

# SEAN HACKETT

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

---

## Summary of Qualifications:

- Actively plans and directs long-term strategy, near-term tactics, and day-to-day operations.
- Leverages statistics and programming to tackle challenging scientific problems. Google Scholar
- Excels at communicating with diverse audiences in both written and oral formats.

**Skills:** *Analytics:* statistics (GLMs, nonlinear regression, maximum likelihood, model comparison, Bayesian statistics), ML (LASSO, random forest, LDA, LLMs), optimization (linear/quadratic programming, gradient-based MAP estimation). *Programming* R (dplyr, purrr, ggplot2, devtools, Shiny, Connect), Python (pandas, Jupyter, igraph, package development) *Tools* Building and querying databases (SQL, GraphQL, BigQuery), Docker, WDL, GCP, CI/CD with GitHub Actions.

---

EDUCATION	Princeton University <b>Ph.D., Quantitative and Computational Biology</b> DOE Office of Science Graduate Fellowship (SCGF)  Cornell University <b>B.S., Biological Sciences</b> <i>Magna Cum Laude</i> with Distinction in Research	Princeton, NJ 2015  Ithaca, NY 2006
RESEARCH		
DIRECTOR	Calico Life Sciences LLC <ul style="list-style-type: none"><li>• Formed the Discovery Data Science team to balance collaborations on basic science programs with focal support and methods development.</li><li>• Established data systems to support the key outputs of the Discovery organization: mechanisms and targets.</li></ul>	S. San Francisco, CA Jan 2023 - Present
ASSOC DIRECTOR	<ul style="list-style-type: none"><li>• Collaborated with key stakeholders to develop long-term strategies for causal inference and systems biology.</li></ul>	Jan 2021 - Jan 2023
MANAGER	<ul style="list-style-type: none"><li>• Helped reorganize the Computing team to improve impact, collaboration, and accountability.</li><li>• Managed 4-6 data scientists, prioritizing high value projects in a problem-rich environment.</li><li>• Led initiatives around computational education, results sharing, and de-duplication of efforts.</li></ul>	Feb 2018 - Jan 2021
DATA SCIENTIST	<ul style="list-style-type: none"><li>• Created a genome-scale mechanistic network connected to known gene-disease associations to support indication prioritization.</li><li>• Improved approaches for finding causal connections in high-dimensional time series using a combination of parametric modeling and LASSO.</li><li>• Developed an automated metabolomics pipeline to streamline data normalization and compound identification.</li></ul>	Jan 2017 - Present
POSTDOCTORAL ASSOCIATE	Princeton University, Lewis-Sigler Institute <ul style="list-style-type: none"><li>• Supervisor: John Storey, Director of the Center for Statistics and ML</li><li>• Used Latent Dirichlet Allocation with Empirical Bayes priors to identify latent variables affecting sparse high-dimensional data.</li></ul>	Princeton, NJ 2015 - 2017
GRADUATE FELLOW	Princeton University, Quantitative and Computational Biology <ul style="list-style-type: none"><li>• Adviser: Josh Rabinowitz, Professor of Chemistry and Genomics</li><li>• Supervised two systems biology graduate students.</li><li>• Developed a scalable algorithm for combining metabolomics, proteomics and fluxes to identify novel allosteric regulators and dissect how metabolite and enzyme concentrations jointly control metabolism.</li></ul>	Princeton, NJ 2010 - 2015