healthcare, LLMs for healthy behavioral changes, and multi-modal data fusion for predictive modeling.

Education

- PhD in Biomedical Data Science, Stanford University, Palo Alto, CA 09.2023 – 09.2028
- B.S in Data Science, UT Dallas, Richardson, Texas (**GPA: 4.00**) 08.2019 – 05.2023

Technical Skills

- `Programming`: C++, Java, Python, R, SQL, Bash/Zsh
- `Tools`: Unix/Linux, Git, Conda, Snakemake, Nextflow, GitHub, Sphinx
- Other languages/software: HTML, Javascript(D3), Markdown, High Performance Computing, LaTeX

Experience

- Student Researcher, Google, Mountain View, CA 01.2022 – Present
  - Designed a biomedical knowledge graph by ingesting and semantically linking 15 publicly available datasets spanning over 300k entities and 950k node-edge-node triples
  - Developed automatically rendered explorer pages based on a Natural Language Interface for biomedical data.
- Undergraduate Research Intern, Institute for Systems Biology, Seattle, WA 06.2021 – Present
  - Performed population based analysis on multi-omics data to identify biomarkers for negative pregnancy outcomes
  - Deployed advanced statistical analysis and models on omics data for correlation studies
- Research Assistant, Prof: Dr Qiwei Li, UT Dallas, Richardson, TX 01.2021 – Present
  - Performed a benchmarking normalization analysis on Spatial Transcriptomics data
  - Deployed several machine learning algorithms for cell type clustering and identification
  - Developed an R package to detect spatially variable genes using BOOST model
- Research Assistant, Functional Genomics Lab, UT Dallas, Richardson, TX 08.2019 – Present
  - Performed ML analysis in Visual Basic to trace phylogenetic branching in interferons
  - Mined several databases to perform variant discovery and detect gene fusion instances in context of neuropathy
  - Performed integrative RNA-seq and CHIP-seq data analysis.

Publications

1. **Bedi S**$^{+}$, Richardson TM^{+}, Jia B^{+}, Saab H, Brinkman FSL, et al. (2022) "Similarities between bacterial GAD and human GAD65: Implications in gut mediated autoimmune type 1 diabetes", DOI: 10.1371/journal.pone.0261103 ***PLOS ONE (2022)***
2. Ferrarini, Mariana, et al. "Global Analysis of Human Sars-cov-2 Infection and Host-virus Interaction." BioHackrXiv, 14 May 2020.

Kelly Brennan

425-221-6100 | kellybrennan35@gmail.com

Education

Stanford University - Computational and Mathematical Engineering Stanford, CA
M.S. Computational and Mathematical Engineering: Data Science & Machine Learning 2021 – 2023

Olin College of Engineering; 3.96 GPA Needham, MA
B.S., Bioengineering

Skills

- **Computing languages:** Python, MATLAB
- **Data Analysis:** Statistics (regression, t-tests, etc.). Machine learning (NLP, LLM, deep learning, data wrangling, etc.). Signal processing.
- **R&D:** Hardware and software testing. Research development, execution, and documentation.
- **Tools:** Snowflake, Git, Postgres, SQL, SQLAlchemy, AWS (EC2, S3, RDS), GCP, PyTorch, TensorFlow.

Engineering Research and Development Experience

Research Data Scientist | **Stanford Medicine** Jan 2023 – Current
Computational Arrhythmia Research Lab Stanford, CA

- Developing NLP model to detect ventricular tachycardia occurrences from clinical notes for semi-supervised learning.
- Using LLM prompt engineering on electronic health records to extract and summarize clinical information.
- Actively engaged in research projects focused on NLP/LLMs of clinical information. Working towards publications in related conferences and journals (Nature Communications, Circulation, AHA, etc.).
- Creating modular research development software tools to organize and analyze data.

Software Engineer | **Form Energy, Inc.** June 2022 – Sept. 2022
Berkeley, CA

- Conceptualized and developed an automated data storage system using Python, SQLAlchemy, and PostgreSQL, and PostgresSQL to replace manual data entry from Excel templates.
- Enhanced the system to accommodate various data types, including metadata and timeseries data.
- Built pip-installable tool for data upload and retrieval via the command line (CLI).

Research Biomedical Engineer | **Kestra Medical Technologies, Inc.** May 2018 – June 2022
Kirkland, WA

- Accomplishments & Impact
  - Analyzed preclinical data using MATLAB and Python, providing visualizations for cross-functional teams, including physicians, engineers, and upper management.
  - Designed, conducted, and reported three pivotal pre-clinical experiments to assess the performance of a wearable defibrillator, supporting FDA IDE and PMA approval.
  - Secured 3 issued patents.
- Actions
  - Collaborated with scientists, engineers, and clinical personnel to make evidence-based product design and research decisions.
  - Developed and validated data processing tools and algorithms in MATLAB and python.

Other activities

Women's Athletics: Mountain biking, cycling (pro-level), skiing, and backpacking.

Kristy A. Carpenter

214 Pine Hill Ct, #203
Stanford, CA 94305

kcarp@stanford.edu
(206)849-5866

Education

Stanford University **September 2020 – present**
Department of Biomedical Data Science Stanford, CA
- PhD candidate in Biomedical Informatics, advised by Prof. Russ Altman
- Proposed thesis work in structure-based algorithms for adverse drug reaction prediction

Massachusetts Institute of Technology **September 2016 – May 2020**
Department of Electrical Engineering & Computer Science Cambridge, MA
Department of Biology
- B.S. in Computer Science & Molecular Biology
- GPA: 5.0/5.0; Phi Beta Kappa inductee

Work Experience

Graduate Research Intern **June 2023 – August 2023**
Merck Research Laboratories South San Francisco, CA
- Intern in Modeling & Informatics Department
- Designed and implemented a machine learning and molecular dynamics workflow for druggability analysis
- Designed and ran atomistic and coarse-grained molecular dynamics simulations in Gromacs and OpenMM
- Prepared and presented a poster

Undergraduate Research Intern **June 2019 – August 2019**
Oak Ridge National Laboratory Oak Ridge, TN
- Implemented, trained, and tuned a novel machine learning architecture in Keras
- Demonstrated that the model can be used to accelerate molecular simulations
- Prepared and presented a poster

Undergraduate Researcher **October 2018 – May 2019**
Massachusetts Institute of Technology Cambridge, MA
- Built and tuned recurrent neural network with Keras to classify antimicrobial peptides
- Ran parallelized classifier training jobs on computing cluster for hyperparameter optimization
- Presented at group meetings

Selected Publications

Carpenter, K.A. & Altman, R. B. (2023). Using GPT-3 to Build a Lexicon of Drugs of Abuse Synonyms for Social Media Pharmacovigilance. *Biomolecules, 13*(2), 387.

Derry, A.*, **Carpenter, K. A.*,** & Altman, R. B. (2022). Training data composition affects performance of protein structure analysis algorithms. *Pacific Symposium on Biocomputing*, *27*, 10–21. *equal contribution

Ingólfsson, H. I., Bhatia, H., Zeppelin, T., Bennett, W. F. D., **Carpenter, K. A**., Hsu, P. C., Dharuman, G., Bremer, P. T., Schiøtt, B., Lightstone, F. C., & Carpenter, T. S. (2020). Capturing Biologically Complex Tissue-Specific Membranes at Different Levels of Compositional Complexity. *Journal of Physical Chemistry B*, *124*(36), 7819–7829.

Carpenter, K. A. & Huang, X. (2018). Machine Learning-based Virtual Screening and Its Applications to Alzheimer's Drug Discovery: A Review. *Current Pharmaceutical Design*, *24*(28), 3347–3358.

Awards

Ruth L. Kirschstein NRSA F31 Diversity Fellow 2023 – present
NSF Graduate Research Fellow 2021 – 2023
GEM Fellow 2021 – 2022

Erin Craig

Ph.D. Candidate, Biomedical Informatics, Stanford

erincr@stanford.edu | 941.527.2156 | erincraig.me

EDUCATION

Ph.D. Biomedical Informatics, Stanford 2019–2024 (expected)

Concentration: development of statistical methods for antibody discovery

Award: Stanford Data Science Scholar ('22-'24), Gabilan Fellow (Stanford Graduate Fellowship, '19-'22)

Advised by: Robert Tibshirani

M.S. Data Science, New College of Florida 2015–2017

B.A. Mathematics, New College of Florida 2005–2009

EXPERIENCE

Data Science Intern, BlackRock AI Labs Summer 2021

Built predictive models with financial data.

