

Detravious Jamari Brinkley

dj.brinkley@ufl.edu | github.com/Brinkley97 | Research Website

Research Interests: Machine Learning, NLP, Multimodal Analysis, Information Extraction, Question Answering, Recommender Systems, Taxonomy Development, Time Series

EDUCATION

University of Florida (UF), Gainesville, FL Aug 2024 – Current
Doctor of Philosophy in Computer Science (CS), 3.83 GPA

University of Southern California (USC), Los Angeles, CA Aug 2021 – May 2024
Master of Science in Computer Science (CS), 3.00 GPA

Claflin University (CU), Orangeburg, SC Aug 2016 – Nov 2020
Bachelor of Science in Computer Science (CS), 3.75 GPA

PUBLICATIONS

1. Yamoah KA, Agyapong G, Scroggins K, Parekh N, **Brinkley D**, Jayaweera C, Gilda S, Dorr B, Dorley E (2026)
An Elicitation-Matrix Approach to Pragmatic Context Modeling in Low-Resource Machine Translation: The Case of Akuapem Twi
The International FLAIRS Conference Proceedings, 39(1)
PDF
2. Driggers-Ellis C, **Brinkley D**, Chen R, Dhawan A, Wang DZ, Grant C (2025)
MultiScript30k: Leveraging Multilingual Embeddings to Extend Cross Script Parallel Data
arXiv preprint [arXiv:2512.11074](https://arxiv.org/abs/2512.11074)
arXiv
3. **Brinkley D**, Johnson E, Feng T, Gil Y (2023)
Reproducibility as a Stepping Stone to Intelligent Assistants for Data Analysis: Recreating a Study of Physical Activity, Sleep, and Work Shift in Nurses
Association for Computing Machinery at the Intelligent User Interfaces Conference
PDF
4. **Brinkley D**, Johnson E, Feng T, Gil Y (2022)
Reproducibility as a Stepping Stone to Intelligent Assistants for Data Analysis: An Analysis of Physical Activity, Sleep, and Work Shift in Nurses
Technical Report at USC Information Sciences Institute
PDF
5. Calles J, Justice I, **Brinkley D**, Garcia A, Endy D (2019)
Fail-safe genetic codes designed to intrinsically contain engineered organisms
Nucleic Acids Research Journal
PDF
6. Calles J, Justice I, **Brinkley D**, Garcia A, Endy D (2019)
Synthetic Genetic Codes Designed to Hinder Evolution
bioRxiv
PDF

RESEARCH EXPERIENCE

- **Christan Grant, UF Data Studio** UF
Ph.D. Student Researcher Aug 2024 – Present
 - Create a working predictive statement taxonomy/codebook to annotate core semantic properties (Source, Target, Date, Outcome, Metadata) and support prediction verification against real-world observations.

- Benchmark a comprehensive hierarchy of model families—including traditional ML, DL networks, Transformer encoders, and Generative LLMs—to extract predictive statements from unstructured text.
- Engineer a modular Retrieval-Augmented Generation (RAG) framework using the Builder design pattern for "LLM-as-a-Judge" Question Answering (QA), integrating vector databases and semantic search to formulate the verification of predictive statements using retrieved observations.
- Architect an automated evaluation pipeline to assess predictive statement veracity (entailment vs. contradiction) and source reliability, establishing a testing framework for multiple embedding strategies (spaCy, SentenceTransformer) and an ensemble majority-vote system measured via Fleiss' Kappa.
- Prototype sequence labeling and entity resolution workflows in Figma and Jupyter Notebooks, exploring Named Entity Recognition (NER), Levenshtein distance, and clustering to identify these core semantic properties.

• **Yolanda Gil, Knowledge Capture And Discovery**

Information Sciences Institute at USC

Graduate Student Researcher

Feb 2022 – Aug 2023

- Reproduced and extended an original behavioral health study by developing the Tereatha Grant Nwankwo (TGN10+) system using a dataset of 113 nurses with demographic, behavioral, psychological, physiological, and time series variables.
- Implemented statistical modeling and hypothesis testing workflows (t-test, ANOVA, linear regression) using Python, R, pandas, NumPy, SciPy, and rpy2 to analyze subgroup differences and predict variables of interest.
- Integrated TGN10+ with DISK, an automated scientific workflow system, using Protege, Docker, Workflow INstance Generation and Specialization, and GitHub Actions to support reproducible computational research.
- **Spearheaded project resulting in two submitted papers (See [3, 4] in PUBLICATIONS).**

