

ANDREW J. BONNEY

636-541-8249 | abonney@wustl.edu | <https://github.com/Bonney96> | <https://www.linkedin.com/in/andrew-bonney/>

BIOINFORMATICS SCIENTIST | MULTI-OMICS & AI INTEGRATION

Bioinformatics Scientist with an M.S. and 3+ years of experience integrating multi-omic data and developing reproducible computational pipelines. expert in Nextflow DSL2, Docker, and Python for harmonizing and QCing complex datasets (RNA-Seq, WGBS, Hi-C). Unique hybrid background combining industrial rigor (Pfizer/Thermo Fisher) with academic innovation, leveraging Computer Vision and Deep Learning to solve biological problems. Currently managing scalable genomic architectures on the RIS/Compute1/2 clusters at Washington University School of Medicine.

TECHNICAL SKILLS

- **Languages:** Python (Advanced), R (Advanced), Bash, Rust, SQL, Java
- **Workflow & Cloud:** Nextflow DSL2, Docker/Singularity, AWS (Batch, S3, EC2), Git/GitHub, Linux/HPC Administration
- **Bioinformatics Tools:** Autodock, GROMACS, GATK, STAR, Bowtie2, SAMtools, BEDtools, Seurat, DESeq2, EdgeR, Bioconductor, Peakachu, Bismark, MACS, HOMER, BLAST, BioPython, BioPerl
- **Data Science & AI:** PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, Explainable AI (SHAP, Captum), Computer Vision (CNNs)
- **Regulatory & QC:** ISO 9001, FDA Validation (21 CFR Part 11), GLP/GMP, Data Integrity & Audit Trails

EXPERIENCE

Bioinformatics Research Assistant, Spencer Lab - Acute Myeloid Leukemia (AML) & Epigenomics Aug 2024 - Present
Washington University School of Medicine St. Louis, MO

- Engineered reproducible, containerized Nextflow pipelines to process large-scale multi-omic datasets (RNA-seq, WGBS, Hi-C), ensuring metadata consistency across patient-derived AML cohorts.
- Developed a modular Deep Learning framework (CNN/PyTorch) to predict DNA methylation states from genomic sequences, integrating meta-data features to improve model generalizability.
- Utilized Peakachu, a supervised loop-calling algorithm, to analyze Hi-C contact maps in AML samples, identifying chromatin loops associated with enhancer-promoter interactions.
- Contributed to targeted methylation and epigenome editing projects (CHARM, CRISPRi, SCGE), integrating CUT&Tag, WGBS, and Hi-C data to identify regulatory elements in AML.
- Orchestrated automated Python/R ETL scripts on AWS Batch and S3 to ingest and harmonize petabyte-scale datasets, optimizing storage costs and compute efficiency.

Analytical R&D Data Scientist (Internship) - Computer Vision & Automation Focus May 2024 - Aug 2024
Pfizer Chesterfield, MO

- Developed and validated automated computer vision algorithms to process high-content microscopy data, successfully extracting phenotypic features from primary fibroblast cell cultures.
- Replaced manual cell counting methods with reproducible R/Python scripts, increasing data analysis throughput by 300% while eliminating inter-operator variability.
- Performed rigorous statistical evaluation of potency assay data, reducing analysis error rates by 40% and establishing robust QC metrics for image-based datasets.
- Partnered with wet-lab biologists to translate complex visual phenotypes into quantifiable digital standards, ensuring assay reproducibility under strict R&D guidelines.

Scientist II - Biochemistry & Instrumental Analysis Jul 2021 - Jan 2024
ThermoFisher Scientific St. Louis, MO

- Wrote custom Python parsers to automate the ingestion of raw instrument data (XML/CSV) from Capillary Gel Electrophoresis (CGE) and PCR systems, structuring unstructured outputs for downstream analysis
- Served as Technical Lead for analytical biochemistry assays, managing the full lifecycle of method transfer and validation for external CDMO clients.
- Executed longitudinal degradation studies (ICH guidelines), utilizing statistical modeling to track analyte stability and predict formulation performance over time.
- Maintained 99.8% data accuracy across 200+ monthly samples by implementing rigorous statistical review processes for complex analytical datasets.

QC Scientist - Analytical Chemistry

May 2021 - Aug 2021

Firmenich

St. Louis, MO

- Processed and verified high-volume GC-MS spectral data for 150+ compounds, maintaining strict data integrity standards for analytical chemistry workflows.

Genetic Quality Assurance Lab Technician

Aug 2020 - May 2021

Bayer

Creve Coeur, MO

- Performed high-throughput DNA extractions and utilized Biomek automation systems to process 500+ samples weekly in an ISO 9001 compliant quality lab.

EDUCATION

Master of Science - Bioinformatics and Computational Biology

Dec 2022 - Dec 2024

Saint Louis University, St. Louis, MO - GPA: 3.97/4.00

- **Graduate Research (Neural Networks):** Applied deep learning to classify structural RNA motifs involved in gene regulation.
- **Graduate Research (Molecular Dynamics):** Conducted molecular docking experiments to model Aripiprazole-D2 Dopamine Receptor interactions and analyze binding site kinetics.
- **Graduate Research (Benchmarking):** Evaluated the "methylIT" pipeline for differential methylation detection in *Oryza sativa*, benchmarking accuracy against standard tools.

Bachelor of Science - Biology University of Missouri-Kansas City

Aug 2015 - May 2019

SELECT PROJECTS & PORTFOLIO

nf-core/scge Cloud-Native Tumor-Normal Pipeline

- I co-developed a custom scalable Nextflow pipeline for the NIH Common Fund's Somatic Cell Genome Editing program, adhering to strict nf-core reproducibility standards.
- Integrated Illumina DRAGEN hardware acceleration on AWS Batch to ingest and process high-depth sequencing data with minimized runtime.

DeepLabCut Imaging Pipeline

- Built a custom machine learning-based pose estimation system to track zebrafish behavior in a custom experimental rig, enabling high-throughput behavioral analysis.

BERT for Sequence Classification

- Leveraged transformer-based language models to classify biological sequences, exploring applications in gene function prediction and regulatory element annotation.

EpiBench: Deep Learning for Methylation

- Created a Python-based framework for predicting DNA methylation, utilizing Pandas for large-scale data manipulation and PyTorch for model training.

PROFESSIONAL AFFILIATIONS

- International Society for Computational Biology (ISCB) Member Since Aug 2024
- SLU Institute for Drug and Biotherapeutic Innovation Member Since Feb 2024