Data Science Consultant, Jonathan Kroner Law Office 2015–2023

Analyze Medicare data, including anomaly detection and visualization, in support of fraud investigations.

Data Scientist, Florence A. Rothman Institute 2017–2018

Used deep learning on clinical notes to predict hospital readmissions.

Data Science Consultant, Haystack Informatics 2017–2018

Anomaly detection on hospital computer use data to identify cases of employee snooping on patients.

Wolfram|Alpha

Oversaw math content development. Architected and led development of Step-by-step Solutions and the Wolfram Problem Generator. Led teams to build and maintain technologies related to education.

Manager/Lead Developer: Educational Software Technologies and Math Content	2011–2015
Developer/Senior Developer: Math Content	2009–2011
Lead Instructor: Mathematica Summer Camp	2010–2014

PUBLICATIONS

Craig, Zhong, and Tibshirani. Survival stacking: casting survival analysis as a classification problem. In review. arXiv preprint: arXiv:2107.13480 (2021).

Craig, Redelmeier, and Tibshirani. Finding and assessing treatment effect sweet spots in clinical trial data. arXiv preprint: arXiv:2011.10157 (2020).

Craig, Arias, and Gillman. Predicting Readmission Risk from Doctors' Notes. *NeurIPS, ML for HC*. Dec 2017.

Sherman, **Craig**, E. Yanovich, Ketko, Kalmanovich, and R. Yanovich. Standardized Interpretation of Heat-Tolerance-Testing Results: Probability of Intolerance Instead of Specialist Judgment. Journal of Athletic Training, V53 N4. pp. 423-430.

Childers, **Craig**, Taha, and Poimenidou. A p-adic Approach to Binomial Identities. The Pi Mu Epsilon Journal, Volume 13, Number 3. Fall 2010. pp. 133-142.

EXTRACURRICULAR ACHIEVEMENTS

Contemporary Dancer: Performed professionally in San Francisco CA, Sarasota FL, Boston MA and Cleveland OH. Studied at the Joffrey Ballet, the San Francisco Conservatory of Dance, the Martha Graham Winter Intensive and Ballet Chicago; performed works by William Forsythe, Jiří Kylián, Ohad Naharin and George Balanchine.

Sohaib Hassan
736 Serra Street | Stanford, CA 94305 | (732) 527 7010 | sohaib@stanford.edu

EDUCATION

Rutgers School of Arts and Sciences
Bachelor of Arts in Genetics — May 2023
Honors Certificate in Computational Genetics
GPA: 3.98

Stanford University School of Medicine — TBD
Biomedical Informatics PhD Candidate

RESEARCH EXPERIENCE

Verzi Lab - Rutgers University *Research Intern* — May 2020 - July 2023
- Investigate gene and protein activity through use of advanced RNA-seq, ChIP-seq, and ATAC-seq computational tools
- Lead computational biology group and collaborate on projects using transgenic and treatment-induced mouse modeling
- Co-authored the Verzi Lab Computational Guide on Bulk and Single-Cell RNA-seq
- Gained high proficiency in R, Python, Bash, and Unix/Linux

Montgomery Lab - Stanford University *Rotation Student* — September 2023 - Present
- Conducting co-localization and fine-mapping of GTEX nerve samples against GWAS ALS

PROJECTS

Finding shared causal variants in GTEX and ALS GWAS — September 2023 - Present
- Creating TensorQTL/SuSiE pipeline to identify causal variants
- Comparing PCA versus PEERs for covariate regression in GTEX pipeline

Discovering Changes and Crosstalk in Epithelium-Mesenchyme Microenvironment During Tumorigenesis — April 2021 - July 2023
- Created and utilized Scanpy/Seurat/CellChat pipeline to determine cell signaling between epithelial and mesenchymal cells
- Discovered novel communication between intestinal cells in EGF and Hippo pathways
- Currently writing manuscript for summer 2023 submission

Genomic analysis of 1,25-dihydroxyvitamin D3 action in Mouse intestine reveals compartment and segment-specific Gene regulatory effects — May 2020 - June 2022
- Applied Kallisto/DESeq2 to analyze differential expression of genes, integration of ChIP-Seq data to discover tissue-specific target genes for vitamin D3, and gene set enrichment analysis to confirm vitamin D_3 as transcriptional activator
- Journal of Biological Chemistry co-first author summer 2022 publication

FELLOWSHIPS AND INTERNSHIPS

Bristol Myers Squibb *Data Science Co-Op* — January 2022 - July 2022
- Developed flask app to analyze biomarker impact on treatment/disease predictions with decision tree-based ML algorithms
- Created hybrid Bulk/Single-cell RNA-Seq pipeline to analyze CRISPR assay cells for differential gene expression
- Utilized machine learning to create model for cell clustering based on gene expression and lipid properties

Rutgers Initiative for Maximizing Student Development *Researcher* — May 2021 - August 2021
- NIH funded summer research program focused on matching students with state of the art laboratories
- Researched and created models on effect of intestinal cell signaling in tumorigenic microenvironment under mentorship of Dr. Langer, Dr. Verzi, and Dr. Pellon-Cardenas

LEADERSHIP

Rutgers-Seq: Bioinformatics Society *President and Founder* — May 2021 - May 2023
- First undergraduate bioinformatics-oriented organization in Rutgers
- Teach students bioinformatic techniques/applications in demo-supported meetings and connect them with computational labs

Collin Hitter

chitter@wesleyan.edu / (631) 358-1242

EDUCATION

Stanford University School of Medicine | *Master of Science Candidate* Stanford, CA, Apr 2023 – e. Mar 2026

- Biomedical Data Science | 4.0 GPA | Selected to join a class of ~10 professionals pursuing the degree part-time

Wesleyan University | *Bachelor of Arts* Middletown, CT, Aug 2014 – May 2018

- *Major*: Molecular Biology & Biochemistry | *Minor*: Economics | *Certificate*: Informatics & Modeling | 3.8 GPA

WORK EXPERIENCE

Moderna | *Manager, Senior Manager: Portfolio Analytics* Cambridge, MA, Nov 2021 – Present

- Support Moderna's R&D organization by managing projects seeking to optimize activities for ~50 preclinical – Phase 3 assets, collaborating with clinical, research, strategy, informatics, and finance functions
- Regularly present to clinical development leadership on project strategies, output, and progress
- In collaboration with digital and clinical functions, produce weekly and monthly R&D progress reports, contextualizing insights for c-suite executives, R&D leadership, and clinical teams
- Support analytics and reporting strategy by assessing which data are most relevant, identifying optimal data sources, and contextualizing new insights for c-suite executives and R&D leadership
- Help drive our global R&D strategy by managing internal and external teams to translate epidemiological and clinical needs into specific data requests, namely evidence generation efforts using NHS England's real-world data
- Led cross-functional workshops to enhance budget cycles, establishing a process and system for real-time reviews and creating a model to phase costs based on study activities, including probabilistic enrollment projections
- Rolled out a new governance framework to ~40 preclinical and clinical development programs, refining decision making forums and establishing a consistent program team framework / operating model

PicnicHealth | *Life Sciences Partnerships Manager* San Francisco, CA, Dec 2020 – Nov 2021

[Patient-centric real-world data startup that is democratizing medical records to advance leading clinical research]

- Managed internal and external initiatives across data science, clinical, engineering, HEOR, and RWE teams to ensure timely data deliveries and impactful publications/abstracts for 9 programs across rare and large market diseases
- Facilitated working sessions with our Chief Medical Officer, Head of Quantitative Science, and a top 5 pharma company's RWE team to determine our analytical approach, ultimately publishing PicnicHealth's first EHA abstract
- Worked directly with the CEO and Chief Medical Officer to develop a de novo abstraction services product-line

ClearView Healthcare Partners | *Analyst, Senior Analyst, Consultant* New York, NY, Aug 2018 – Dec 2020

[Boutique strategy consultancy consistently ranked in Vault's Top 10 Firms for Health Sciences]

