

# GABRIEL EFFANGHA

Email: gabrielnoah129@gmail.com

LinkedIn: [\(17\) Gabriel Effangha | LinkedIn](#)

GitHub: [Geff115 \(Gabriel Effangha\)](#)

Mobile: +2349053171386

## PROFESSIONAL SUMMARY

Software Engineer with experience building scalable web applications, data-driven systems, and machine learning solutions. Skilled in Python, TypeScript, and cloud platforms, with applied projects in bioinformatics and AI/ML. Passionate about solving complex problems and delivering production-ready systems.

## EDUCATION

**ALX Africa**  
Software Engineering

Nairobi, Kenya  
**Nov 2023 - Jan 2025**

**Stanford University, Department of Genetics**  
Fundamentals of Data Science in Precision Medicine & Cloud Computing, AI/ML

Stanford Data Ocean, Remote  
**May - July 2025**

## SKILLS SUMMARY

- **Languages:** Python, JavaScript, SQL & NoSQL, C, R, TypeScript
- **Frameworks/Libraries:** Flask, Django, FastAPI, NextJS, Pandas, Numpy, Scikit-Learn, Matplotlib, PyTorch, TensorFlow
- **Databases/Tools:** PostgreSQL, MySQL, MongoDB, Supabase, Prisma, SQLAlchemy, Redis, Docker, Git
- **Platforms:** Visual Studio Code, Jupyter Notebook, Vi / Emacs, Hugging Face
- **Soft Skills:** Proficient in Communication, Problem-solver, Time management & organization, Team player, Adaptability

## WORK EXPERIENCE

**In-House Group | Software Engineer | [LINK](#)**

**November 2024 - August 2025**

- Designed and implemented a monthly user matching algorithm using Prisma and TypeScript, achieving >90% matching efficiency across approved members.
- Built an enhanced scoring system with weighted specialties/interests, anti-starvation, prioritization, and recency-aware matching to ensure fairness and inclusivity.
- Engineered scalable batch processing logic capable of handling 500+ users per cycle without performance degradation.
- Developed smart randomization with deterministic test-mode sorting, balancing reproducibility during testing with diversity in production matches.
- Automated match notifications and feedback workflows, sending 100+ personalized emails per cycle with batched delivery for performance and rate-limit compliance.
- Implemented robust retry and advisory lock mechanisms, guaranteeing fault-tolerant execution under database or networking instability.
- Created odd-user and unmatched-user handling strategies, including admin fallback pairing and priority boosts in future cycles, reducing member exclusion.
- Integrated feedback survey scheduling post-match to gather engagement insights, helping improve user experience and retention.
- Enhanced system observability with structured logging, match statistics, and efficiency tracking, enabling quick debugging and transparent reporting.

## PROJECTS

**Breast Cancer Prediction Model (Logistic Regression)**

**September 2025**

- Built and evaluated a predictive model for breast cancer patient survival outcomes (Alive/Dead) using Logistic Regression.
- Applied preprocessing (OneHotEncoding for categorical data, StandardScaler for numerical features) via scikit-learn pipelines.
- Trained and tested the model on a 70/30 split, achieving ~89% accuracy on both train and test datasets.

**Stanford Data Ocean Projects (Bioinformatics, AI/ML)**

**May 2025 - July 2025**

- Built CNN-based image classification models using TensorFlow to classify medical images (e.g., Pneumonia detection).
- Visualized transcriptomic and epigenetic data using heatmaps, volcano plots, and PCA in Python and R.
- Queried and analyzed the 1000 Genomes project using PyAthena and AWS S3, and implemented heat maps to draw insights from the data.
- Applied dimensionality reduction techniques (PCA, UMAP, t-SNE) on omics data.

- Built and deployed the YouTubeIntel platform for YouTube channel discovery and analytics, through a robust batch channel discovery algorithm, capable of handling 5-15 million channels.
- Developed a robust credit-based monetization system with tiered pricing and Redis-backed rate limiting.
- Designed a secure authentication system with JWT, Google OAuth, device/session tracking, and HTML email flows (Resend).
- Built an AI-powered channel analytics API to give deeper insights into videos and channel metadata.
- Implemented WebSockets for real-time notification updates.
- Integrated an eventlet background worker with Gunicorn to handle asynchronous operations for non-blocking threads.

**Image Classifier CNN Model: “Bird or Forest”****June 2025**

- Collected and curated image datasets of birds and forests using DuckDuckGo Search API.
- Preprocessed images to 192 x 192 resolution, handled corrupted files, and ensured clean data ingestion.
- Trained a ResNet-18 CNN model using FastAI, achieved a 0% validation error, and no signs of overfitting.
- Successfully deployed model using .pth format for Hugging Face compatibility.

**Brand Identity API****April 2025**

- Built an API to extract brand elements (logo, palettes) from websites using Playwright and ColorThief.
- Integrated GPT-4o for intelligent tagline and identity generation.
- Added Redis caching and rate-limiting for scalability and performance optimization.
- Tested extensively on dynamic, JS-heavy sites like Airbnb.com.

---

**CERTIFICATES****Fundamentals of Data Science in Precision Medicine and Cloud Computing – Stanford University****May - July 2025**

- Developed a comprehensive understanding of omics data and various stages involved in data analysis.
- Queried health data using R and Python programming tools to draw insights, and visualized the insights using heat maps and volcano plots.
- Trained CNN and linear regression models on health data for predictions of patterns or trends, as well as curating insights from datasets.

**Software Engineering – ALX Africa | [CERTIFICATE](#)****February 2025**

- Mastered fundamental Python syntax, proficiently utilizing control flow, loops, functions, and data structures.
- Acquired expertise in procedural programming paradigms, system design, and associated logical concepts, enhancing capabilities.
- Mastered object-oriented programming in Python, including MVC and MTV web frameworks.
- Mastered fundamental C programming and implemented data structures like Linked Lists, Binary Trees, Stack & Queues, Pointers, Arrays, and Heaps.