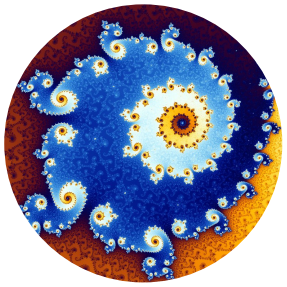

Remy **Ochei**

Software Engineering · Machine Learning · Data Science · Research Engineering

469-553-8882 | do.infinitely@gmail.com | linktr.ee/aleph_zero | doInfinitely

Eadem mutata resurgo.
"Changed, yet the same, I rise again."

Education

University of North Texas — *Denton, TX*
M.S. in Artificial Intelligence — *August 2023 -*

University of Texas at Dallas — *Dallas, TX*
Ph.D. Candidate, Computer Science — *2018 - 2019*

University of Texas at Dallas — *Dallas, TX*
Masters Cert. in Systems Engineering and Management — *2017 - 2018*

Baylor College of Medicine — *Houston, TX*
Medical Degree — *2015 - 2017*

Duke University — *Durham, NC*
A.B. in Computer Science, B.S. in Neuroscience, Music Minor, Japanese Minor — *2009 - 2013*

Experience

Modulate AI — *Cambridge, MA*
Senior Research Engineer — *08/2022 - Present*

- Owned, researched, and improved toxicity detection scoring algorithms at Modulate, combining signals from transcription, emotion, demographics, and other models evaluating multi-participant conversations, to understand when harm is occurring, categorize it, and escalate it with the appropriate severity/confidence.
- Developed novel scoring algorithms and improved existing models to ever more accurately evaluate conversations in voice chat. Models in particular take transcriptions, emotion labels, demographic information, and other signals as inputs in order to classify conversations
- Ensured high precision of the scoring system in complicated conversational contexts, such as uses of reclaimed language, by maintaining robust and diverse training, evaluation, and test datasets.
- Developed tools to evaluate models on representative production data, allowing Modulate to ship model updates without causing accuracy regressions to existing customers.

Python, Numpy, Gitlab, Natural Language Processing, NoSQL/SQL

Contract Work (Fund for Alignment Research, Home Depot, Marcillac, TripleByte) — *Remote*
Machine Learning Engineer, Software Engineer, Subject Matter Expert - General Coding Logic — *07/2020 - 05/2023*

- Conducted machine learning experiments for the Fund for Alignment Research using PyTorch and Weights and Biases
- Built an end-to-end machine learning pipeline to monitor the state of shelves in Home Depot stores.
- Designed a proof assistant that semi-automatically verifies the correctness of computer programs with Marcillac
- As a Subject Matter Expert for TripleByte, I developed new content for multiple-choice skill assessments in general coding logic. Through these questions I tested understanding of various concepts, models, technologies and practical implementation considerations spanning a range of skill levels — from Entry-level through Principal.
- Built an adaptive testing recommendation engine for nurses
- Implemented a GPU version of a CPU only computer vision library
- Built a scraper to parse equity news sites for IPO evaluations
- Wrote a lesson on using bootstrapping to compute Upside Potential Ratios
- Taught a programming bootcamp
- Tutored computational physics

Python, Javascript, Numpy, Github/Git, NLP, Computer Vision, Docker, Kubernetes, Object Detection, PyTorch, Weights and Biases

Dolthub — *Dallas, TX*
Software Engineer — *12/2020 - 07/2021*

Dolt adds revision control capabilities and git semantics to SQL databases. I preside over data bounties, contests in which our community members collaborate and compete to construct large and intriguing datasets.

- Interfaced with our Discord community, answering bounty participant questions and reviewing their pull requests.
- Designed the data bounties, from the contest copy to the sql schemas, and prototyped the collection process to anticipate participant difficulties.
- Wrote programs to facilitate working with and understanding massive datasets.

Python, Go, Github/Git, Shell Scripting, Pandas, Web Scraping, Markdown, MySQL

Self-Decode
Dallas, TX

Bioinformatician
12/2019 - 12/2020

I worked as a bioinformatician (bioinformaticist) for Self-Decode, which involved understanding and transforming primary literature in medicine, statistics, and computer science.

- Worked with massive genomic files.
- Calculated polygenic and behavioral risk scores for serious illnesses, such as cancers, heart disease, and COPD.
- Designed procedures to convert these scores to "intuitive, human useful" measures.
- Worked on a risk assessment for COVID-19, which estimated the severity of a case of coronavirus, given background risk factors such as age and pre-existing conditions.

Python, R, NumPy, Statistics, Pandas, Decision Trees

Google
Mountain View, CA

Software Engineer, Virtual Assistants and Home Automation
06/2014 - 07/2015

As a member of the Third Party Actions team, I enabled interactions between third-party developers and Google Assistants.

- Built the integration of Nest Smart Thermostats with Google services.
- Engaged in grammar engineering to bind search queries to the appropriate response.
- Attended TGIFs, weekly all-hands meetings on the global mission and activities of Google.
- Learned and used internal and open-source tools for maintaining and modifying the largest codebase in human history.

Python, Java, C++, Javascript, HTML, Natural Language Processing, MySQL, Protocol Buffers

Research

Gabianni Theoretical Neuroscience Lab
Houston, TX

Researcher, Probability Theorist, Programmer
2016

I studied mathematical models of sensory information processing in insects at Rice University. Most of my time was spent pouring through "Mathematics for Neuroscientists," Professor Gabianni's textbook.

Nicolelis Brain/Machine Interface Labs
Durham, NC

Undergraduate Researcher, Information Theorist, Programmer
2011 - 2013

The Nicolelis Laboratory is best known for pioneering studies in neuronal population coding, Brain Machine Interfaces (BMI) and neuro-prosthetics in human patients and non-human primates.

- Used my programming and mathematical skills to assist in the prototyping of neuroprosthetics.
- Created linear regression models and implemented Kalman Filters to predict the motor intentions of monkeys from neuronal data.
- Wrote code to integrate a mini mind-controlled car with the BMI system.

My **senior thesis** was the development of an Information Theoretic model of an Artificial Tactile Stimulation Task. The task itself was a discriminatory task in which rhesus macaques learned to distinguish digital objects based on their artificial texture which they sensed via direct brain stimulation when they interacted with the objects on a computer screen. My thesis analyzed their performance on these tasks to produce quantitative measures of the channel capacity of BMI systems.

Johns Hopkins Center For Language and Speech Processing
Baltimore, MD

Undergraduate Researcher, Computational Linguist
2012

Each summer, CLSP organizes and host a few international teams for an intensive 6-week research workshop on speech and language engineering. These very successful workshops have had a widespread impact on the Human Language Technology community.

- Developed natural language processing technologies as part of a team of researchers.
- Created automatic extractive summarizers of multi-speaker audio documents.
- Implemented a reduction of Graph Edit Distance to Binary Linear Programming.
- Computed Kullback-Leibler and Jensen-Shannon Divergence for parse trees.
- Participated in seminars for advanced techniques in Natural Language Processing and Computer Vision.
- Presented research to sponsors at the NSF, Office of the Director of National Intelligence, and Google Research.