- Collaborated with clients and teams of 2 – 8 people to execute engagements with 7 large pharmaceutical and 9 small/mid-sized biotech companies seeking guidance on commercial, corporate, M&A, and R&D strategies
- Conducted extensive research through syndicated databases, internet searches, and medical journals; mentored analysts; interviewed KOLs; managed workflow and client dynamics to drive project deliverables and presentations
- Assisted with due-diligence efforts and created multiple models for investors and small & large-cap companies, including a 30+ asset P&L tool that encompassed marketed, pipeline, and business development targets for a fortune 500 company's vaccines and anti-infectives business unit
- Analyzed the universe of genetically-linked neuromuscular diseases for a leading gene therapy company, developed revenue models for top indications, and ultimately recommended they expand into alternative therapeutic areas
- Characterized the AML landscape across scientific, clinical, pipeline, and market access dynamics, eventually outlining inflection points and recommendations for a preclinical cell therapy asset's R&D path

OTHER EXPERIENCE, SKILLS, & PERSONAL

- **The Taylor Lab (Wesleyan)**: Drug Discovery Research Assistant | Conducted in vitro antibiotic design experiments
- **The Patricelli Center:** Mentor and Fellow | Advised students solving social issues via entrepreneurial ventures
- **Programming:** Python (including pandas), C++, R, SQL, Tableau, Excel VBA
- **Syndicated Databases:** Citeline Informa, Global Data, AdisInsight, Evaluate Pharma, IHS Markit, Syneos Health
- **Certificates:** London School of Economics | Macroeconomics; Moderna/Carnegie Mellon University | AI Awareness
- **Achievements:** NCAA DIII Lacrosse Champion & 3-yr starter; Moderna Clinical Development "Star" of the Month
- **Interests:** Skiing, Hiking, Running, Oyster Farming, Travel, Tennis, Beach, Freakonomics Radio Podcasts

Marie Huynh

US Permanent Resident | marie-huynh | portfolio | mahuynh@stanford.edu | +1(914)349-7286

Education

Stanford University (California, USA), *Academic Research Master in Biomedical Data Science* 2022 - present

GPA: 3.98/4.0
Relevant Coursework: Design and Analysis of Algorithms, Machine Learning, Algorithms for Computational Molecular Biology, NLP with Deep Learning, Statistical Inference, Deep Learning for Computer Vision, Data Science for Medicine.

Ecole Polytechnique (Paris, France), *Bachelor of Science in Mathematics & Computer Science* 2019 - 2022

GPA: 3.97/4.0, Magna Cum Laude
Minor: Biology – Relevant Coursework: Probability, Statistics, Linear Algebra, Topology and Multivariable Calculus, Networks, Object-Oriented Programming, Computer Architecture, Concurrent Programming, Cell Biology, Molecular Genetics.

Research and Work Experience

Wall Lab, Stanford University (California, USA), *Graduate Research Assistant* Jun 2023 - present

Work under the supervision of Dr. Dennis Wall (Principal Investigator).

- Developed an advanced protocol to filter high-quality videos from 3000 structured videos of children with and without autism spectrum disorder (ASD) to predict ASD using eye gaze, head pose, and facial landmarks.
- Constructed individual GRU/LSTM-based models using eye gaze, head pose, and facial landmarks to predict ASD.
- Employed late and intermediate fusion techniques to construct ensemble models to predict ASD with SATO performance.

Biomedical Data Science Dpt., Stanford University (California, USA), *Teaching Assistant* Sep 2023 - Dec 2023

Montgomery Lab, Stanford University (California, USA), *Graduate Research Assistant* Jan 2023 - Sep 2023

Work under the supervision of Dr. Stephen Montgomery (Principal Investigator).

- Clustered metabolites and proteins into modules using weighted correlation network analysis (WGCNA) to investigate the major axes of variance explained by the metabolome and proteome. Interpreted unknown metabolites.
- Currently investigating gene x environment interactions at the molecular level using eigenmetabolites and eigenproteins as proxies for environmental exposure.

Moderna Therapeutics (Cambridge, USA), *Bachelor Thesis* Jan 2022 – Apr 2022

Work under the supervision of Dr. Mihir Metkar and Dr. Eric Ma. (Grade : A+)

- Built and trained a transformer model using Pytorch Lightning to predict a mRNA sequence for a given protein sequence using publicly available human transcriptome data.
- Conducted a comprehensive analysis of codon usage patterns within the predictions, identifying existing biological patterns and exploring the utilization of less common codons.

Awards

Women in Science Scholarship, École Polytechnique (Paris, France) 2019–2020

Outstanding Student Distinction, École Polytechnique (Paris, France) 2019–2020

US President's Volunteer Service Award (New York, United States), *Bronze (2017), Gold (2018)* 2017–2018

Leadership & Volunteerism

École Polytechnique's Bachelor Students' Organization (Paris, France), *Treasurer* Jun 2020 – Jun 2021

- Found and negotiated sponsorships, managed the budget, and oversaw the subscriptions and expenses.
- Collaborated with 8 other members to manage Bachelor student-life on campus, including events, sports, job fairs, ...

Safe Water Cube (Paris, France), *Active Member/Fundraiser* Jun 2020 - Jun 2021

- Collaborated with the association Safe Water Cube and raised 1.5k during the pandemic to contribute to building an innovative potable-water fountain in Laos.

Miscellaneous

Languages	French (native speaker), English (fluent, Toefl IBT: 112), Spanish (advanced level, C1)
Interests	Tennis, Rugby and Literature
Skills	Python (PyTorch, Tensorflow), R, C/C++, Git/Github

Riley Juenemann

720-438-8077 | rjuene@stanford.edu | linkedin.com/in/rjuenemann/

Education

Stanford University, Stanford, CA
Doctor of Philosophy, Computational & Mathematical Engineering (ICME) June 2026
Doctoral Dissertation Advisor: Markus Covert | GPA: 4.0

Tulane University, New Orleans, LA
Bachelor of Science, Summa Cum Laude with Honors in Mathematics May 2021
Majors: Mathematics and Computer Science | Minor: Management | GPA: 4.0

Research Experience

Stanford University, Stanford, CA

- Whole-Cell Modeling of *Escherichia coli* September 2022 – Present
  - Introducing genetic engineering capabilities to the *E.coli* whole-cell model, which utilizes nearly 20,000 parameters from heterogeneous experimental data sets to simulate many distinct cellular processes and capture their complex interactions on a system-wide level. The model is publicly available at https://github.com/CovertLab/wcEcoli
  - Analyzed sets of over 4000 simulations on Stanford's Sherlock compute cluster

Tulane University, New Orleans, LA

- Undergraduate Research in Computational Biology May 2018 – May 2021
  - Developed statistical and machine learning tools in R and Python to automatically categorize particle trajectories from live-cell data with Prof. Scott McKinley
  - Systematically executed simulations and used topological data analysis to investigate the formation of ring channels in cells for submission of publication

Duke University, Durham, NC

- NSF REU for Meeting the Grand Challenges May 2019 – July 2019
  - Investigated the internalization of titanium dioxide nanoparticles into human lung cells in Prof. Christine Payne's Lab
  - Developed skills in live-cell culture, fluorescence and transmission electron microscopy

Internship Experience

Eli Lilly and Company, Indianapolis, IN

- Information and Digital Solutions (IDS) Intern May 2021 – August 2021
  - Proposed integration strategy for genomics data within existing cross functional frameworks
  - Created heuristic methods in R to extract relevant text from unstructured documents
- Information and Digital Solutions (IDS) Intern May 2020 – August 2020
  - Devised a manufacturing value stream model in R using discrete event simulation to assess interrelationships and enhance the ability to analyze proposed improvements
  - Deployed R Shiny application to be used by manufacturing personnel

Languages and Skills

Python, C++, R, Shiny, LaTeX, GitHub, CUDA, Parallel Computing, MATLAB, Java, SQL, C, JavaScript, D3, Rails, Ruby on Rails, HTML, CSS, Haskell, ImageJ

Publications

MV Ciocanel, **R Juenemann**, AT Dawes, SA McKinley. Topological data analysis approaches to uncovering the timing of ring structure onset in filamentous networks.
Bulletin of Mathematical Biology, **83**(10), (2021). DOI.

