

SEAN HACKETT

✉ seanmchackett@gmail.com |  GitHub |  shackett.org |  607-592-2678

Summary of Qualifications:

- Uses programming and statistics to approach challenging scientific problems - Google Scholar.
- Flexibly transitions between surveying what is possible, and focusing on long-term goals.
- Excels at communicating with diverse audiences both in written and oral formats.

Skills: *Analytics:* statistics (GLMs, likelihood, MCMC,), machine learning (LASSO, random forest, LDA), optimization (LP, MILP, QP, gradient-based MAP estimation).

Programming R (dplyr, purrr, ggplot2, devtools, shiny, Connect), Python (pandas, jupyter, TensorFlow, igraph)

Tools SQL (MySQL, SQLite), Docker, GCP, Git, GitHub Actions.

EXPERIENCE

	Calico Life Sciences	S. San Francisco, CA
MANAGER	<ul style="list-style-type: none">• Helped to reorganize the Computing team to improve impact, collaboration and accountability.• Managed 4-6 data scientists providing both continued development and supporting prioritization in a problem-rich environment.• Hired and onboarded for 6 different roles.• Led initiatives around computational education, results sharing, and de-duplication of efforts.	Feb 2018 - Present
PRINCIPAL	<ul style="list-style-type: none">• Collaborated with other stakeholders to develop long-term strategies for causal inference and systems biology.	Jan 2017 - Present
DATA SCIENTIST	<ul style="list-style-type: none">• Improved approaches for finding causal connections in high-dimensional time-series using a combination of parametric modeling and LASSO.• Developed an automated metabolomics pipeline which streamlined data normalization and compound identification. Improved reliability using Docker and GitHub Actions. This initiative improved the quality of routine results and decreased time spent on manual processing.	
POSTDOCTORAL	Princeton University, Lewis-Sigler Institute	Princeton, NJ
ASSOCIATE	<ul style="list-style-type: none">• Supervisor: John Storey, Director of the Center for Statistics and ML• Used Latent Dirichlet Allocation with Empirical Bayes priors to identify latent variables affecting sparse high-dimensional sports data.	Dec 2015 - Jan 2017
GRADUATE	Princeton University, Quantitative and Computational Biology	Princeton, NJ
FELLOW	<ul style="list-style-type: none">• Adviser: Josh Rabinowitz, Professor of Chemistry and Genomics• Supervised two graduate students performing systems biology research.• Developed a scalable algorithm for combining high-dimensional datasets with scientific databases to provide novel insight into how metabolism is controlled. This led to a first-author article in Science.• Invented a data-driven method for estimating metabolic rates based on quadratic programming.	Sep 2010 - Dec 2015
EDUCATION	Princeton University Ph.D., Quantitative and Computational Biology <ul style="list-style-type: none">• Thesis: <i>Quantitative Analysis of Metabolism and Protein Abundance Using Integrative 'Omics</i>• DOE Office of Science Graduate Fellowship (SCGF): 2012 - 2015	Princeton, New Jersey November 2015
	Cornell University B.S., Biological Sciences <ul style="list-style-type: none">• Thesis: <i>Candidate gene analysis of German shepherd dogs to identify genes contributing to arrhythmogenesis</i>• Concentration in Genetics and Development• <i>Magna Cum Laude</i> with Distinction in Research	Ithaca, New York June 2006