

# RAZVAN IOAN PANEA

Contact: [contact@razvanpanea.com](mailto:contact@razvanpanea.com) | (917) 755-6090  
Website: [razvanpanea.com](http://razvanpanea.com)  
GitHub: [github.com/ripanea](https://github.com/ripanea)  
LinkedIn: [linkedin.com/in/rpanea](https://linkedin.com/in/rpanea)

---

## Summary

- Initiative-taking with strong knowledge and experience in computer science and bioinformatics.
- Developed from scratch a cloud-based workflow management system for bioinformatics in Python.
- Led a group of software engineers to develop a complete cloud-based infrastructure using Kubernetes and Django.
- Processed a terabytes-scale genomics dataset of Burkitt lymphoma and performed analysis using R.

---

## Education

09/2014 – **Duke University | Durham, NC, USA**  
05/2020 Ph.D. in *Computational Biology and Bioinformatics*  
Research: *Cloud computing and cancer genomics*

09/2011 – **Jacobs University Bremen | Bremen, Germany**  
06/2014 Bachelor of Science in *Applied and Computational Mathematics*  
Specialization: *Bioinformatics and Computational Biology*

---

## Research and Work Experience

09/2014 – **Duke University | Durham, NC, USA**  
Present *Ph.D. Candidate* – Advisor: Sandeep Dave, MD  
► Developed a cloud-based infrastructure for cancer genomics  
Technologies used: *Python, Bash, Git, SQL, Docker, Kubernetes, Google Cloud, AWS, Apache libcloud*

- Developed CloudConductor, a cloud-based, modular, scalable, parallelizable, and extensible workflow engine to generate and execute analysis pipelines. Available at [github.com/labdave/CloudConductor](https://github.com/labdave/CloudConductor)
- Led a team of 5 software engineers and bioinformaticians throughout the SDLC process to extend CloudConductor to a complete cloud-based infrastructure for bioinformatics analysis
- Improved CloudConductor to a platform-agnostic workflow engine using Apache libcloud
- Integrated and evaluated the workflow engine on Google Cloud Platform and Amazon Web Services
- Integrated 50+ bioinformatics tools using Docker and utilized them to analyze terabytes of genomics data

► Characterized the mutational patterns in Burkitt lymphoma subtypes  
Technologies used: *R, Python, Bash, SQL*

- Sequenced the whole genome and transcriptome of 101 Burkitt Lymphoma cases of different subtypes
- Identified genetic alterations in the sequencing data using the newly developed cloud-based infrastructure
- Identified mutations in driver genes, such as *MYC*, *ID3* and *DDX3X*, the presence of the Epstein-Barr virus (EBV) in the samples, and translocations characteristic to Burkitt Lymphoma
- Identified associations between lymphoma subtypes and EBV status
- Work has been published in the *Blood* journal (*Panea et al., Blood, 2019*)

06/2013 – **Max Planck Institute for Marine Microbiology | Bremen, Germany**  
08/2014 *Software Developer* - Supervisor: Dr. Frank-Oliver Glöckner  
► Improved features of the SILVA Project ([www.arb-silva.de](http://www.arb-silva.de)) software using Python and C++ binding  
Technologies used: *C++, Python*

- Researched and tested SWIG, a software for Python and C++ binding
- Implemented SWIG in the project to bind and call methods implemented in Python on C++ objects

► Tested the performance and efficiency of a phylogenetic tree generating software  
Technologies used: *C++*

- Benchmarked a new phylogenetic tree generator software
- Compared the performance and quality of the new tree generator software to the one currently used
- Deployed the new, more efficient software

03/2013 – **Jacobs University Bremen | Bremen, Germany**  
06/2014 *Research Assistant* – Advisor: Dr. Georgi Muskhelishvili  
► Determined the interaction between two neighboring genes in *E.coli*  
Technologies used: *Java*

- Developed a software in Java to aggregate, analyze and report the expression levels of all neighboring genes at different time points in *E.coli* life cycle
- Identified a correlation between the differential expression between two neighboring genes and their transcriptional orientation to each other

---

## Teaching Experience

01/2016 – **Duke University | Durham, NC, USA**  
05/2016 *Teaching Assistant* – CompSci 260: Introduction to Computational Genomics

- Designed and graded course assignments
- Organized and presented tutorial sessions for programming in Python
- Held office hours to answer students' questions and provide supplemental instructions

09/2012 – **Jacobs University Bremen | Bremen, Germany**  
06/2014 *Teaching Assistant* – NatSciLab Unit Computer Science I/II

- Assisted computer science professors in teaching courses based in C and Python
- Offered supplemental instruction and assisted students in planning their programming projects
- Assessed student progress and offered weekly reviews to help improve student performance

---

## Publications

09/2019 **Panea R\***, Love C\*, Shingleton J\*, Reddy A\* et al., "The whole genome landscape of Burkitt lymphoma subtypes", *Blood*, 2019

02/2020 Shingleton J., [ and 16 authors, including **Panea R.**], "Non-Hodgkin Lymphomas: Malignancies Arising from Mature B Cells", *Cold Spring Harbor Perspectives in Medicine*, 2020

11/2019 Li X., [ and 75 authors, including **Panea R.**], "Whole Exome and Transcriptome Sequencing in 1042 Cases Reveals Distinct Clinically Relevant Genetic Subgroups of Follicular Lymphoma" *Blood* (2019), (Manuscript in preparation).

---

## Posters and Presentations

07/2018 **Google Next '18 | San Francisco, CA, USA**  
*Presentation and Panel Discussion*  
Title: Speeding up Research in Genomics  
Link: <https://youtu.be/goBFt3B976A?t=11m31s>

06/2016 **Intelligent Systems for Molecular Biology | Orlando, FL, USA**  
*Poster Presentation*  
Title: Defining the Microbiome of Lymphomas

10/2017 **Duke Cancer Institute | Durham, NC, USA**  
*Poster Presentation*  
Title: A Cloud-Based Framework for Cancer Genomics

---

## Skills

**Programming:** Python, R, Bash, C, Java, C++, SQL, Git  
**Libraries/Packages:** Apache Libcloud, Plotly, Pandas, SQLAlchemy  
**Computing Platforms:** Google Cloud, SLURM, Kubernetes, AWS  
**Operating Systems:** Linux, Windows, OS X, Chrome OS  
**Development Software:** PyCharm, Jupyter Notebook, RStudio, MySQL Workbench, Adobe Illustrator  
**Bioinformatics Tools:** BWA, Samtools, BCFtools, VCFtools, GATK, Picard, STAR, RSEM, Trimmomatic, Diamond, Annovar, BEDtools