Awards

- NSF Graduate Research Fellowship September 2021
- Stanford Graduate Fellowship: Inventec Fellow September 2021
- Enhancing Diversity in Graduate Education Doctoral Fellowship September 2021
- Stanford ICME Xpo Research Symposium Third Place Poster Award May 2023
- First Place Prize for Undergraduate Research November 2020
  NSF-Simons Center for Quantitative Biology
- William Wallace Peery Medal (top academic honor for Tulane undergraduates) May 2021

Conferences

- American Mathematical Society (AMS) SE Sectional Meeting (Talk) November 2021
- Heidelberg Laureate Forum September 2020
- Emerging Research Trends in Computer Science August 2019
  The Cornell, Maryland, Max Planck Pre-doctoral Research School

Cyrus Knudsen

Education

2021–now **Bioengineering, Ph.D.**, *Stanford University*.
Collaboratively building the world's first comprehensive model of E. coli to facilitate computer-aided rational design in biology (Covert lab). GPA: 3.6 (US scale).

2015–2021 **Biotechnology, M.Sc.**, *Technical University of Denmark*.

2019 **Visiting Student**, *Harvard University*.

Research experience

2021–now **Covert lab**, *Stanford University*.
Creating complex systems models of large biochemical networks based on massive amounts of heterogeneous data

- Developed new method based on convex optimization that infers parameters by integrating high-dimensional data from various physical scales.

Publications

Milne, N., Thomsen, P., **Knudsen, C**, Rubaszka, P., Kristensen, M., and Borodina, I. Metabolic engineering of Saccharomyces cerevisiae for the production of psilocybin and related tryptamine derivatives. *Metabolic Engineering* (2020).

- **88th** percentile of most cited papers in journal, **99th** percentile of mentions in newspaper articles and social media posts (three year period).

Posters and presentations

2023 **Stanford Data Science Conference**, *Lightning Talk & Poster*, Stanford, CA.
Using optimization to model metabolic networks

2023 **Synthetic Biology: Engineering, Evolution & Design**, *Poster*, Burbank, CA.
Describing enzyme function and regulation using convex optimization

2020 **iGEM 2020 Conference**, *Poster and presentation*, Virtual.

Awards and fellowships

2019 **Fulbright Scholar**, *Harvard University*.

Work experience

2019 **Bioinformatics Intern**, *Novo Nordisk*.

- Designed and developed a deep-learning based segmentation model (PyTorch).

2018–2019 **R&D Intern**, *Biosyntia*.

- Proposed, developed and maintained a server-based data warehouse and data processing pipeline of experimental data (Python, Django, PostgreSQL).

ROHAN VISHESH KOODLI

koodli.com ◇ koodli@stanford.edu ◇ /RK900 ◇ /RohanKoodli

EDUCATION

PhD, Biomedical Informatics, Stanford University 2022 - Present
Advisor: Ron Dror

MS, Electrical Engineering & Computer Science, University of California, Berkeley 2021 - 2022
Advisor: Nir Yosef

BA, Computer Science, University of California, Berkeley 2018 - 2021
Graduated with Distinction

SKILLS

Languages	Python, Java, C/C++, R, SQL, RISC-V Assembly
Libraries	PyTorch, JAX/Flax, TensorFlow, NumPy, Pandas, scikit-learn, Biopython, PyMOL
Tools	Jupyter Notebooks, Git, LaTeX, AWS, GCP, Docker

RESEARCH EXPERIENCE

Dror Lab, Stanford University, Stanford AI Lab — April 2023 - Present
Yosef Lab, UC Berkeley, EECS & Center for Computational Biology (CCB) — April 2020 - August 2022
Das Lab, Stanford University, Department of Biochemistry — December 2016 - March 2020

INDUSTRY EXPERIENCE

Microsoft Research (Office of the CSO), Research Intern — May 2023 - August 2023
Atomic AI, ML Engineering Intern — January 2022 - April 2022
Google Search, Software Engineering Intern — May 2020 - August 2020
Google Cloud, Software Engineering Intern — June 2019 - August 2019

PUBLICATIONS & PREPRINTS

MultiVI: deep generative model for the integration of multimodal data
Tal Ashuach, Mariano I. Gabitto, **Rohan V. Koodli**, Giuseppe-Antonio Saldi, Michael I. Jordan, Nir Yosef
Nature Methods, June 2023

PolyVI: Deep Generative Models for Gene Expression, Chromatin Accessibility, and Surface Protein Expression Data
Rohan Koodli
Master's Thesis, May 2022

Redesigning the Eterna100 for the Vienna 2 folding engine
Rohan V. Koodli, Boris Rudolfs, Hannah K Wayment-Steele, Eterna Structure Designers, Rhiju Das
bioRxiv, August 2021

EternaBrain: Automated RNA design through move sets and strategies from an Internet-scale RNA videogame
Rohan V. Koodli, Benjamin Keep, Katherine R Coppess, Fernando Portela, Eterna participants, Rhiju Das
PLOS Computational Biology, June 2019

AWARDS & FELLOWSHIPS

Stanford HAI Google Cloud Grant ($15,000) 2023
NSF Graduate Research Fellowship 2022
Regents' and Chancellor's Scholarship 2018
Valedictorian 2018
Siemens Competition National Semifinalist 2017

Jodie Meng

658 Escondido Rd, Stanford, CA 94305 | (630) 881-9990 | jodieme@stanford.edu

EDUCATION

Stanford University **Stanford, CA**

Honors B.S. Candidate in Biomedical Computation Expected Graduation: June 2024

M.S. Candidate in Biomedical Informatics Expected Graduation: June 2025

Relevant Coursework: Applied Statistics, Nonparametric Statistics, Data Structures and Algorithms, Probability, Linear Algebra and Multivariable Calculus, Operating Systems

SKILLS

R, Python, MS Office Suite Products (Excel, Powerpoint)

Flow cytometry analysis, cell culture, transfection/transformation, RNA extractions, polymerase chain reactions, gene knockdowns, and titrations.

RESEARCH EXPERIENCE

Kara Davis Lab, Pediatric Hematology and Oncology **Stanford, CA**

Research Assistant *January 2022–present*

- Investigate mechanisms of leukemia relapse following CAR-T cell therapy by characterizing the effects IKAROS and PAX5, two related developmental proteins, on CD19 antigen loss. Presented at 2022 Bio-X Conference at Stanford Medicine.
- Develop independent honors thesis project on identifying gene expression signatures and leukemia cell populations associated with patient relapse across B-ALL subtypes, with a goal of informing therapeutic strategies for high-risk patients. Implement single-cell sequencing pipelines in R with patient data.

Stanford Medicine, Primary Care and Population Health **Stanford, CA**

Fellow *July 2021–December 2022*

- Propose, create, and execute IRB-approved qualitative research study to evaluate the health and safety needs of Mandarin-speaking Chinese elders in Bay Area during COVID-19 pandemic. Recruit participants from Stanford Health Care; conduct and transcribe interviews in Mandarin on patients' social, mental, and emotional well-being. Paper-in progress.

Frock Lab, Radiation Oncology **Stanford, CA**

Research Assistant *July 2021–December 2021*

- Investigate backup repair pathways of cancer cells for potential chemotherapeutic applications. Generate High-Throughput Genome-Wide Translocation Sequencing (HTGTS) libraries to locate and measure cellular DSB patterns. Create visualizations of DNA recombination in Non-Homologous End Joining (NHEJ) and Alternative End Joining (A-EJ) repair mechanisms, review paper published in *Biomolecules* (2021).

LEADERSHIP EXPERIENCE

Camp Kesem at Stanford **Stanford, CA**

Director, Operations Coordinator, Camp Counselor *December 2020–present*

Cardinal Free Clinics **San Jose, CA**

Mandarin Interpreter Chair, Patient Outreach Chair *January 2022–present*

Stanford Data and Mapping for Society **Stanford, CA**

Data Journalism Editor *June 2020–December 2021*

PERLA MOLINA

perlam@stanford.edu | linkedin.com/in/perlamolina/ | https://perlamolina.me/ | 775-450-8323

EDUCATION

Stanford University — Stanford, CA *Jun 2023 - Current*
- **Ph.D.:** Biomedical Data Science

University Of San Francisco — San Francisco, CA *Aug 2019 - May 2023*
- **Bachelor of Science:** Major in Data Science *GPA 3.66*
- **Awards & Honors:** Provost Scholarship; Cum Laude Honors; Pi Mu Epsilon Inductee 2022; 2022 & 2021 Undergraduate Newmark Women in Tech Scholarship; Hispanic Scholarship Fund Scholarship 2022; Dean's Honor Roll (Spring 2020 - Spring 2023)

EXPERIENCE

Redwoods Analyst *Jun - Aug 2022*
DaVita Kidney Care - San Francisco, CA
- Worked with the DaVita Clinical Research team on research tasks involving qualitative research and Natural Language Processing
- Programmed survey for Care Partner Stress on Qualtrics and aided in the launching process
- Conducted preliminary and analytical research on the Care Partner Stress survey results & presented findings to the team

