

# Tony Kabilan Okeke

kabilan108.github.io

GitHub/kabilan108

Linkedin/kabilan108

tonykabilanokeke@gmail.com

Philadelphia, PA

Education	<b>Drexel University, Philadelphia PA</b> B.S. & M.S. IN BIOMEDICAL ENGINEERING Concentrations: Bioinformatics, Neuroengineering, Bioimaging	Sep 2019 - Jun 2024 GPA: 4.00
Skills	Programming: Python, R, Shell, C++, MATLAB, SQL, AWK, PHP, JavaScript ML Frameworks: TensorFlow, PyTorch, HF Transformers, Scikit-learn, LangChain, LlamaIndex Tools & Packages: Seurat, SquidPy, CellProfiler, FastAPI, Flask, React.js, Docker	
Experience	<b>Data Scientist (Bioinformatics)</b> IIVT NON-CLINICAL SAFETY, GSK - COLLEGEVILLE PA	Apr 2022 - Sep 2022, Apr 2023 - Sep 2023
	<ul style="list-style-type: none"><li>Enhanced spatial transcriptomics (10X Visium) data analysis by introducing advanced software and machine learning models to improve quality control, visualizations, cell type annotations and accurate quantification of gene expression across different tissue types.</li><li>Facilitated the integration of internally generated spatial transcriptomics (ST) data with publicly available single-cell data to enable comprehensive analysis and interpretation of ST data in conjunction with broader cellular information.</li><li>Developed Python scripts to streamline the identification of off-target genes expressed in specific tissues by automating the import of tissue-specific gene expression data from public domains like NCBI and Bgee.</li><li>Collaborated with the HESI eSTAR carcinogenicity working group to develop a neural network model that predicts molecular initiating events associated with liver carcinogenicity, leveraging rat transcriptomics studies.</li><li>Utilized UMAP and t-SNE clustering methods on a comprehensive ChEMBL dataset to group HESI compounds by global embeddings, and developed machine learning models for biomarker identification within each cluster to assess carcinogenicity, enabling more targeted and insightful evaluations.</li><li>Spearheaded the development of an interactive analytics and visualization platform for spatial transcriptomics data (GeoMx), supporting future acquisition plans and facilitating in-depth, interactive reporting.</li><li>Developed an interactive web-app for performing statistical analyses and biomarker discovery on high-content imaging datasets using <i>Dash</i> and <i>Flask</i>; scaled computations using <i>Dask</i> and deployed on <i>RStudio Connect</i>.</li><li>Implemented machine learning algorithms including Decision Trees, Random Forests, and SVM for biomarker discovery in high-content imaging datasets.</li><li>Implemented various deep learning models (Noise2Void, Cellpose) to improve the performance of Cell Profiler image segmentation and feature extraction pipelines.</li></ul>	
	<b>Computational Research Assistant</b> INVENIO LAB, PENN MEDICINE - PHILADELPHIA PA	Mar 2021 - Aug 2022
	<ul style="list-style-type: none"><li>Led rigorous analytical efforts resulting in four peer-reviewed publications, focusing on the immune response in COVID-19 patients, including comprehensive studies on serum markers, urinary biomarkers, and neurological impacts.</li><li>Conducted extensive data analysis to unveil temporal changes in serum markers related to organ failure and anti-viral treatments, contributing to a deeper understanding of COVID-19 progression and patient management strategies.</li><li>Developed Python and R scripts to automate ingestion, quality control, and analysis of data generated from ELISA and Next Generation Sequencing assays; performed differential methylation and pathway enrichment analyses on NGS data sets.</li></ul>	
	<b>Research Assistant</b> ZHOU LAB, CHILDREN'S HOSPITAL OF PHILADELPHIA - PHILADELPHIA PA	May 2020 - Jun 2021
	<ul style="list-style-type: none"><li>Contributed to the development of the 'SeSAmE' package which provides utilities to support the analysis of Infinium DNA methylation data sets.</li><li>Developed tests to validate package functionalities against publicly available datasets from the Gene Expression Omnibus (GEO).</li></ul>	

## Projects

### Meddibia

MAR 2023

Python, TensorFlow, Flask, Flutter

- Pioneered a cross-functional team to win a \$5,000 prize at a hackathon, leveraging diverse skills to develop and deploy a medical diagnostic tool with a user-friendly mobile app interface.
- Engineered Python scripts for fine-tuning state-of-the-art neural networks like VGG16 and EfficientNet, achieving a 70% classification accuracy for skin lesion analysis from Kaggle datasets using Google Colab for computation.
- Integrated GPT-3 for natural language processing to distill symptoms from patient descriptions, coupled with a Naive Bayes classifier for disease prediction, and encapsulated the ML pipeline within a Flask API for real-time mobile app queries.

