# Varenya Jain

646-306-1672 | varenyajj@gmail.com | Linkedin | Github | Website

## Education

**University of Illinois Urbana-Champaign** 2020-2024
*Bachelor of Science in Integrative Biology, Minor in Computational Science and Engineering* *Urbana, IL*
- *Dean's List (Spring 2024)*

## Experience

**Research Assistant** | **Robinson Lab** Feb. 2025 - Present
*Rahel Hirsch Center fr Translationale Medizin, Berlin Institute of Health Charit* *Berlin, Germany*
- Develop a federated genotype-phenotype database for the Fetal Sequencing Consortium to study Rare Mendelian diseases across Columbia University, The Broad Institute, and the University of New South Wales
- Train large language models on Charité servers to construct machine-readable, ontology-aware phenotype representations
- Write software to transform raw prenatal sonography data into standardized GA4GH Phenopackets

**Contract Bioinformatician** Dec. 2022 - Present
*Advanced Genomics Institute and Laboratory Medicine* *Paramus, NJ*
- Optimize Illumina ICA and BaseSpace workflows, adding SFTP integration to streamline WGS/WES analyses
- Perform targeted enrichment, structural variant detection, and CNV calling to improve clinical interpretation of germline samples
- Utilize Franklin by Genoox databse to prioritize causal variants for clinical reporting

**Student Researcher** Aug. 2023 - Dec. 2024
*Illinois Natural History Survey - Tan Lab of Biodiversity Genomics* *Urbana, IL*
- **Morphological Diversity Project (2023):**
  - Researched adaptation in Sisoridae and Amphilidae catfishes within fast-water environments.
  - Presented findings showing significant shape disparity linked to adhesive organ presence, providing insights into evolutionary constraints on morphology.
- **Phylogenomics Project (2024):**
  - Investigated ultraconserved elements (UCEs) in zebrafish genomes using low-coverage whole-genome sequencing (lc-WGS) to study evolutionary relationships.
  - Built and optimized phylogenomic assembly workflows using QC utilities, De Novo Assembly, and the CAPTUS toolkit.

**Student Researcher** | **Metabolomics & Proteomics Core Facilities** June 2023 – Aug. 2023
*Roy J. Carver Biotechnology Center* *Urbana, IL*
- Developed an untargeted metabolomics pipeline for post-processing, Quality Assurance/Quality Control, and data analysis.
- Supported for data processing on LC-MS Untargeted Metabolite Profiling.

**SLC Conference Planning Committee** | **IEEE** Jan 2022 - Jan 2023
*The Institute of Electrical and Electronics Engineers* *Chicago, IL*
- Managed conference logistics, scheduling, and coordinating with stakeholders.
- Worked with Midwest Region 4 Committee members to develop and implement strategies for a successful conference.

**SPIN Research Intern** | **The NEAT Project v4.0** Aug. 2021 – Sep. 2022
*National Center for Supercomputing Applications* *Urbana, IL*
- Developed an NGS toolkit on the HAL cluster, improving parallel processing performance by 7%.
- Implemented alignment algorithms (Smith-Waterman, BLAST, localized string alignment) to enhance mutation and sequencing model pipelines.
- Presented research at the NCSA Exhibition, Engineering Open House, and FoDOMMaT/SPIN Showcase.

**Outreach Committee Lead** | **Pulse 2022** Aug 2021 - Aug 2022
*Department of Electrical and Computer Engineering* *Urbana, IL*
- Coordinated materials and event location logistics with the ECE department during COVID-19.
- Develop Software to exhibit Computer Engineering principles: C++ Data Structures, Polymorphism, Command-Line Interface, Stack/Heap Memory Management, Address Space, etc.
- Design Hardware activities to guide freshmen through simulated Electrical Engineering projects: series vs parallel circuits, Pulse Width Modulation motor control, Thermistor and LDR implementation

**Phys 211 Experienced Learning Assistant** Jan 2021 - Dec 2021
*Loomis Laboratory Of Physics* *Urbana, IL*

- Continued the study of Physics pedagogy by instructing PHYS 211 labs, providing guidance to 30+ students.
- Answered student questions, clarified critical lecture material, and built/corrected IOLab setups.
- Developed advanced teaching, communication, and leadership skills through interactions with students and seminar presentations alongside lab staff.

**IOT Research Lab Assistant** | **Caesar Lab** Jan. 2021 – Jun. 2021
*Coordinated Science Laboratory* *Urbana, IL*

- Remodeled a Reinforcement Learning System intended for UAV-Assisted Emergency Response.
- Upgraded the fully-distributed communication environment for USAF usage.
- Implemented communication trees via Python Message Passing Interface (MPI) standard in under 6 months.

**Research Intern** | **Drs. Spitalnik, Hod, La Carpia** June 2018 - Aug 2018
*Columbia University Medical Center - Lab of Transfusion Biology* *New York, NY*

- Investigated the effects of transfusional iron overload on gut microbiota due to intravenous infusion.
- Conducted initial studies with a mouse model to retrieve data on iron-deficient erythropoiesis in blood donors and red blood cell recovery after transfusion of hematopoietic red blood transplant.
- Utilized basic Spearman Correlation meta-analysis of bacteria communities, performed blood analysis tests, and used Flow Cytometry to collect sample cell data.