PROJECTS

Predicting Pre-Term Birth From Proteomics Pregnancy Data - Lab Rotation Project *Jun - Aug 2023*
- Used XGBoost and LASSO machine learning methods to create algorithms to predict pre-term birth from proteomic pregnancy data
- Analyzed performance and significance between trimesters 1 and 3, and their delta
- Results from XGBoost model showed better performance and p-value in trimester 3 and the trimester delta
- Results from LASSO model revealed dimension reduction doesn't contribute to significance of model for any trimester modality

Linear Regression Methods on Cardiovascular Disease Dataset - Class Project in R *Dec 2022*
- Created functions using LASSO, Ridge Regression, and Ordinary Least Squares methods to perform predictive and explanatory analysis on any dataset
- Selected the best predictive model for high systolic blood pressure of patients with cardiovascular health problems
- Conducted diagnostics and transformations on the data to enhance analysis of models
- Discovered high sensitivity of the data by outliers and the necessity to remove them to optimize performance of regression functions
- Explanatory model revealed the most significant factors associated with high blood pressure are gender, height, weight, cholesterol, activity level, and the presence of cardiovascular disease
- Concluded further study needs to be done to examine efficacy of treatment plans that include diet changes and/or increased activity level to target weight and cholesterol

Performing Data Mining Methods on Ovarian Cancer Data - Class Project in R *Nov - Dec 2022*
- Applied classification analysis to predict a Benign Ovarian Tumor or Ovarian Cancer
- Performed cluster analysis to categorize ovarian tumors
- Results showed high accuracy of the decision tree for classification analysis and overall two types of ovarian tumors in the dataset with low total variance
- Highlighted the need for more data in this particular set due to low total variance in the cluster analysis

EXTRACURRICULAR ACTIVITIES

Association of Women in Mathematics: Vice President from *May - Dec 2021*; President from *Jan 2022* *Aug 2020 - May 2023*
- Created a website for students and faculty to access information about the club, upcoming events, and important announcements (https://sites.google.com/view/awmusfca)
- In charge of administrative tasks and overseeing activities including but not limited to coordinating meetings, organizing operational tasks of all other executive board members, and communicating with faculty advisors and school administration to approve events
- Administered an initiative to revive the club after the pandemic and encourage women in mathematics and other related fields

Data Science Association *Aug 2020 - May 2023*
- Successfully funded 12 students to attend Data Science Conference 2023 at the San Francisco Data Science Institute

ACV (Anime, Culture, Video Games) Club: Executive Board Member from *May 2022* *Aug 2019 - May 2023*

Women in Tech Club *Aug 2019 - Mar 2022*

SKILLS/INTERESTS/OTHER AFFILIATIONS

- **Coding Languages**: HTML (Expert), CSS (Expert), R (Proficient), Python (Proficient), SQL (Proficient), Java (Proficient), Javascript (Beginner)
- **Software & App Programs**: Adobe (Photoshop, Illustrator, Flash, Dreamweaver), Google Suite, Microsoft Excel, Qualtrics, Github
- **Language**s: Spanish (Native)
- **Conferences**: Women in Data Science (2022), Hack the Gap Hackathon (2021), Nebraska Conference for Undergraduate Women (2021), Tapia Conference (2020)

Sayuri Monarrez Yesaki

Biochemist with a career in Data Engineering - 5 years of experience in the biotech industry, with a passion for developing data workflows and computational tools to enable scientists to develop ground-breaking therapies.

WORK EXPERIENCE

Denali Therapeutics | South San Francisco CA
05/2022 - Present

Associate Data Engineer

- Enable capturing, managing, integrating, and analyzing scientific data using a combination of open-source, in-house, and commercial solutions. Define and document best practices for capturing and entering scientific data and metadata, and educate scientists about these standards.
- Maintain and enhance Denali's enterprise-level lab informatics management solutions - Dotmatics and Benchling suits, and provide ELN user training.
- Analyze and fix ETL and pivot failures and data abnormalities, and ensure scientific stakeholders have access to high quality and up to date data.
- Work with bench scientists and application scientists to understand their data needs and create scripts to parse, reformat, transform, verify and upload data and metadata to Denali's Data Warehouse.
- Explain information content and context to scientists and generate fit-for-purpose datasets.

Macromoltek | Austin TX
09/2019-09/2021

Scientific Software Engineer

- Planned, designed, implemented, and tested new software approaches for computer-assisted drug discovery technologies specialized in antibodies.
- Successfully built internal tools using Google Apps Scripts to manage, process, and archive project data. **Saves the Principal Scientist 8 hours per week.**

University of Texas Austin | Austin TX
01/2019-07/2019

*Teaching Specialis*t

- Taught 140 Genetics students per semester across seven different discussion sections per week + office hours. Proctored and graded quizzes and exams.

Reliant Immune Diagnostics | Austin TX
01/2018-05/2018

Lab Technician Intern

- Performed tests, collected and verified data for precision and reproducibility study for FDA application for a medical device.

PROJECTS

University of Texas Austin
Fall 2018

Data Analytics Course

- Trained various classification models using scikit-learn to predict whether or not a patient has diabetic retinopathy. Models included k-Nearest Neighbors, Naïve Bayes, SVM, Neural Network (MLP), and Boosted Decision Trees.
- Applied several clustering techniques in scikit-learn to medical data in order to group different types of seizures. Algorithms included K-Means, DBSCAN, and HAC.
- Performed association analysis on a retail dataset using mlxtend to find the frequently purchased itemsets and strong association rules.

(+1) 512-766-9320
sayuri.monarrez@gmail.com
GitHub: SayuriMY
LinkedIn: SayuriMY

EDUCATION

Stanford
MS in Biomedical Informatics
Fall 2023 - TBD
Enrolled part-time

University of Texas Austin
B.S. Biochemistry
Certificates:
• Elements of Computing
• Scientific Computing and Data Science
GPA: 3.8
08/2016-12/2018

University of Melbourne
Study Abroad
07/2015 - 07/2016

AWARDS

Undergraduate Research Fellowship
University of Texas Austin
Spring 2018

TECHNICAL SKILLS

Proficient: Python, SQL, C#

Basic: JavaScript, Google Apps Script, Pandas, Numpy, Scikit-learn

Familiar: R, BigQuery

(Familiar) Data Mining: Classification, Clustering, Association analysis, Dimensionality reduction, Supervised and Unsupervised Machine Learning

ENTERPRISE SOFTWARE

Dotmatics
Benchling
Lucidchart
pyMol

KEITH E. MOORE

kem1@stanford.edu | linkedin.com/in/keith-moore-hp

BIOMEDICAL MACHINE LEARNING AND DATA SCIENCE

An 'in the trenches' technology leader with a track record or turning patents into products. Trusted at all levels of the organization. Strategic thinker and inventor who can see the forest and the trees.

AREAS OF EXPERTISE

Applied Research | Intellectual Property Strategy | Program Management |
Machine Learning | Embedded Systems | Product Development

PROFESSIONAL EXPERIENCE

HP Inc., Palo Alto, CA

Head of Microfluidics and Life Sciences, HP Labs **2018-2021**
Created the Life Sciences Lab and managed research in molecular diagnostics and small molecule detection.

- Established the Microfluidics lab and created new research projects in digital biology.
- Recruited multi-disciplinary team focused on single cell dispense and nano-fabrication of molecular sensors.
- Created technology for single cell dispense for HP digital dispense/titration products

Head of Print Adjacencies and 3D Lab, HP Labs **2015-2018**
Created the 3D Printing lab at HP labs and managed physical science research and laboratories..

- Delivered results that compelled HP to enter 3D polymer and metals market (multi-billion $ opportunity).
- Established strong IP portfolio for protecting HP's 3D business.

IP Strategist and Deputy Director, HP Labs **2014-2015**
Separated HP Labs as HP separated into 2 companies.

- Co-authored the HP Intellectual Property Strategy
- Partitioned over 19,000 patents between HP Inc and Hewlett Packard Enterprise.
- Created the structure and research portfolio for the new HP Labs at HP Inc.

Chief Technologist, LaserJet and Enterprise Solutions **2011-2014**
Defined the interconnect protocols to ensure HP products (old and new) work with latest operating systems.