### MLGO: Machine Learning for Predicting GO Enrichment

SEP 2022

Python, R, PyTorch, limma

- Orchestrated an end-to-end R pipeline for preprocessing and differential expression analysis of over 11,000 RNA sequencing datasets from DEE2, leveraging Google Cloud VMs for high-compute capacity.
- Developed and tuned an autoencoder neural network to reconstruct log fold-change values from the datasets, explored latent space with PCA, UMAP, and t-SNE, and predicted Gene Ontology (GO) enrichment, utilizing tools such as TensorBoard and Keras Tuner for optimization.

### Identification of Connectomic Biomarkers for Autism using Machine Learning

AUG 2022

Python, sklearn, bctpy

- Developed Python scripts to compute a collection of graph theory measures on a dataset of structural connectomes from patients with Autism Spectrum Disorder (ASD); Employed machine learning models (LASSO regression, Support Vector Machine) for feature selection and diagnostic accuracy assessment.
- Executed a comprehensive study correlating connectomic biomarkers with clinical outcomes such as autism severity, social communication, and intelligence, utilizing data visualization tools to represent the intricate relationships within a multimodal neuroimaging dataset.

## Publications

Laudanski, K., Liu, D., **Okeke, T.**, Restrepo, M. and Szeto, W.Y., 2022. *Persistent Depletion of Neuroprotective Factors Accompanies Neuroinflammatory, Neurodegenerative, and Vascular Remodeling Spectra in Serum Three Months after Non-Emergent Cardiac Surgery*. Biomedicines 10.10 (2022): 2364.

Laudanski, K., **Okeke, T.**, Siddiq, K., Hajj, J., Restrepo, M., Gullipalli, D. and Song, W.C., 2022. *A disturbed balance between blood complement protective factors (FH, ApoE) and common pathway effectors (C5a, TCC) in acute COVID-19 and during convalesce*. Scientific Reports 12.1 (2022): 13658.

Laudanski, K., **Okeke, T.**, Hajj, J., Siddiq, K., Rader, D.J., Wu, J. and Susztak, K., 2021. *Longitudinal urinary biomarkers of immunological activation in covid-19 patients without clinically apparent kidney disease versus acute and chronic failure*. Scientific Reports, 11(1), p.19675.

Laudanski, K., Hajj, J., Restrepo, M., Siddiq, K., **Okeke, T.** and Rader, D.J., 2021. *Dynamic changes in central and peripheral neuro-injury vs. Neuroprotective serum markers in COVID-19 are modulated by different types of anti-viral treatments but do not affect the incidence of late and early strokes*. Biomedicines 9.12 (2021): 1791.

Laudanski, K., Jihane, H., Antallosky, B., Ghani, D., Phan, U., Hernandez, R., **Okeke, T.**, Wu, J., Rader, D. and Susztak, K., 2021. *Unbiased analysis of temporal changes in immune serum markers in acute COVID-19 infection with emphasis on organ failure, anti-viral treatment, and demographic characteristics*. Frontiers in Immunology, 12, p.650465.

## Abstracts

**Okeke, T.**, Siddiq, K., Restrepo, M., Tadikonda, P. and Laudanski, K., 2022. 181: PATTERNS OF URINARY BIOMARKERS OF IMMUNOLOGIC ACTIVATION AND SEPSIS IN PATIENTS WITH COVID-19. Critical Care Medicine, 50(1), p.74.

Restrepo, M., **Okeke, T.**, Siddiq, K., Tadikonda, P. and Laudanski, K., 2022. 174: LONGITUDINAL CHANGES OF NEURO-SPECIFIC SERUM PROTEINS IN COVID-19 PATIENTS. Critical Care Medicine, 50(1), p.71.

## Honors

Drexel University Dean's List

Fall 2019 - Present

HESI GTTC Professional Development Award, \$1000

Spring 2023

Philly CodeFest Collaborative Team Award (Meddibia), \$5000

Winter 2023

## Service & Leadership

Drexel Computational Design, Vice President & Cofounder

Mar 2021 - Present

Drexel Society of AI, Member

Sep 2022 - Present

Mid-Atlantic SOT, Member

Oct 2023 - Present

Tau Beta Pi, Member

Dec 2021 - Present