• **Deidra Morrison Wells, CS and Shannon Smith, Chair of Nursing**

Claflin University

Undergraduate Student Researcher

Feb 2019 – Jul 2021

- Applied data analysis and statistical methods to study user engagement in a native software application, analyzing preferences from 26 participants seeking like-minded fitness partners.
- Designed and analyzed 8 survey polls grounded in the Theory of Planned Behavior to extract insights on exercise habits, user motivation, and the influence of accountability partners.
- **Completed senior honors thesis through the Alice Carson Tisdale Honors College, culminating in an applied research project on fitness partner matching and user motivation.**

• **Deidra Morrison Wells, CS and Shannon Smith, Chair of Nursing**

Claflin University

Undergraduate Student Researcher

May 2020 – Jul 2020

- Designed and evaluated AccountaFitBuddy, a web application built with HTML, CSS, JavaScript, Python, Django, Heroku, and Amazon S3 to connect users with nearby fitness partners.
- Conducted a user experience study with 6 participants using the User Experience Questionnaire, with 66.7% expressing motivation with the current version and interest in an updated version based on recommended changes.
- **Developed the technical application component of the senior honors thesis project, connecting applied software development with user-centered evaluation.**

• **Joshua Sherfield, Domingo Colon**

Lawrence Livermore National Laboratory

Undergraduate Student Researcher

May 2019 – Jul 2019

- Improved usability of NeMS, an in-house network mapping tool, by replacing command-line interaction with a web-based GUI developed using HTML, CSS, Python, and Flask.

• **Drew Endy, Endy Lab**

Stanford University

Undergraduate Student Researcher

Jul 2018 – Aug 2018

- Developed Python and Biopython scripts to search the UniProt database and identify functional proteins encoded with a reduced set of amino acids.
- **Substantiating project resulted in two submitted papers (See [5, 6] in PUBLICATIONS).**

- **Belinda Wheeler, English**

Undergraduate Student Researcher

Clafin University

Apr 2017 – Jan 2018

- Executed a data-driven community research study in Orangeburg County by surveying 164 students and conducting interviews with 19 business owners, managers, and community leaders.
- Applied data analysis and statistical methods to extract insights supporting community development and strategic decision-making.
- **Selected as one of the oral presentation winners during Clafin's Annual Research Day.**

APPLIED RESEARCH, INDUSTRY EXPERIENCE, AND PROJECTS

- **Your Therapeutic Cookbook: Cancer Patient Diet RAG System**

Gainesville, FL

Hacking Cancer with Human Centered Data Science

Sep 2025

- Designed and implemented a dual-embedding retrieval architecture using Chroma Vector Store and HuggingFace embedding models to generate personalized nutrition plans for cancer patients based on cancer type, treatment phase, and symptoms.
- Engineered document embedding pipelines to process dietary datasets and patient information into vector representations, enabling semantic similarity search for evidence-based dietary recommendations.
- Developed query embedding system that combines patient calories, weight goals, food allergies, and preferences in the same vector space to identify top-K most relevant dietary guidelines grounded in authoritative medical knowledge.

- **Colgate-Palmolive**

Piscataway, NJ

Ph.D. Student – Data Scientist and Engineer

Jun 2024 – Aug 2024

- Developed an automated ETL pipeline for 2,000+ body wash samples, centralizing experimental data to support faster testing, analysis, and visualization workflows for chemists.
- Automated data extraction from Google Drive, transformed and updated only newly added data to avoid redundant reprocessing, and loaded processed outputs into Google Looker Studio using Python.
- Designed a centralized real-time dashboard in Google Looker Studio, enabling chemists to compare and analyze sample properties of interest across historical and newly processed data.

- **TSlearn: Time-Series Learning Library**

Los Angeles, CA

Project Lead

Feb 2022 – Aug 2024

- Developed a time-series learning library with core statistical utilities, including autocorrelation and stationarity analysis, alongside functions for transforming time-series data into supervised ML and DL problems.
- Analyzed, modeled, and forecasted datasets such as daily Yahoo Finance stocks and monthly airline passengers using AR, MA, ARMA, ARIMA, MLP, CNN, and LSTM models.

- **Catalyst Robotics**

Remote/Los Angeles, CA

Co-Founder & Technical Lead

Apr 2021 – Aug 2024

- **Perception and Navigation Team Lead:** Applied computer vision (CV) techniques (object detection, tracking, semantic segmentation, road/lane detection) using Python, NumPy, OpenCV, and TensorFlow. Researched open-source models (KITTI, YOLO, Kaggle), implemented 2 urban semantic segmentation models, evaluated simulation platforms like CARLA, taught foundational computer vision to 7 interns, and collected rural datasets to answer the research question: "How do CV Models for Autonomous Driving Fare in Rural Areas?"