- Architected security offerings justifying claims of "most secure printer."
- Dramatically lowered warranty cost with new test approach that won President's Quality Award.
- Recognized as HP Fellow (only 50 in 270,000 HP employees) for technical contributions.

PATENTS

US6142685, US5774377, US5918235, US6173309, US6282581, US6189046, US6347319, US6781069, US6408342, US6564304, US6880724, US7065137, US7028188, US7856657, US7730452, US7716688, US7183902, US7164268, US7143392, US8438259, US8681352, US8792110, + 20 pending applications

EDUCATION

Stanford School of Medicine, Stanford, CA
Master of Science (MS), Biomedical Data Science (expected 2025) – focus on MRI image processing.

Stanford University, Stanford, CA
Master of Science (MS), Computer Science, HP Resident Fellow – focus on distributed and parallel systems.

Tufts University, Medford, MA
Bachelor of Science (BS), Electrical Engineering, Summa Cum Laude

Selina Pi

sjpi@stanford.edu
843-327-0542

EDUCATION

Stanford University, Stanford, CA — Sept 2022 – present
Ph.D. Student in Biomedical Data Science
Research interests: Statistical and machine learning methods leveraging observational data to evaluate and improve clinical and health policy decisions
Completed coursework: Modern Applied Statistics: Learning (STATS 315A), Mathematical Models and Medical Decisions (BIOMEDIN 219), Data Science for Medicine (BIOMEDIN 215), Causal Inference and Machine Learning (ECON 293)
Honors: Lusted Award Finalist, Quantitative Methods and Theoretical Developments, 45th Annual Meeting of the Society for Medical Decision Making, October 2023

Princeton University, Princeton, NJ — June 2019
B.S.E. in Operations Research and Financial Engineering — GPA: 3.94
Certificate in Global Health and Health Policy

PROFESSIONAL EXPERIENCE

Analysis Group, New York, NY — Sept 2019 – May 2022
Healthcare Senior Analyst — Jan 2021 – May 2022
Healthcare Analyst — Sept 2019 – Dec 2020

- Worked on health economics and outcomes research projects analyzing insurance claims and chart review data to assess demographics, treatment patterns, healthcare resource use, and clinical and economic outcomes among patients with cancer, gastrointestinal disorders, and rare diseases

RESEARCH EXPERIENCE

Manuscripts

[1] Sanchez L, Chari A, Cheng M, Cherepanov D, DerSarkissian M, Huang F, Stull DM, Dabora J, Young M, Noga SJ, **Pi S**, Zhang M, Banatwala A, Duh MS, Ailawadhi S. Comparison of health care costs and resource utilization for commonly used proteasome inhibitor-immunomodulatory drug-based triplet regimens for the management of patients with relapsed/refractory multiple myeloma in the United States. *J Manag Care Spec Pharm. 2023;29*(10):1-14.

Conference Presentations

[1] **Pi S**, Masterson J, Ma SP, Corbin CK, Milstein A, & Chen JH. Using case mix index within diagnosis-related groups to evaluate variation in hospitalization costs at a large academic medical center. *Oral presentation at AMIA 2023 Annual Symposium (November).*

[2] **Pi S**, Goldhaber-Fiebert JD, & Alarid-Escudero F. A decision-analytic loss function for asymmetrical misclassification costs. *Poster presentation at the 45th Annual Meeting of the Society for Medical Decision Making 2023 (October).*

SKILLS

- Programming languages: SAS, R, MATLAB, Java, SQL, Python
- Data sources: Administrative claims (IBM Marketscan, JMDC, Optum, IQVIA), electronic health records
- Languages: Proficient in French (written/spoken)

Elana Simon

Email Google Scholar LinkedIn

Education

STANFORD UNIVERSITY PhD in Biomedical Informatics, 2022 - present
HARVARD UNIVERSITY B.A. in Computer Science, 2014 - 2018

Publications

Honeyman JN*, **Simon EP***, et.al. Science: 2014 Feb 28; 343(6174) doi: 10.1126/science.1249484. Detection of a Recurrent DNAJB1-PRKACA Chimeric Transcript in Fibrolamellar Hepatocellular Carcinoma. *co-first authors

Simon EP, et.al. Proceedings of the National Academy of Sciences: 2015 Nov 3; 112(44) doi: 10.1073/pnas.1424894112. Transcriptomic Characterization of Fibrolamellar Hepatocellular Carcinoma.

Shin JE, Riesselman AJ, Kollasch AW, McMahon C, **Simon EP**, et.al. Nature Communications: 2021 April 23; 12(2403) doi: 10.1038/s41467-021-22732-w. Protein design and variant prediction using autoregressive generative models.

Spyvee M, Kallenbach JM, Gupta A, **Simon EP**, Grand G. International Patent: Naptho [2,1 -D] Thiazole Derivatives, Compositions thereof and Methods of Treating Disorders.PCT:PCT/US2020/061796. Issued May 27, 2021

Farber BA, Lalazar G, **Simon EP**, et.al. Oncotarget. 2017 Dec 15;9(12):10211-10227. doi: 10.18632/oncotarget.23325. Non coding RNA analysis in fibrolamellar hepatocellular carcinoma.

Ahmad W*, **Simon EP***, Chithrananda S, Rhamsundar B. ELLIS Machine Learning for Molecule Discovery Workshop: 2021 Dec 13. ChemBERTa-2: Towards Chemical Foundation models, *co-first authors

Work Experience

Reverie Labs Senior Machine Learning Engineer, 2018 - present

- Implemented graph convolutional neural networks (GCNNs) and 3D CNNs to model small molecule ADMET properties and protein-ligand binding to design kinase inhibitors
- Used models to prioritize compounds for virtual screens and lead-optimization. Designed and used tools for de-novo molecular generation. Led to patented CDK9 inhibitor
- Optimized and analyzed methods for molecular pre-training with transformers leading to state of the art performance on multiple MoleculeNet benchmarks. Contributed pre-trained models to HuggingFace and gave talk at DeepChem meet-up
- Designed and implemented algorithm for determining ligand's binding mode in protein
- Wrote blog post using UMAP to analyze chemical datasets and detect biases. Performed detailed analysis with dimensionality reduction methods revealing unique properties of molecular embeddings

Harvard Medical School Student Researcher, Marks Lab, 2017

- Implemented autoregressive language models to model proteins
- Predicted phenotypic effects of genetic variation and generated novel protein sequences

Harvard College Computer Science Teaching Fellow, 2016-2018

- CS181 Machine Learning (2018), CSCI P-14200 Intro to Programming (2017), CS51 Abstraction/Design in Computation (2016)

Google Research Google Accelerated Science Intern, 2016

- Predicted bioactivity of molecules with GCNNs
- Discovered biases in experimental assays by analyzing model failure modes

Rockefeller University / Memorial Sloan Kettering Cancer Center Student Researcher, 2012-2016

- Initiated research project to characterize a rare liver cancer and analyze tumor transcriptome
- Discovered new oncogene (DNAJB1-PRKACA chimera) and inspired repurposing of therapeutics / an FDA approved clinical trial

Facebook iOS Development Summer Intern, FBU for Engineering, 2015

Awards / Honors

Gave keynote speech at American Association for Cancer Research in front of 15k cancer researchers

Won Inaugural AACR Young Champion in Cancer Research award, served as National Cancer Research Month Spokesperson

Introduced President Obama at White House Precision Medicine Announcement

Introduced the "Moonshot on Cancer" talks at the Vatican with Joe Biden and Pope Francis

Gave keynote speech at "Innovators in Adolescent and Young Adult Cancer" Case Comprehensive Cancer Center

Jessy Song

Email: jessys@stanford.edu | Phone: +1(415)688-0886 | LinkedIn: @jessy-song-1102

Professional Experience & Internships

06/2022 – today
South San Francisco, CA

AI Engineer/Data Scientist, Frontier Medicines

- Research and develop machine learning models and cheminformatics tools for covalent drug discovery.
- Developed and managed internal library, molecule similarity search tool, contributed to chemical synthesizability and predictive property model development.
- Developed tools are used by scientists across all projects, improving efficiency for molecular data analysis, computational molecule screening and property assessment.

05/2021 – 08/2021
Kitchener, ON

Medical Device Algorithm R&D, Intellijoint Surgical Inc

- Prototyped for a robot integration project with Intellijoint's surgical navigation system and third-party robot to demonstrate robotics-assisted total knee replacement surgery.
- Implemented MATLAB optimization algorithms for robot spatial control and command pose change using camera-tracker system.

