Last Updated: 6 March 2023
Current Copy 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 data quality, reproducible research, scientific communication, and finding novel solutions to real-world problems using machine learning.

EXPERIENCE

OCTOBER 2021 – FEBRUARY 2023
CEREBRAL, SAN FRANCISCO CA (REMOTE)
DATA ENGINEER (10/22 – 2/23), DATA PRODUCT MANAGER (3/22 – 12/22), DATA SCIENTIST (10/21 – 9/22)

- 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 clinicial decision-making data product to expose an API 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 transformations), 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 git/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 (e.g., Inpatient Deterioration Index) for clinical and operational use in Epic's EHR system.

MAY 2016 – SEPTEMBER 2019
GRADUATE STUDENT, LEVINE LAB, PRINCETON NJ

- Investigated how transcription occurs across large genomic distances during early development in *Drosophila* embryos.
- Developed an image processing pipeline to segment nuclei and identify transcriptional foci from traditional confocal and super-resolution microscopy.

EDUCATION

SEPTEMBER 2015 – SEPTEMBER 2019
PH.D. QUANTITATIVE AND COMPUTATIONAL BIOLOGY, PRINCETON UNIVERSITY
GPA: 3.94, NSF GRFP – Honorable Mention
AUGUST 2011 – MAY 2015
B.S. BIOLOGY (HONORS) AND COMPUTER SCIENCE, UNIVERSITY OF RICHMOND
GPA: 3.92, Phi Beta Kappa, Barry M Goldwater Scholarship, Arnold and Mabel Beckman Fellowship

SKILLS

Python (*pandas, scikit-learn,* etc.); R (*tidyverse, tidymodels,* etc.); **SQL** (*Postgres, Snowflake*); **Statistics**; **Machine Learning**; **Containerization** (*Docker*); **Version control (***git, subversion*); **Workflow orchestration** (*Apache Airflow*); **CI/CD** (*Github Actions*); **Infrastructure as Code** (*Terraform*); **Amazon Web Services** (*Lambda, S3, MSK,* etc.); **Data transformations** (*DBT*); **Data Observability** (*Monte Carlo*); **Shell scripting** (*Bash, Zsh)*; **BI Tools** (*Looker*); **Mentorship**