## Projects

**Phenopacket Processor** | ***Personal*** July. 2025 – Aug. 2025

- Construct extensible CLI for working with clinical data and the Human Phenotype Ontology
- Convert existing digital phenotypic records into GA4GH Phenopackets
- Enable genotype-phenotype database curation by partners at the University of New South Wales

**BadgeDev** | *SIGPWNY* Aug. 2024 – Dec 2024

- Aid in software development for FallCTF participant, sponsor, and staff badges
- Assemble and perform QC on Pico soldering and test firmware before distribution
- Fix and replace badge firmware/hardware during event

**AM Radio** | *ECE 210* Jan. 2022 – May 2022

- Construct a functioning AM Radio using a Superheterodyne Receiver; Convert digital .wav audio input to analog 3.5mm speaker output
- Utilize Fourier Transforms to convert Time domain signals to Frequency domain responses

**Virtual Gloves** | ***ECE 120 Honors*** Jan. 2021 – May 2021

- Develop a "virtual keyboard" by moving fingers attached to flex sensors and provide haptic feedback
- Collect data from flex resistors in Arduino Studio and use C++ to send output confirmations to LEDs

**LoopKit** | ***Personal*** July. 2020 – Sep. 2021

- Adapt the LoopKit source code to customize a program for personal monitoring functionalities
- Modify Swift code to circumvent authentication requisites for bolus delivery within the Loop system
- Implement precision adjustments to default carb absorption parameters, optimizing glycemic control dynamics

## Leadership and Involvements

**Conferences**

- NCSA Letter of Outstanding Student Leadership Recognition (2023-2024)
- NCSA 2nd Annual Student Research Conference - Planning Committee, Panel Moderator, and Industry Chair (2024)
- ACMG Annual Clinical Genetics Meeting (Metro Toronto Convention Center), 2024
- ISPD 28th International Conference on Prenatal Diagnosis and Therapy (Westin Copley Place, Boston), 2024
- NCSA Student Research Conference - Planning Committee and Panel Moderator (2023)
- IEEE Nexus Region 4 Student Leadership Conference - Conference Planning Lead (2022-2023)

**Presentations**

- University of Illinois - Undergraduate Research Symposium (2024)
- National Center for Supercomputing Applications - SPIN Lightning Talks (2021-2022)
- National Center for Supercomputing Applications - Engineering Open House (2021)

**Teaching**

- ECE PULSE - Outreach Committee Lead (2022)
- Loomis Laboratory Of Physics - "Expert Learning Assistant" for Mechanics (2021)

**Volunteering**

- SIGPWNY - Contributed to solving cybersecurity CTF ("Capture the Flag") competitions and developed badge software for FallCTF
- De Dilse Charitable Inc. - Co-Founder and Lead Volunteer (2019-2022)
- Kaplen Jewish Community Center on the Palisades - Senior Science Counselor (2019)
- American Wheat Mission - Volunteer (2018)

**Student Organization Membership**: IEEE@UIUC, iRobotics, MRDC, Vex Robotics, ACM

## Technical Skills

**Wet Lab Experience**: Flow Cytometry, PCR, RT-qPCR, Western blotting, recombinant DNA techniques, cell culture (bacterial and mammalian), microscopy (fluorescence, confocal), ELISA, protein purification, CRISPR/Cas9 gene editing, histology, and immunohistochemistry.

**Bioinformatics / Computational Biology**: OMIM, UCSC Genome Browser, Ensembl, Illumina Connected Analytics, BaseSpace SequenceHub (DRAGEN), GATK, VCFtools, STAR aligner, BLAST+, Smith-Waterman, Chromosome Analysis Suite (ChAS). Familiar with RNA-seq and single-cell RNA-seq data analysis, differential expression and pathway enrichment analysis, and genome assembly and annotation.

**Biostatistics / Data Analysis**: Experience with R (statistical modeling, ANOVA, regression, PCA, clustering, and Generalized Procrustes Analysis [GPA]) and Python (NumPy, pandas, seaborn, and scikit-learn). Familiarity with experimental design and power analysis, and applied statistical interpretation of biological data.

**Programming / Computational Skills**: Exposure to C/C++, CUDA, Python (including Biopython and data analysis libraries), R, and MATLAB. Experience with data wrangling and visualization (pandas, ggplot2) and applying machine learning methods to biological datasets (scikit-learn, TensorFlow). Proficient with developer tools such as Git, VirtualBox, KVMs, VSCode, and Pycharm.

**Developer Tools**: Git, GitHub, GitLab, Docker, VirtualBox, Kernel-based Virtual Machines (KVMs), Digital Ocean Droplets, VSCode, IntelliJ

**Systems / Hardware and Electronics**: Experience with breadboards, Raspberry Pi (Zero/3b/4), Arduino UNO, and Intel MAX 10 FPGA kits. Familiarity with oscilloscopes, network analyzers, RLC circuits, operational amplifiers, band-pass filters, Quartus Prime, ModelSim, and SystemVerilog.

**GNU/Linux**: Debian, Fedora, CentOS, AlmaLinux, Arch Linux, Ubuntu Server, EndeavorOS

**Research / Professional Practices**: Scientific writing and data presentation; familiarity with reproducible research practices; knowledge of FAIR data principles; collaboration in interdisciplinary teams; literature review and hypothesis generation; AGILE software development and project management.