

Brett Van Tassel

brettt.vantassel@gmail.com ❖ (609) 575-7711 ❖ Flagstaff, AZ ❖ [GitHub link](#)

WORK EXPERIENCE

The Translational Genomics Research Institute (North)

January 2021 – January 2025

Bioinformatician

Flagstaff, AZ

- **Sequencing & ETL Pipeline Development:** Built robust sequencing pipelines for COVID-19 and Group A Strep using Python, Pandas, GeoPandas, Plotly, R, and shell scripting, with GitHub for version control. Designed and maintained a full ETL pipeline to extract raw data from sequencing instruments and internal databases, transform it through demultiplexing and alignment-to-consensus conversion using ASAP (BBDuk, Bowtie2), and convert XML outputs to HTML via XSLT. Parsed results with BeautifulSoup and ElementTree into structured pandas dataframes, normalized metadata formats, and loaded processed data into a PostgreSQL database. Enabled streamlined analysis, automated Excel report generation, and integration into public-facing interactive dashboards.
- **Data Visualization:** Developed automated, interactive dashboards to facilitate real-time visualization of sequencing data. Initially managed interactive dashboards in Tableau after self-learning the platform, then migrated them to Plotly to eliminate licensing costs while preserving functionality. Dashboards enable real-time visualization of sequencing data (<https://pathogen-intelligence.tgen.org/covidseq-tracker/>)..
- **Metagenomic Microbiome Analysis:** Collaborated closely with research partners to guide analysis strategy for metagenomic microbiome projects, conducting data validation, visualization, and statistical analyses using QIIME2, LEfSe, PICRUSt2, and more while ensuring PHI protection and extracting actionable insights from complex datasets.
- **Version Control and CI/CD Pipelines:** Managed GitHub repositories with strong version control, leveraging GitHub Actions for CI/CD workflows and cross-repository automation via dispatch events.
- **High-Performance Computing Expertise:** Extensive experience running large-scale bioinformatics workflows on SLURM clusters, with automation via cronjobs and parallelization using Parsl for efficient Python-based scripting.
- **Software Dev and QIIME2 Contributions:** Passionate about software design, contributing to the QIIME2 ecosystem by developing Python-based plugins that integrate advanced trimming and alignment tools to enhance sequencing preprocessing. Built a QIIME2 wrapper for the ASAP, storing visualizations as artifacts and streamlining internal workflows for the public health team to identify COVID variants and potential outbreaks.
- **Bioinformatics Support Across Projects:** Provided comprehensive bioinformatics support to multiple teams, assisting with data analysis, pipeline development, and troubleshooting across diverse research projects, ensuring timely and accurate results to advance scientific objectives.
- **Documentation:** Authored highly detailed, user-friendly documentation using Docusaurus and React in TGen Knowledge base.

EDUCATION

Ramapo College of New Jersey

January 2024

Master of Science, Data Science

Mahwah, NJ

- Thesis Title: Examining Disease through Microbiome Data Analysis

Ramapo College of New Jersey

January 2020

Bachelor of Science, Bioinformatics

Mahwah, NJ

- Mathematics Minor

SKILLS & INTERESTS

- **Skills:** Python, AWS EC2 and S3, Airflow, ETL Data Management, R, bash, Linux, Machine Learning, Statistics, Data Visualization, React, HPC, SLURM, Version Control (Git/GitHub), CI/CD, QIIME2, Docker, Metagenomics, Flask, MySQL, PostgreSQL.
- **Interests:** Logic Puzzles; Topological Puzzles; Hiking; Guitar and Singing; LLMs