

Mitzi Morris

Address 128 East 7th Street, New York, NY 10009, USA

Email mitzi212@gmail.com
GitHub github.com/mitzimorris

Employment History

2017 - present [Stan Project, Columbia University, New York, NY](#)

Senior Staff Associate III / Research Programmer

Core member of the [Stan development team](#). Stan is an open-source probabilistic programming platform written in C++ with interfaces for Python, R, Julia, and the Unix command line.

- Python programming - [CmdStanPy](#) - lightweight Python interface to Stan.
- C++ programming - added features to Stan and CmdStan:
 - Generate new predictions based on existing sample (“standalone generated quantities”)
 - JSON data parser, used for user-specified initialization of Stan’s NUTS-HMC sampler
- Other programming - GNU Make makefile build scripts; Python build scripts for versioned Markdown-based documentation
- Statistical modeling - In collaboration with epidemiologists, wrote Stan programs for hierarchical models with spatial and non-spatial factors, and R scripts for analysis and visualization.
- Talks and teaching - presentations on Stan and Bayesian workflow at conferences, workshops, and meet-ups. User support via Stan Forums.

2014 - 2016 [Lucidworks, San Francisco, CA](#)

Technical Writer

Wrote blogposts and user manuals for Lucidworks Fusion, a search and analytics system built on the Apache Solr search engine and Apache Spark cluster computing framework.

2013 [Text Processing in Java](#)

Author

Wrote book *Text Processing in Java*, which covers multilingual text processing, from the basics of bytes and chars to search and classification with Lucene and Solr. Source code and data for all examples freely available via [GitHub](#).

2012 [AT&T Research, Florham Park, NJ](#)

Contractor

Implemented product search for a mobile shopping app consisting of a RESTful API to a Lucene/Solr backend and a JQuery-based UI allowing for autocomplete and suggestions.

2011 - 2012 [Department of Biomedical Informatics, Columbia University, New York, NY](#)

Programmer/Analyst Bioinformatics

Wrote Java programs to find mentions of diseases and symptoms in free text medical notes using Lucene for search and indexing and a MySQL database to aggregate results.

Employment History (continued)

- 2009 - 2011** [Center for Genomics and Systems Biology, New York University, New York, NY](#)
Programmer/Analyst Biology
Analyzed very large genomic datasets (100M items per experiment), using Java, R, and Python.
- 2008** [Alias-i, New York, NY](#)
Software Engineer
Designed and implemented LingMed, a Java-based system for indexing bio-medical databases used to annotate and cross-reference genes and diseases across MEDLINE articles.
- 2006 - 2007** [Epigenomics Lab, department of Psychiatry, Columbia University, New York, NY](#)
Senior Research Programmer
Wrote Java programs, R scripts, and MySQL databases for analysis, visualization, and storage of large datasets from high-throughput genomics experiments.
- 2005** [Buzzmetrics, New York, NY](#)
Software Engineer (contractor)
Wrote multi-threaded Java programs to spider the blogosphere, storing feeds and content in a MySQL database.
- 2002 - 2004** [Columbia Genome Center, Columbia University, New York, NY](#)
Senior Programmer/Analyst
Developed programs and databases for Geneways a system to extract gene interactions from biomedical articles. Wrote Java Swing GUIs, Oracle databases.
- 2001**
Software Engineer (contractor)
Designed and implemented multi-threaded Java application to process XML documents and store data in both a local and a remote database.
- 1999 - 2001** [New York Times Company Digital](#)
Software Engineer
Wrote Java programs to translate newsfeeds from proprietary data formats into XML.
Reimplemented news-by-email infrastructure: Java, JSP front-end; JDBC, Oracle back-end.
- 1997 - 1998** [Viacom, New York, NY](#)
Director of Quality Assurance and Testing
- 1993 - 1997** [NOMOS, Pittsburgh, PA](#)
Director of Quality Assurance and Testing
Developed software in C and C++ for radiation therapy treatment planning.
- 1990 - 1993** [CHILDES project, Carnegie Mellon University, Pittsburgh, PA](#)
Research Programmer
Designed and implemented software in C for annotation of natural language corpora.

Publications

Peer-reviewed Journal Articles

Morris M, et al. Bayesian Hierarchical Spatial Models: Implementing the Besag York Mollié Model in Stan. *Spatial and Spatio-temporal Epidemiology* 2019;31:100301 DOI: doi.org/10.1016/j.sste.2019.100301.

first author, modeling, programming

Gerstein MB, et al. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science*. 2010;330(6012):1775-87. PubMed PMID: [21177976](https://pubmed.ncbi.nlm.nih.gov/21177976/)

programming, analysis, and visualizations, (one of 18 first authors)

Rzhetsky A, et al. GeneWays: a system for extracting, analyzing, visualizing, and integrating molecular pathway data. *Journal of Biomedical Informatics*. 2004 Feb;37(1):43-53. PubMed PMID: [15016385](https://pubmed.ncbi.nlm.nih.gov/15016385/)

software development

Books

Text Processing in Java. New York: Colloquial Media, 2014. 328pp.

Online Articles

Stan Tutorial

- [Bayesian Workflow Illustrated Using BRMS](#)

Stan Case Studies

- [Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data](#)
- [Multilevel regression modeling with CmdStanPy and plotnine](#)
- [Stan Notebooks in the Cloud](#)

[Lucidworks Blog Posts](#) Various articles on configuration and security for Enterprise Search

Education

Coursework in Statistics
2006, 2009

Columbia University, NYU

Coursework in Computer Science and Mathematics
1990-1993, 1995

Carnegie Mellon University

M.A. in Linguistics
1989

University of California at Santa Cruz

B.A. in French with Honors
1981

University of Michigan