

(909) 993-3077  
Los Angeles, CA  
lorenzosca7@gmail.com

# Lorenzo Scaturchio

Data Scientist / AI Consultant

Portfolio: lscaturchio.xyz  
github.com/gr8monk3ys  
linkedin.com/in/lorenzo-scaturchio

---

## SKILLS

<b>Languages</b>	Python, SQL, Mojo, R
<b>Tech Stack</b>	Zsh, Pytorch, Langchain, FAISS
<b>MLOps</b>	Git, CI/CD, Docker, Google Cloud (GCP), AWS, A/B testing

---

## WORK EXPERIENCE

### GENERATIVE AI ENGINEER

June 2024 - Present

*Sizzle* *Los Angeles, CA*

- Actualize full pipeline for data processing, model training, and deployment of both nutritional factors and macro nutrients of any given recipe.
- Assimilate AI solutions into Google Cloud services using VertexAI to streamline operations and improve data analysis capabilities from custom REST API endpoint, resulting in easily scalable solution.

### AI/ML FREELANCER

August 2022 - Present

*Upwork, Fiverr* *Los Angeles, CA*

- Develop and deploy custom AI/ML models for various clients, improving business processes and outcomes using both open-source and paid LLMs, resulting in an average efficiency increase of 25%.
- Engineered end-to-end MLOps solutions, including CI/CD pipelines, model monitoring, and infrastructure automation, attaining a 98% customer satisfaction rate.
- Collaborate with cross-functional teams to integrate AI solutions into existing systems, enhancing overall efficiency and functionality, diminishing operational costs by up to 20%.

### BIO-INFORMATICS DATA ANALYST

May 2020 - August 2021

*Joint Genome Institute* *Berkeley, CA*

- Utilized R with a cross-functional team of researchers to construct workflows for genetic homology comparison, slashing processing time by 90%.
- Constructed a model based on plant species datasets for microbial interactions, reducing data dimensionality and achieving 98% accuracy.
- Created and optimized genome reference pipelines for model, decreasing processing time from 4 hours to 10 minutes.
- Showcased and presented technologies to over 1000 end clients and senior leadership members in yearly user meeting.

### RESEARCH SCIENTIST

January 2020 - May 2021

*Vice Lab* *Merced, CA*

- Devised and performed Bash scripts to run hydro-power models on a proprietary supercomputer, succeeding a 50% reduction in processing time through asynchronous execution deployed with Slurm.
- Drove creation of a data visualization web application for hydro power models operated with Vue.js, increasing user engagement by 25% and improving data interpretation.
- Optimized model implementation and resolved bottlenecks using Cython, further cutting down execution time by 20%.

---

## EDUCATION

### Masters of Science in Applied Data Science

August 2022 - May 2024

*University of Southern California, Viterbi* *Los Angeles, CA*

### Bachelors of Science in Computer Science & Engineering

August 2018 - May 2022

*University of California, Merced* *Merced, CA*

---

## PROJECTS

### TALKER

February 2024 - May 2024

- Implemented and collaborated on an open source teaching assistant RAG leveraging Ollama2 with a FAISS knowledge base, answering students' questions based on class criteria, syllabus, and slides, increasing student engagement by 30%.
- Included features of providing multi-modal results to queries from classrooms such as YouTube videos with video time queues based on question context, enhancing learning efficiency by 25%.

### BLOGBOT-AI

January 2024 - March 2024

- Constructed an automated blog content generation system using OpenAI's GPT-3.5-turbo model and Langchain
- Accomplished workflows for generating SEO-optimized blog titles, descriptions, and detailed sections
- Utilized Python and ConversationBufferMemory for maintaining context and orchestrating tasks