

# Nikolai Vetr

CONTACT Phone: (602) 578-9196  
INFORMATION Email: nikolai.vetr@gmail.com

LinkedIn: [linkedin.com/in/nikolai-vetr](https://www.linkedin.com/in/nikolai-vetr)  
GitHub: [github.com/NikVetr](https://github.com/NikVetr)

**Education** **Postdoc**, Stanford University *Current*  
Computational Biology; Depts: Pathology, Genetics

**PhD**, University of California, Davis *2020*  
GPA: 4.0/4.0, Anthropology, Population Biology, Data Science & Informatics

**BA**, Vanderbilt University *2013*  
Earth & Environmental Sciences; Ecology, Evolution & Organismal Biology  
GPA 3.9/4.0, Departmental Honors, *summa cum laude*

- Projects** **Dissertation Work**
- Developed and implemented more efficient algorithms for computing phylogenetic likelihoods, mixing over high-dimension correlation matrices, and modularizing particular matrix operations under multivariate Brownian diffusion models of character evolution.
    - Reducing runtime for a typical analysis from weeks to hours
  - Extended these models to discrete characters under the multivariate probit and explored their approximation in the context of truncated biogeographic dispersal.
  - Applied these methods empirically to modern human, catarrhine, and fish datasets, also investigating their performance through extensive simulation experiments.

## Recent Personal Projects

I've developed, fit, interpreted, described, and visualized results from:

- high-dimensional multilevel generalized linear mixed models (GLLMs) in Stan to heart transplant patient immune response data for time-series classification of rejection risk using non-HLA antigens straddling an acute rejection episode.
- multilevel GLMMs in Stan to consumer dietary and attitudinal response datasets measuring the effects of exposure to an advertisement advocating dietary change.
- a multilevel univariate Ornstein-Uhlenbeck model in Stan to the 'evolution' of nitrogen concentrations in manure ponds across California to predict values after an arbitrary amount of time has passed in an arbitrary pond.
- an efficient, conditional multivariate probit model in R that provides personalized movie rating predictions by exploiting basic properties of Schur complements.
- many dozens of minor scripts (< 100 LOC) devoted to data visualization, text mining, web scraping, replicating published analyses or algorithms from scratch, and exploring off-the-cuff ideas proposed by myself or colleagues.

## Skills

- Generalized Linear Models (e.g. spline, probit, Gaussian process, robust, etc.)
- Principle Components Analysis
- Missing Data Imputation
- Monte Carlo Methods
- Random Forests
- Diffusion Processes / Time Series
- Support Vector Machines
- Model Comparison (e.g. via information theory, cross-validation, marginal likelihoods)
- Neural Networks
- Expectation-Maximization
- Clustering / Unclustering
- Markov Chains (CTMC / DTMC)
- Measurement Error Modeling
- Parallel & Distributed Computing

**Languages** **Programming:** R, RevBayes, Stan, BASH, Python, C++  
**Natural:** Russian, English, Spanish

<b>Leadership</b>	<b>Director</b> , Board of Directors, Wild Animal Initiative <span style="float: right;"><i>Current</i></span>
	– I serve on the Board of Directors at the <i>Wild Animal Initiative</i> , a non-profit dedicated to re-searching the lives of wild animals.
	<b>Founder &amp; Lead</b> , Applied Bayesian Statistics Research Cluster <span style="float: right;"><i>2019 - 2020</i></span>
	– Founded, coordinated, and supervised an interdisciplinary research cluster of 70+ scientists and statisticians across numerous career stages (PhD Student, Postdoc, PI, Industry Researcher). – Financially and spiritually sponsored by <i>Data Science &amp; Informatics</i> .
<b>Teaching</b>	<b>Coordinator</b> , Various Reading Groups <span style="float: right;"><i>2014 - 2020</i></span>
	– Founded or co-founded groups on Python, Linear Algebra, Bayesian Data Analysis, Deep Learning, Computational Molecular Evolution, Quantitative Genetics, and Machine Learning.
	<b>Associate Instructor</b> , UC-Davis <span style="float: right;"><i>2015 - 2020</i></span>
	– Taught three terms of an upper-division paleoanthropology course, two terms of an upper-division evolutionary primatology course, and one term of an intro course on human evolution. – Created or modified lab and lecture materials, designed and graded the assignments and tests, mentored students, delivered lectures, and supervised teaching assistants.
	<b>Teaching Assistant</b> , UC-Davis & Vanderbilt University <span style="float: right;"><i>2012 - 2020</i></span>
	– TA'd for the above, as well as the lab for an introductory cell biology course at Vanderbilt. – Designed and supervised labwork and lab-related assessments.
	<b>Outreach Lecturer</b> , UC-Davis <span style="float: right;"><i>2013 - 2020</i></span>
	– Gave multiple yearly talks on human evolution to elementary and middle school students on location and visiting UC-Davis and during campus-wide events (e.g. <i>Picnic Day</i> ). – Volunteered for various workshops targeted at adults, e.g. to medical professionals on Natural Language Processing during UC-Davis Data Science Health Day.
	<b>Course Coordinator</b> , Workshop in Applied Phylogenetics <span style="float: right;"><i>2019</i></span>
	– Served as a course coordinator for a world-renowned, widely-attended, week-long workshop in applied computational Bayesian phylogenetics at the Bodega Marine Laboratory.
<b>Field &amp; Labwork</b>	<b>Carpentries Instructor</b> , Data & Software Carpentries <span style="float: right;"><i>2019</i></span>
	– Completed Instructor Training and Checkout for The Carpentries organization. Integrated techniques learned into my own teaching on e.g. Bayesian statistics & phylogenetics.
	<b>Archaeological / Paleontological Excavator</b> <span style="float: right;"><i>2013</i></span>
	Helped in summer excavation at La Ferrassie, a Neandertal fossil and artifact dig site.
	<b>Paleoecologist</b> <span style="float: right;"><i>2012 - 2013</i></span>
	Examined how ecological and environmental conditions were recorded in enamel stable isotope ratios and dental microwear textures for marsupial taxa across Australasia.
	<b>Water Quality Analyst</b> <span style="float: right;"><i>2011</i></span>
<b>Grants &amp; Awards</b>	Analyzed water quality across a large set of streams in the North Island of New Zealand.
	Excellence in Data Science Community Training and Outreach (\$1.5k) <span style="float: right;"><i>2019, 2020</i></span>
	Outstanding Graduate Student Teaching Award Nominee <span style="float: right;"><i>2016, 2019, 2020</i></span>
	1st Place Picnic Day Exhibit Award in “Secrets of Nature” Category <span style="float: right;"><i>2017</i></span>
	UC-Davis Summer Research Grant (\$3k) <span style="float: right;"><i>2014, 2015</i></span>
	NSF Graduate Research Fellowship (\$138k) <span style="float: right;"><i>2015</i></span>
	Graduate Scholars Fellowship (\$56k) <span style="float: right;"><i>2013</i></span>
	Conference Travel Award (\$500) <span style="float: right;"><i>2012, 2018</i></span>
	Eugene H. Vaughan Undergraduate Research Assistantship in Geology (\$13k) <span style="float: right;"><i>2012</i></span>
	Vanderbilt Undergraduate Summer Research Grant (\$5k) <span style="float: right;"><i>2012</i></span>
	Ross Family Scholarship (\$60k) <span style="float: right;"><i>2012</i></span>
	National Merit Scholarship (\$20k) <span style="float: right;"><i>2009</i></span>