- **Biomedical Team Lead:** Researched and implemented 3 open-source facial pulse estimation models, deployed them to obtain readings over time, processed and standardized sampled time features using Python, pandas, and NumPy, and visualized the resulting pulse signals with Matplotlib. Identified next steps to compare with ground-truth pulse readings, build forecasting models, and evaluate model strengths and weaknesses.
 - Recruited and onboarded 22+ interns from multiple universities across business research and development, biomedical, perception and navigation, drone, and mechanical design teams.
 - Conducted over 300+ customer discovery interviews through programs including NSF I-Corps (Zap, Boom, Regional) and the Min Family Challenge at USC to evaluate product-market fit, and contributed to early prototype development that helped raise \$200,000 in funding.
- **CSCI-544: Natural Language Processing** Los Angeles, CA
Graduate Course Project *Jan 2024 – May 2024*
 - Implemented NLP models for sentiment analysis, sequence decoding, and named entity recognition, including Perceptron, Support Vector Machines, Logistic Regression, Multinomial Naive Bayes, Feedforward Neural Networks, CNNs, Hidden Markov Models, BLSTM, GloVe-enhanced BLSTM, and BLSTM-CNN models.
 - Built text classification and sentiment analysis pipelines for 250K Amazon reviews using data cleaning, feature extraction, scikit-learn, and PyTorch.
 - **CSCI-535: Multimodal Probabilistic Learning of Human Communication** Los Angeles, CA
Graduate Course Project *Jan 2023 – May 2023*
 - Implemented individual projects focusing on inter-rater agreement, emotion classification (incorporating early fusion, late fusion, and hyperparameter tuning), fairness and bias analysis, and statistical hypothesis testing using Fleiss' Kappa and t-tests.
 - Leveraged machine learning techniques to extract and analyze audio, text, and visual modalities across solo, dyadic, and multiparty communication settings to optimize multimodal models.
 - **CSCI-567: Machine Learning** Los Angeles, CA
Graduate Course Project *Aug 2021 – Dec 2021*
 - Implemented foundational ML algorithms including K-nearest neighbors, regression, classification, neural networks, K-means clustering, and Hidden Markov Models, with emphasis on linear algebra, calculus, probability, statistics, and optimization.

PROGRAMMING SKILLS

Languages: Python, R, HTML, CSS

Machine Learning/NLP: PyTorch, Transformers, RAG, LLM-as-a-Judge, NER, sentence-transformers, scikit-learn, Keras

Data/Scientific Computing: pandas, NumPy, SciPy, statsmodels, STATA, rpy2, Jupyter Notebook

Tools/Infrastructure: UF HiPerGator, Docker, GitHub Actions, Protege, OpenCV, Django, Google Looker Studio, vector databases

FELLOWSHIPS, SCHOLARSHIPS, AWARDS

- Herbert Wertheim College of Engineering Dean's Research Award (\$32,000/year) Fall 2024 – Fall 2025
- GEM Fellow at UF 2024 (\$80,000) Fall 2024 – Present
- GEM Fellow at USC 2021 (\$80,000) Fall 2021 – Spring 2024
- Alice Carson Tisdale Honors College Scholarship (\$42,000) Fall 2017 – Fall 2020
- Intel Scholars Program and Scholarship (\$5,000) Fall 2020
- Claflin University - Claflin Carolina Cluster Career Pathways Initiative (\$2,400) Summer 2020
- Claflin Consortium Enabling Cybersecurity Opportunities and Research (\$5,000) Summer 2019
- The Leadership Alliance Summer Research Early Identification Program (\$2,000) Summer 2018
- South Carolina Independent Colleges and Universities (\$2,000) Apr 2017 – Jan 2018

- President's List, Dean's List, and Honor Roll (CU) 9 Semesters
- Campus Honoree for the South Carolina Student Service Achievement Award Fall 2018
- Certification from the BB&T Emerging Leaders Program Fall 2017
- Inducted to Sigma Nu Tau Entrepreneurship Honor Society Fall 2017

LEADERSHIP AND SERVICE

- **Corporal, United States Marine Corps – Fox Company 4th LAR Eastover, SC** Led 14+ Marines to achieve mission success through job proficiency and physical fitness, resulting in performance improvement. Acquired transferable skills (ie: communication, adaptability, self-discipline, etc) for enhanced team versatility.
- **National Society of Black Engineers (NSBE) Founder Archive Collection Team (FACT)**
Collaborated with 1 of the NSBE Founders and 8 other team members to review historical founding documents. Sought to improve web archive interface with accurate information, a more friendly user design, a robust searching algorithm, and additional features.
- **Chapter Vice President and Pre-College Initiative Chairperson, CU (NSBE)** Guided a chapter of 20+ members, coordinating executive board meetings and organizing diverse events such as coding workshops and hackathons for general body members and K-12 students.
- **MSc Mentor, Google x USC Viterbi AI Community Project** Served as a MSc mentor to underrepresented undergraduates at USC, focused on teaching AI to the local Los Angeles community.