Bryan Ramirez-Gonzalez

Los Angeles, California | (626) 391-0502 | bryanram2024@gmail.com | linkedin.com/in/bryanrg22 | github.com/bryanrg22 | bryanram.com

EDUCATION

The University of Southern California **Los Angeles, CA**
Bachelor of Science in Computer Science, Merit Scholar, Honors Engineering Research Track
- Relevant Coursework: Data Structures, Algorithms, Object-Oriented Programming, Embedded Systems, Discrete Mathematics, Linear Algebra

SKILLS

Programming Languages: C, C++, Python, Java, JavaScript/TypeScript, SQL, OCaml
Frameworks/Tools: LangChain, PyTorch, OpenCV, NumPy, Pandas, Flask, FastAPI, REST APIs, Matplotlib, React + Vite, Tailwind CSS, Firebase (Auth, Firestore, Hosting, Cloud Functions), Google Cloud Run, Docker, Git & GitHub, GitHub Actions (CI/CD), AWS, Linux, Bash

EXPERIENCE

Jane Street **New York, NY**
Undergraduate Fellow (UNBOXED 2024; FOCUS 2025 — Discovery Fellowships) Jul 2024; May 2025
- Selected 1 of 37 (UNBOXED '24) and 1 of 14 (FOCUS '25); Wrote over 150+ lines of SQL for data analysis and awarded a $2,000 scholarship.
- Completed 30+ hours of coursework in statistics, market-making activities, and computer programming, achieving 15% return improvement in fast-paced market simulation games under mentorship of researchers and software engineers.

USC Information Sciences Institute (ISI) — HUMANS Lab **Los Angeles, CA**
Undergraduate Research Intern — LLM-assisted AI for TikTok Eating-Disorder Dataset (EDTok) [arxiv] August 2024 - May 2025
- Published EDTok, an ethics-compliant, multimodal (video + text) TikTok dataset of 43,040 videos and 577,071 comments (2019–2024), by building an ingestion/curation pipeline using the TikTok Research API with PykTok.
- Raised dataset precision by 24% using a two-stage filter—weak-supervision rules + LLM-assisted zero-shot relevance classification (Google Gemini) with prompt engineering validated via manual review samples (~99% on 200 videos ; 100% Eating Disorder-related on 300 videos).
- Surfaced platform-scale insights on dataset (537M views, 79.9M likes, 962k shares across 10.9k users) via longitudinal analysis, BERTopic topic modeling, and multi-label emotion classification, linking themes to affect (e.g., recovery to optimism/joy; body-image to fear/sadness).

University of Southern California — Melady Lab **Los Angeles, CA**
Undergraduate Research Intern — Interpretable AI for Image–Text Misinformation Detection July 2024 - August 2024
- Improved out-of-context image–text detection to 68% accuracy (AUC ≈ 0.73) on NewsCLIPpings by adding a learned query ranker over pre-trained Vision-Language Models (CLIP, BLIP-2) with CLIP-retrieved hard-negatives to create challenging pairs using PyTorch.
- Trained a lightweight 4-class evidence ranker that fuses image, caption, and query embeddings (CLIP/BLIP-2) to score evidence supportiveness.
- Outperformed transfer-learning baselines (fine-tuned CLIP) using a feature-based approach: frozen CLIP/BLIP-2 encoders with zero-shot checks feeding a learned query ranker, improving accuracy by 5.6% (62.6% to 68.2%) and AUC by 0.058 (0.67 to 0.73) on NewsCLIPpings.

ACTIVITY / EXTRACURRICULAR - 3x Hackathon Winner

- **Awards:** 2025 Caltech HackTech Winner, 2024 HackHarvard Winner, 2024 AstroHacks Winner, 2024 FTC Robotics Regional Semifinalist
- **Selective Programs & Campus:** 2025 D. E. Shaw Connect Fellowship, 2025 Two Sigma New Seekers Summit, 2025 Susquehanna (SIG) Discovery Day; **Software Developer @ USC Quant Club** (QuantSC), USC's Premier Startup Incubator (LavaLab), USC Hackathon Team

PUBLICATIONS - [Google Scholar]

- C. Bickham, ***B. Ramirez-Gonzalez***, M.D. Chu, K. Lerman, E. Ferrara. *EDTok: A Dataset for Eating Disorder Content on TikTok*, **ICWSM 2025**

PROJECTS

Basis *(Award-Winning AI Workflow Automation Platform @ **Fall 2025 USC LavaLab Winner** - github.com/bryanrg22/Basis_Info)*
Co-Founder & Lead Developer September 2025 - Present
- Built a full-stack AI-assisted cost segregation workflow platform for residential-focused firms that ingests appraisal PDFs and large photo sets, applying YOLO, CLIP, Places365, and OpenAI Vision plus IRS-grounded RAG for asset classification, achieving 95% accuracy in early testing.
- Improved IRS audit-ready accuracy and speed by combining semantic search, deterministic rules, confidence-aware outputs, and engineer-in-the-loop checkpoints with parallel processing across room/object/asset stages, acquiring 2 paying users from top-5 firms.

Swerve *(Award-Winning Agentic Procurement Platform Using LangChain Agents @ **Caltech Hackathon** - github.com/bryanrg22/swerve)*
Lead Developer April 2025
- Reduced procurement cycle time with "Hugo," an event-driven LangChain tool-using agent that routes queries through OpenAI GPT-3.5-turbo (intent + tool selection) and o4-mini (multi-step reasoning) to generate structured, audit-traceable low-stock and reorder recommendations.
- Implemented Python/Flask APIs with CRUD to ingest CSVs, CAD files, and uploads into a structured Firestore (orders, parts, inventory, sales, supply) and wired Slack automation to surface real-time low-stock insights; Invited by Dryft to SF Neo offices for post-hackathon collaboration.

Lambda Rim *(Full-Stack Statistical & ML Hub for NBA Fantasy Sports Betting - github.com/bryanrg22/lambda-rim / lambdarim.com)*
Founder & Sole Developer January 2025 - Present
- Engineered a full-stack forecasting platform (Poisson, Monte Carlo, GARCH) for NBA sports betting picks, achieving a 78%+ win rate and profit growth from **$10 to $3,000** in documented runs shipped via React with a Flask API in Docker on Cloud Run with GitHub Actions CI/CD.
- Cut analysis time from 15 min to seconds by building an automated OCR to prediction pipeline (OpenCV + Tesseract, NBA API, Web-Scraped Injury Reports) with ML-derived probabilities and OpenAI ChatGPT rationales using requests + pdfplumber, nba_api, pandas + NumPy.
- Designed an event-driven system on Google Cloud (GCP) (Cloud Scheduler to Cloud Functions to Cloud Run) and a read-optimized Firestore schema (Active, History, Users) with a settlement pipeline that auto-archives results, reducing redundant reads and improving page latency.