01/2020 – 04/2020
Waterloo, ON

Software Engineer, BlackBerry Enterprise - Architecture Team

- Prototyped a web platform user interface with React and JavaScript that provides User Management functionalities.

05/2019 – 08/2019
Waterloo, ON

Software Engineer, PointClickCare - API Architecture Team

- Implemented RESTful APIs using Spring framework in Java to be used by over 100 vendors and product management teams to track client application and activations.

Research Experience

01/2023 – 04/2023
University of Waterloo

Computational Neuroscience Research Student, Bio-Robotics, Artificial Intelligence and Neuroscience (BRAIN) Lab
Supervisor: Dr. Bryan Tripp; BME 499 research course project

- Research project focused on improving MouseNet - a biologically-constrained convolutional neural network that models mouse visual cortex.
- Designed Unity training video dataset that model mouse vision, trained network with contrastive learning and analyzed network representational similarity to mouse brain.

01/2022 – 12/2022
National Research Council Canada & University of Waterloo

Deep Learning Research Assistant, Vision and Image Processing (VIP) Lab
Supervisors: Dr. Ashkan Ebadi and Dr. Alexander Wong

- Researched and developed neural network with prototypical learning for few-shot classification to detect COVID-19 with ultrasound images.
- Implemented the prototypical-convolutional network and conducted experiments in PyTorch to optimize classification network performance, and assessed network explainability with visual analysis tools.
- Crafted manuscripts for conference presentation and journal paper. Discussed results within the COVID-Net initiative team weekly.

Education

09/2023 – toady

M.S, Biomedical Informatics, Stanford University
Part-time via Honors Cooperative Program

09/2018 – 05/2023

B.A.Sc, Honours Biomedical Engineering, University of Waterloo
With Distinction; Neural Engineering Specialization, Computing Option
Relevant Courses: SYDE 599 Special Topics in Deep Learning, SYDE 566 Simulating Neurological Systems, SYDE 552 Computational Neuroscience

Awards and Distinctions

University of Waterloo President's Research Award (2021)
University of Waterloo President's Scholarship of Distinction (2019)

Thodsawit Tiyarattanachai, MD PhD

Master's Student, Biomedical Data Science Program
Stanford University School of Medicine

Email: ttiya@stanford.edu
Phone: (650) 445-2138
LinkedIn: thodsawit

Physician-Scientist specializing in computational methods with 7 years of research experience

EDUCATION

Doctor of Medicine, First Class Honors
Chulalongkorn University Class of 2020

PhD in Clinical Sciences
Chulalongkorn University Class of 2023

AWARDS

Fulbright Thai Graduate Scholarship (2023)

Outstanding Alumni Award, Faculty of Medicine, Chulalongkorn University (2023)

Prince Mahidol Award Youth Program Scholarship (2019)

Winner, Chiang Mai University International Medical Challenge (2016)

Bronze Medal, Thailand Physics Olympiad (2012)

National Outstanding Youth Award, Ministry of Education, Thailand (2011)

Bronze Medal, World Robot Olympiad (2010)

CODING SKILLS

Python, R, Stata, MATLAB, SQL
Deep learning
Machine learning

RESEARCH INTEREST

artificial intelligence
medical imaging, ultrasound
cancer screening and surveillance
cancer prediction models
cancer biomarkers
liver cancer

PROFESSIONAL EXPERIENCE

Physician-Scientist, Chulalongkorn University (2020-2023)

- Developed and deployed AI system for real-time detection and characterization of focal liver lesions during ultrasound exams
- Lead a clinical trial to evaluate applicability of the AI system in assisting operators during ultrasound exams in clinical settings

Visiting Instructor, Radiology, Stanford University (2021-2022)

- Develop motion compensation and quantitative methods for analysis of focal liver lesions in contrast-enhanced ultrasound
- Conducted a multicenter study to evaluate stability over time of Ultrasound LI-RADS Visualization Score

SELECTED PUBLICATIONS

Artificial intelligence assists operators in real-time detection of focal liver lesions during ultrasound: A randomized controlled study.
Tiyarattanachai T, Apiparakoon T, Chaichuen O, et al.
Eur J Radiol. 2023;165:110932. doi:10.1016/j.ejrad.2023.110932

The feasibility to use artificial intelligence to aid detecting focal liver lesions in real-time ultrasound: a preliminary study based on videos.
Tiyarattanachai T, Apiparakoon T, Marukatat S, et al.
Sci Rep. 2022;12(1):7749. doi: 10.1038/s41598-022-11506-z.

Development and validation of artificial intelligence to detect and diagnose liver lesions from ultrasound images.
Tiyarattanachai T, Apiparakoon T, Marukatat S, et al.
PLOS ONE. 2021;16(6):e0252882.

Multicenter Study of ACR Ultrasound LI-RADS Visualization Scores on Serial Examinations: Implications for Surveillance Strategies.
Tiyarattanachai T, Fetzer DT, Kamaya A.
Am J Roentgenol. 2022;219(3):445-452. doi: 10.2214/AJR.22.27405

Ultrasound Liver Imaging Reporting and Data System (US LI-RADS) Visualization Score: a reliability analysis on inter-reader agreement.
Tiyarattanachai T, Bird KN, Lo EC, et al.
Abdom Radiol (NY). 2021. doi: 10.1007/s00261-021-03067-y.

A Comprehensive Motion Compensation Method for In-Plane and Out-of-Plane Motion in Dynamic Contrast-Enhanced Ultrasound of Focal Liver Lesions.
Tiyarattanachai T, Turco S, Eisenbrey JR, et al.
Ultrasound Med Biol. 2022;48(11):2217-2228.

SHAI **WALDRIP**, PHD

Palo Alto, CA | swaldrip@stanford.edu

SUMMARY OF QUALIFICATIONS

- Health equity champion dedicated to reducing the global cancer burden through research
- Committed to bridging the gap in underserved and underrepresented communities through a clinical informatics approach

EDUCATION AND TRAINING

2023-2025 **Stanford University School of Medicine:** Stanford, CA
Postdoctoral Scholar: Biomedical Informatics

2019-2022 **Morehouse School of Medicine:** Atlanta, GA
Doctor of Philosophy: Biomedical Sciences (Oncology)

2014-2018 **Georgia Gwinnett College:** Lawrenceville, GA
Bachelor of Science: Biology (Cell Biology & Biotechnology)

EXPERIENCE

2019-2022 **Dissertation Research:** *Morehouse School of Medicine* (Atlanta, GA)

- Utilized biological and bioinformatics tools to validate potential mechanism-based targets for the early detection and personalized therapy for BRCA1-associated triple-negative breast cancer (TNBC):
  - Analyzed next-generation sequencing (mRNA sequencing) data from wildtype and BRCA1 mutant TNBC cells to identify significant differentially expressed transcripts
  - Discovered several potential downstream targets that could be used as druggable targets for patients with BRCA1 mutant TNBC

2018 **Summer Intensive Research Initiative Biomedical Science Intern:** *Morehouse School of Medicine* (Atlanta, GA)

- Performed cell culturing, brain slicing, and immunohistochemistry with the goal of attenuating Cerebral Malaria (CM) using Neuregulin-1 (NRG-1)
- Examined 15 post-mortem CM-infected brain stem, cerebral cortex, and cerebellum patient samples via microscopy with the purpose of observing *Plasmodium falciparum* sequestration
- Utilized ImageJ to measure the intensity of NRG-1 expression in post-mortem CM-infected brains compared with the brains of those who died of non-malarial causes
- Concluded that NRG-1 could serve as a biomarker for assessing prognosis of CM

2018 **Bioinformatics Research Project:** *Georgia Gwinnett College* (Lawrenceville, GA)

- Independently explored the connection between Aggregative Adhesion Fimbriae Type I (AAF/1) and Invasin in various bacterial species using MEGA7 and the NCBI website to discover if 1) any proteins played a similar role to AAF/1, and 2) what bacterial species those proteins belonged to
- Resulted in identification of biofilm formation regulator HmsP *Edwardsiella hoshinae* and *Edwardsiella anguillarum*

2018 **Biotechnology Lab Project:** *Georgia Gwinnett College* (Lawrenceville, GA)

