Last Updated: 28 October 2024
Resume available here
CV available here

# TYLER HEIST, PH.D.

email@tylerheist.com · +1 (859) 815-9234

I'm Tyler, a data engineer and scientist. My background is largely in healthcare and microscopy data, but I am always interested in gaining expertise in new domains. I have a passion for enabling data-driven decision-making, data quality, reproducible research, and scientific communication.

## EXPERIENCE

**MAY 2023 – PRESENT**

**SENIOR DATA ENGINEER, JOYN INSURANCE,** REMOTE

- Architected a re-design of Joyn's entire data ecosystem– moving from a series of tightly-coupled serverless functions to a modern data stack approach using Apache Airflow, DBT, DLT, Fivetran, and Snowflake, decreasing downtime by >90% and increasing confidence in data across the organization.
- Implemented reverse ELT integrations to clients (e.g., insurance brokers), vendors (e.g., CRM, Billing), and regulatory reporting agencies (e.g., local and state insurance offices), ensuring compliance and saving tens of human-hours weekly while providing high-quality data.
- Mentored data analyst in writing high quality SQL and enabled them to rapidly write their own Apache Airflow DAGs without prior experience in Python via paired work and templates for internal reporting.

**OCTOBER 2021 – APRIL 2023**

**DATA ENGINEER, DATA PRODUCT MANAGER, DATA SCIENTIST, CEREBRAL**, REMOTE

- Architected and implemented data infrastructure to enable an NLP model to quickly detect and surface patient messages in Cerebral's Electronic Medical Record (EMR) that contained likely suicidal ideation, allowing over 4,000 patients to receive aid from crisis specialists in under 10 minutes (baseline >12 hours).
- Designed and managed a clinical decision-making data product to expose data-driven APIs for use by Cerebral's EMR that assisted over 500 clinicians to provide optimal treatments to over 10,000 patients.
- Designed and implemented robust ELT processes ingesting data from multiple, disparate sources (e.g., Electronic Health Record (EHR), marketing, user behavior data, messaging) using a variety of processes (e.g., custom, Fivetran, Stitch, Apache Kafka, DBT) orchestrated in Apache Airflow.
- Led data quality efforts for the data team, which included implementing various tools like DBT (data modeling), Monte Carlo (data observability), and Atlan (data catalog) as well as creating and organizing twice-quarterly one-day hackathons to address data quality and infrastructure gaps, resulting in a reduction in up to 75% of tech debt tickets per quarter.
- Mentored data analysts and data scientists in software development best practices, such as version control, documentation, and testing.

**SEPTEMBER 2019 – SEPTEMBER 2021**

**DATA SCIENTIST, EPIC SYSTEMS**, MADISON WI

- Designed, led, and executed research using Electronic Health Record (EHR) data to provide insights into the COVID-19 patient population, with a focus on comorbidities and clinical outcomes, collaborating with several external groups (e.g., FDA COVID-19 evidence accelerator, Kaiser Family Foundation) in addition to researchers at academic institutions (e.g., University of Chicago).
- Designed, trained, and validated predictive models for clinical and operational use in Epic's EHR system.

## EDUCATION

**SEPTEMBER 2015 – SEPTEMBER 2019**

**PH.D. QUANTITATIVE AND COMPUTATIONAL BIOLOGY,** PRINCETON UNIVERSITY **GPA: 3.94**

**AUGUST 2011 – MAY 2015**

**B.S. BIOLOGY (HONORS) AND COMPUTER SCIENCE,** UNIVERSITY OF RICHMOND **GPA: 3.92**

## SKILLS

**Python** (*pandas, dlt,* etc.); **R** (*tidyverse, tidymodels,* etc.); **SQL** (*Postgres, Snowflake, MySQL*); **NoSQL** (*MongoDB*); **Streaming** (*Apache Kafka, Red panda*); **Statistics**; **Machine Learning**; **Containerization** (*Docker, Kubernetes*); **Version control** (*git, subversion*); **Workflow orchestration** (*Apache Airflow*); **CI/CD** (*Github Actions, Azure DevOps*); **Infrastructure as Code** (*Terraform*); **Cloud Computing** (*AWS, Azure*); **Data modeling** (*DBT*); **Data Observability** (*Monte Carlo*); **Shell scripting** (*Bash, Zsh)*; **BI Tools** (*Looker, PowerBI*); **Mentorship**