# Martin Skarzynski

Data Science · Epidemiology · Genomics · Machine Learning · Public Health · Technical Management · Software Engineering

240-595-3460 | mskar@dr.com | mskar.github.io | mskar | mskar | marskar

## Summary

My primary research interest is in understanding health risk factors by combining scientific expertise from diverse fields with machine intelligence. I believe I am uniquely equipped to bridge the gaps between scientific disciplines and deliver on the promise of data science in health research. My preferred tools are R and Python, open source programming languages kept on the cutting edge by their active and supportive communities. Through research and teaching, I am constantly improving my ability to obtain, tidy, explore, transform, visualize, model, and communicate data. I aim to utilize my technical skills and science background to become a leader among the next generation of multidisciplinary data scientists.

## Recent Experience

### Artificial Intelligence Engineering Manager

Booz Allen Hamilton — *October 2019 - Present*

- Leads a team of data scientists and software developers working on a cyber intelligence application
- Led interdisciplinary COVID-19 statistical modeling and genomic analysis efforts

### Lead Instructor

Data Society — *October 2019 - Present*

- Teaches courses such as:
  - Machine Learning
  - Deep Learning
  - DevOps
  - Data Science Bootcamp
  - Data Science Academy
  - Data Visualization
  - Natural Language Processing
  - Python Programming
  - R Programming

General Assembly — *July 2019 - Present*

- Teaches courses such as:
  - Data Science 5000
  - Python Programming
  - Tech Excellence Data Science
  - React Development
  - Data for Leaders
  - Front-End Web Development

### Adjunct Professor

Virginia Tech — *September 2021 - Present*

- Teaches graduate Data Analytics course

### Treasurer

Data Community DC — *April 2020 - Present*

- Manages the finances of a non-profit organization that supports eleven Meetup groups

### Cancer Prevention Fellow

Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute — *June 2017 - April 2021*

- Integrated clinical, laboratory, epidemiologic, genomic, and medical imaging data
- Combined deep learning and statistical inference using stacked ensembles

### Bioinformatics and Data Science Department Co-Chair

Foundation for Advanced Education in the Sciences — *January 2014 - May 2020*

- Co-administered an academic program with over twenty faculty members
- Taught three graduate data science courses:
  - Introduction to Python
  - Introduction to Text Mining
  - Applied Machine Learning
- Taught nine graduate biotechnology workshops:
  - Pharmacometric Analyses using R
  - Junior Scientist Training Program
  - Immunochemistry and Antibodies
  - Recombinant DNA Methodology
  - Research Tools for Studying Disease
  - Cancer Prevention
  - Cellular Immunology
  - Genetics Institute
  - Flow Cytometry

### Online Course Developer

DataCamp — *October 2018 - Present*

- Developed an online course called Creating Robust Python Workflows (part of Coding Best Practices track)

### Postdoctoral Fellow

National Cancer Institute — *July 2015 - June 2017*

- Conducted genomic analysis of immune and cancer cells
- Developed and tested immunotherapeutic agents
- Mentored a trainee from the NIH High School Scientific and Training Enrichment Program

### Adjunct Professor

George Washington University — *January 2015 - May 2015*

- Taught two undergraduate courses:
  - Biology of Organisms
  - Women and Leadership

### Predoctoral Fellow

National Institues of Health — *July 2009 - October 2015*

- Tested kinase inhibitor and monoclonal antibody therapeutic efficacy and drug interactions
- Engineered and assayed therapeutic immunotoxins
- Quantified cancer cell signaling pathways
- Mentored trainees from three different NIH training programs:
  - Summer Internship Program
  - Medical Research Scholars Program
  - Biomedical Research Training Program for Individuals from Underrepresented Groups

## Select Publications

### SARS-CoV-2 genome-based severity predictions correspond to [...] higher viral load

medRxiv — *Nov. 2021*

### Variants in SARS-CoV-2 associated with mild or severe outcome

Evolution, Medicine, and Public Health — *June 2021*

### Using prediction models to reduce [...] disparities in [...] lung cancer screening [...]

Journal of the National Cancer Institute — *Jan. 2021*

### Interactions between ibrutinib and anti-CD20 antibodies [...] combination therapy

Clinical Cancer Research — *Jan. 2016*

### Designing the furin-cleavable Linker in recombinant immunotoxins [...]

Bioconjugate Chemistry — *May 2015*

### Health disparities in the immunoprevention of [HPV] [...] associated malignancies

Frontiers in Public Health — *Dec. 2015*

### Harnessing the Fcμ receptor for potent and selective cytotoxic therapy of [CLL]

Cancer Research — *Oct 2014*

## Education

### MPH, Epidemiologic and Biostatistical Methods for Public Health and Clinical Research

Johns Hopkins University School of Public Health — *May 2018*

- Data Science Certificate

### PhD, Tumor Biology

Georgetown University-NIH Graduate Partnerships Program — *May 2015*

- Teaching Certificate

### MS, Biotechnology

Jagiellonian University — *June 2009*

### BA, Biology

St. Mary's College of Maryland — *May 2007*

## Select Awards

### Trans-Fellowship Research Award

Cancer Prevention Fellowship Program — *June 2019*

### Cancer Research Training Award

National Cancer Institute — *November 2015*

### Fellows Award for Research Excellence

National Institutes of Health — *June 2015*

### Director's Award

National Heart Lung Blood Institute — *October 2014*

### Orloff Science Award

National Heart Lung Blood Institute — *January 2014*