- Analyzed possible genetic modification of non-genetically modified organism (GMO) labeled soy-based protein powders using DNA extraction, Nanodrop 2000, Polymerase Chain Reaction, Agarose gel electrophoresis, and Enzyme-Linked Immunosorbent Assay (ELISA)
- Determined no genetic modification in Sprouts Farmers Market Natural Vanilla soy protein powder or Jarrow Formulas Iso-Rich soy dietary supplement samples, therefore validating non-GMO labeling

LEADERSHIP

- Morehouse College Fall 2022 STEM Research Symposium Judge (2022)
- Annual Biomedical Research Conference for Minoritized Scientists Judge (2022)
- Co-Creator and Co-Facilitator for MozFest 2022 (2022)
- Stanford ICME-AUC Biomedical Data Science Summer Research Program (2022)
- Founder of BODEGA Biomedical Data Science Summer Research Program (2022)
- Vice President and Co-Ambassador of the Bioethics Club (2021-2022)
- President of the Atlanta University Center Data Science Club (2021-2022)
- Vice President of the Atlanta University Center Data Science Club (2021)
- Graduate Education in Biomedical Sciences Peer Mentoring Program (2020-2021)

PUBLICATIONS

- Lama Nazer, Razan Zatara, **Shai Waldrip**, Janny Xue Chen Ke, Mira Moukheiber, Ashish Khanna, Rachel Hicklen, Lama Moukheiber, Dana Moukheiber, Haobo Ma, and Piyush Mathur. Bias in Artificial Intelligence Algorithms and Recommendations for Mitigation. PLOS Digital Health 2023.
- **Shai Waldrip**, Jingyao Xu, Yunlong Qin, Danita Eatman, E Shyam P Reddy, and Veena N Rao. A Systems Biology Approach to Validate Potential Targets for Early Detection and Personalized Therapy for BRCA1-Associated TNBC. Open Acc J Oncol Med 5(3)- 2022.
- Jingyao Xu, Zerak Kabir, Kartik Aysola, Yuli Chai, Nina Wyatt, **Shai Waldrip**, Alexis Clark, Michelle Lee, Vaishali Reddy, Manan Shah, Eric Chang, Joel Okoli, E Shyam P Reddy, and Veena N Rao. p21 WAF1/CIP1 is a Downstream target of ELK-1 Growth / Tumor Suppressor Pathway in Breast and Androgen-independent Prostate Cancers. Open Acc J Oncol Med 4(5)- 2021.

CHRISTINE YIWEN YEH

+1(626) 202-6402 ◇ Stanford, CA
cyyeh@stanford.edu ◇ linkedin.com/in/christine-yiwen-yeh

EDUCATION

MD/PhD Biomedical Data Science, Stanford University 2020 -

- ***Thesis:*** *Investigating spatiotemporal dynamics of solid-tumor cancer tissues: an integrated data-driven approach*
- ***Co-advisors:*** *Livnat Jerby, Sylvia Plevritis*

MS Biomedical Informatics, Stanford University 2016 - 2017
BSH Molecular, Cellular, and Developmental Biology (Honors), Stanford University 2012 - 2016

PROFESSIONAL EXPERIENCE

Computational Biologist (Associate) Jul 2017 - Aug 2020
D. E. Shaw Research *New York, NY*

- Lead the computational biology arm of several proprietary drug discovery programs including the development of a SHP-2 inhibitor for advanced/metastatic solid tumors; in phase 1 clinical trials (NCT05487235)
- Developed and deployed machine learning algorithms for novel computational biochemistry analyses
- Deconvolved the catalytic mechanism of a well-validated type II diabetes protein using molecular dynamics simulations and machine learning leading to 2 first (co)-author publications

Bioinformatics Research Intern Jun 2016 - Sep 2016
Genentech *South San Francisco, CA*

- Discovered the functional and biochemical functions of protein arginine methyltransferases in the context of various cancers and complex neurological disorders
- Built computational analysis infrastructure for mass-spectrometry data in an early drug discovery pipeline

FEATURED PUBLICATIONS

Yeh, CY*; Aguirre, K*; Laveroni, O*; Kim, S; Wang, A; Liang, B; Zhang, X; Han, LM; Valbuena, R; Plevritis, SK; Bassik, MC; Snyder, MP; Howitt, BE; Jerby, L. Mapping ovarian cancer spatial organization uncovers immune evasion drivers at the genetic, cellular, and tissue level. ***bioRxiv*** (2023) *Equal Contribution*

Yeh, CY; Izaguirre, J; Greisman, J; Willmore, L; Maragakis, P; Shaw, DE. A conserved motif controls the kinetics of the Protein Tyrosine Phosphatase 1B catalytic cycle. **Journal of Chemical Information and Modeling** (2023)

Yeh, CY*; Adusumilli, R*; Kullolli, M; Mallick, P; John EM; Pitteri, SJ Assessing biological and technological variability in protein levels measured in pre-diagnostic plasma samples of women with breast cancer. **Biomarker Research** (2017) *Equal Contribution*

SELECT AWARDS

Stanford Graduate Fellowship, Stanford Vice Provost of Graduate Education 2020 - 2025

- *5-year full support for outstanding students pursuing a doctorate in science and engineering*

PhD Abstract Competition Finalist, Human Proteome Organization World Congress 2017
Firestone Medal for Excellence in Research, Stanford Vice Provost of Undergraduate Education 2016

SKILLS

Python, R, Bash, Slurm, Git, Cancer Biology, (Spatial) multi-omics, Translational Bioinformatics, Machine Learning, Drug Discovery, Medicine - Clinical Oncology, Biomarker Discovery, Teaching

JULIE ZHANG

736 Serra Street, Stanford, CA, 94305
(425)-772-1308 ⋄ jyz27@stanford.edu

EDUCATION

Stanford University *Sept. 2020 - Present*

- Ph.D. student in Statistics; NSF GRFP and Enhancing Diversity in Graduate Education Fellowship

University of Washington *Sept. 2015 - June 2020*

- Bachelors of Science in Statistics and Mathematics with Honors, Magna Cum Laude
- UW College of Arts and Sciences Dean's Medalist in the Natural Sciences

Robinson Center Transition School *Sept. 2014 - June 2015*

PUBLICATIONS

- **Zhang, J.**, Preising, G. A., Schumer, M., and Palacios, J. A. CRP-Tree: A phylogenetic association test for binary traits. *Journal of the Royal Statistical Society Series C: Applied Statistics* (2023). https://doi.org/10.1093/jrsssc/qlad098
- Ouedraogo, A.L., **Zhang, J.**, Tinto H., Valéa I., and Wenger, E.A. A microplanning model to improve door-to-door health service delivery: the case of Seasonal Malaria Chemoprevention in Sub-Saharan African villages. *BMC Health Serv Res* **20**, 1128 (2020). https://doi.org/10.1186/s12913-020-05972-2

RESEARCH AND WORK EXPERIENCE

Bill and Melinda Gates Foundation Seattle, WA
Research Scientist - Institute for Disease Modeling *May 2023 - Present; Jan. - Aug. 2022*

AbbVie North Chicago, IL
Data and Statistical Sciences Experiential Intern, Immunology *June - Sept. 2023*

American Institute for Mathematics
Dynamics and Data in the COVID-19 Pandemic Virtual Workshop *June 2020 - Aug. 2020*

Institute for Disease Modeling Bellevue, WA
Research Intern - Malaria Team *June 2019 - Sep. 2019*

Institute for Pure and Applied Mathematics Los Angeles, CA
Research Intern - The Aerospace Corporation *June 2018 - Aug. 2018*

TA EXPERIENCE

- Stanford University: DATASCI 192A/B, STATS 202, STATS 116 (twice), CS 229T, STATS 191
- University of Washington: Math 134/5/6: Honors Accelerated Calculus (twice)

CONFERENCES AND PRESENTATIONS

- **"CRP-Tree: A phylogenetic association test for binary traits."**: IMS Annual Meeting (06/2022), International Conference on Statistics and Data Science (12/2022), Probabilistic Modeling in Genomics (03/2023)
- **"Comparing Object Correlation Metrics for Effective Space Traffic Management."**: Joint Mathematics Meetings (01/2019), Nebraska Conference for Undergraduate Women in Mathematics (01/2019), Statistics and Data Science Symposium (05/2019)

ADDITIONAL SKILLS

- Proficient in R, Python, LaTeX, Git
- Basic SQL, Matlab, Java, ArcMap
- Native English, proficient Mandarin Chinese