

# Mahesh Vangala

A programming enthusiast with full stack expertise bringing ideas to life, using but not limited to, Angular/Python/Beam/Kafka/(Neo4j, Cassandra)/Docker/Kubernetes/(AWS, GCP).

With 10+ years of hands on experience in academic, research and clinical settings working with genomic and clinical data, I bring broad spectrum of skills ranging from Data Engineering, Data Science and Dev/Ops.

Creative and goal oriented, I take immense pride in end-to-end automation of operational tasks and a big believer of the quote "you don't know it until you teach it to a computer".

Looking forward to the opportunities to prove and enhance my leadership and technical expertise.

## Contact Info

- ✉ [vangalamaheshh@gmail.com](mailto:vangalamaheshh@gmail.com)
- 🌐 [linkedin.com/in/maheshvangala](https://www.linkedin.com/in/maheshvangala)
- 🐙 [github.com/vangalamaheshh](https://github.com/vangalamaheshh)
- 🌐 [vangalamaheshh.github.io](https://vangalamaheshh.github.io)
- ☎ [\(443\)326-1957](tel:(443)326-1957)

## Experiences

- 2022 | present  
● **Sr. Software Eng. - Data Analytics & Software Development**  
Institute for Protein Innovation 📍 Boston, MA
- 2017 | 2022  
● **Data Scientist**  
UMass Chan Medical School 📍 Worcester, MA
- 2014 | 2017  
● **Sr. Computational Biologist**  
Dana-Farber Cancer Institute 📍 Boston, MA
- 2012 | 2014  
● **Bioinformatics Research Associate**  
Vermont Genetics Network, University of Vermont 📍 Burlington, VT
- 2010 | 2012  
● **Bioinformatics Software Engineer**  
Institute for Genome Sciences, Univeristy of MD School of Medicine 📍 Baltimore, MD

## Selected Projects

- 2020 | 2022  
● **National COVID Cohort Collaborative (N3C)**  
UMass Chan Medical School
  - Led the project from plan to prototype to production in under 2 months.
  - Designed and delivered end-to-end automation of weekly data transfers.
  - Continuous integration of feature updates using Docker and AWS Fargate stack.
  - Automated email notifications of data quality metrics.
- 2015 | 2017  
● **Visualization Pipeline for RNA-seq data (VIPER)**  
Dana-Farber Cancer Institute
  - A comprehensive solution that performs most standard RNA-seq analyses quickly and effectively.
  - Published work in peer reviewed journal. <https://doi.org/10.1186/s12859-018-2139-9>
  - Played a vital role in devising and developing module based design pattern.
- 2010 | 2012  
● **CloVR: a virtual machine for automated and portable sequence analysis from the desktop using cloud computing**  
Univeristy of MD School of Medicine
  - CloVR supports use of remote cloud computing resources to improve performance for large-scale sequence processing.
  - Published work in peer reviewed journal. <https://doi.org/10.1186/1471-2105-12-356>
  - My work into comparative genomics pipeline in CloVR resulted in further funding of the project.

## Education

- 2008 | 2010  
● **Professional Science Masters in Bioinformatics (Genomics track)**  
Virginia Commonwealth University 📍 Richmond, VA
- 2005 | 2007  
● **M.S in Biophysics**  
University of Madras 📍 Chennai, India
- 2001 | 2004  
● **B.S. in Biotechnology**  
Osmania University 📍 Hyderabad, India

## Skills

- Rust
- Python
- Java
- Perl
- Bash
  
- Angular
- GraphQL
  
- Docker
- Kubernetes
  
- Snakemake
- Apache Beam
- Apache Spark
  
- RabbitMQ
- Redis
- Apache Kafka
  
- SQL Server/MySQL
- Neo4J
- Cassandra
- Bigquery
- Athena
- DynamoDB
  
- AWS
- GCP