Kenneth V. Domingo

⚲ Bulacan, Philippines   ✉ hello@kvd.studio   🔗 kvd.studio   in kvdomingo   🐙 kvdomingo

Education

University of the Philippines-Diliman   *Quezon City, PH*
B.S. in Applied Physics (Major in Instrumentation)   *Aug 2015 – July 2020*

- **Thesis**: Compressive sensing: Applications from 1-D to N-D
- **Research Lab**: Video and Image Processing, Instrumentation Physics Laboratory
- **Relevant Coursework**: Computer Methods in Physics, Physical Electronics, Control Systems Approach to Physics Modeling, Signal/Video/Image Processing, Neural Networks and Machine Learning

Work Experience

AI Software Engineer   *Cologne, DE*
eyva.ai   *Dec 2024 – present*

Software Engineer / Data Engineer   *Taguig City, PH*
Thinking Machines Data Science   *Jan 2023 – Oct 2024*

- Led a cross-functional team of 7 engineers to architect and develop the UNICEF Giga DataOps Platform ↗, a fully open-source ↗ master data management platform built in coordination with Giga ↗, a UNICEF-ITU joint initiative whose mission is to connect all schools to the internet by 2030. V1.0 of the Platform was launched in September 2024 and was presented at the Open Source Summit ↗ in Vienna later that month. At launch, the Platform managed data for over 2.1M school records across 141 countries, with near-real-time connectivity data for 93k schools updating every few minutes. The Platform was also being leveraged by 2 downstream applications with 5k monthly users.
- Developed a data analytics platform for a Thailand-based agricultural company: optimized the data lake setup for near-realtime updates of logistic movements and validation of the delivery of burnt produce. Daily tasks involved orchestration and scheduling of ELT tasks and SQL queries on BigQuery via Dagster, and automating replication of multi-cloud deployment environments via Terraform.
- Developed GenAI applications ↗ for several enterprise clients in the financial sector, leveraging Azure OpenAI. Daily tasks involved designing responsive chat interfaces, and image processing to remove non-text elements in financial documents in order to prepare them for vector embedding for use in downstream RAG applications.
- Developed the SDG Impact Tool ↗ for Gold Standard ↗ to digitize their process of reporting, validating, verifying, and tracking the contribution of project activities to the UN Sustainable Development Goals. Daily tasks involved backend development in Django, frontend development in React, and automating infrastructure provisioning across multiple Google Cloud deployment environments using Terraform.
- Developed an internal geospatial analytics platform for a Philippine-based telecommunications company: involved in migrating user interface components to a modern tech stack and increasing test coverage.

Software Developer (Freelance)   *Remote*
Stevn Books   *Nov 2022 – Jan 2023*

- Developed the API and infrastructure for an inventory management and profit calculator system deployed on DigitalOcean: integrated with PayMongo to process customer payments and AWS Cognito as a B2C identity provider.

Software Engineer / Data Analyst   *Taguig City, PH*
Cobena Business Analytics and Strategy (Demand Science)   *Nov 2020 – Nov 2022*

- Developed a unified frontend enabling users to experiment with deployed ML models being developed in-house.
- Developed the frontend for an internal data lakehouse platform allowing users to perform queries on access-controlled data, apply filters, save filters, and export their search results.
- Guided the adoption of MLOps principles within the AI/ML team.
- Spearheaded the company-wide adoption of modern frontend tooling using Figma and TypeScript.
- Developed a supply chain performance survey tool for UK-based marketing consultancy Ideas and Action ↗: led backend development with Django, Redis, and PostgreSQL, deployed to an AWS environment.
- Developed Gateway, a B2B SaaS geospatial analytics platform: led weekly code reviews to maintain quality standards and optimize the team's delivery.
- Developed internal tooling for a multinational FMCG client: initiated the migration from manual testing to automated API testing using Python scripts, reducing testing time from 2 days to 30 minutes.

Backend Developer (Volunteer)   *Quezon City, PH*
DetectPH   *Mar 2020 – May 2020*

- Developed the API for a COVID-19 contact-tracing mobile application written with Express and MongoDB.

Bioinformatics Intern *Quezon City, PH*

Philippine Genome Center *Jan 2020 – Feb 2020*

- Underwent training in bioinformatics; developed a command-line tool and web application for designing primers for site-directed mutagenesis.

Student Researcher *Quezon City, PH*

Instrumentation Physics Laboratory *Oct 2017 – July 2020*

- Research focused on processing and analysis of videos, images, and other signals using conventional signal processing methods, as well as experimental methods such as compressive sensing and machine learning.

Student Assistant (Contract) *Quezon City, PH*

Philippine Coral Reef and Mangrove Remote Sensing Project *Dec 2018 – Jan 2019*

- Performed ground-truth validation of images involving corals classified by a deep neural network.

Projects

Giga DataOps Platform *July 2023 – Sept 2024*

- An open-source master data management platform developed with Giga, a UNICEF×ITU initiative to connect all schools around the world to the internet by 2030.
- Giga Sync ↗ (GitHub ↗) is the data ingestion portal, allowing users—such as country representatives, government agencies, and NGO partners—to upload school data in a variety of formats. Here, users can also monitor the progress of ingestion and quality check reports of their previously uploaded data.
- Giga Dagster ↗ (GitHub ↗) is the heart of the platform, responsible for orchestrating data tasks such as ingestion, transformation, and data quality checks. Data is stored in the data lake in the Delta Lake open table format, allowing various access methods and maintaining a full change history via the change data feed.
- Giga Data Sharing ↗ (GitHub ↗) allows open access to the school master data, subject to role-based access control. The API is compliant with the Delta Sharing Protocol, allowing users to interact with the data using the Delta Sharing client libraries.
- Giga Superset ↗ (GitHub ↗) is the business intelligence platform, allowing data analysts to explore, visualize, and extract insights from the data stored in the data lake.
- Giga Datahub ↗ (GitHub ↗) is the data catalog, allowing for the discovery of data stored in the data lake.
- Giga Trino (GitHub ↗) is the query engine, allowing Delta Lake tables to be queried as SQL for business intelligence applications such as Superset.
- Giga Monitoring (GitHub ↗) provides for the monitoring and logging of all deployed applications on UNICEF Giga's Kubernetes cluster.
- **Backend**: Python, FastAPI, PostgreSQL, Redis
- **Frontend**: TypeScript, Vite, React, Zustand, TailwindCSS, Carbon Design System, Ant Design, Tanstack Query, Tanstack Router, Swagger UI
- **DevOps**: GitHub, Azure DevOps, Prometheus, Grafana, Sentry
- **DataOps**: Dagster, Apache Spark, Apache Hive Metastore, Delta Lake, Delta Sharing, Datahub, Trino, Apache Superset
- **Infrastructure**: Docker, Helm, Azure Data Lake Storage, Azure Kubernetes Service, Azure Entra ID

Farm Data Analytics Platform *May 2024 – Aug 2024*

- A data platform for validating the delivery of burnt agricultural products using various data sources such as vehicular data, geospatial farm data, GISTDA roads, and NASA FIRMS active fire data. The platform used Google BigQuery and Google Cloud Storage as the primary data warehouse.
- **Backend**: Node.js, Express, PostgreSQL, BoxySAML, Prisma
- **Frontend**: TypeScript, Swagger UI
- **DataOps**: Dagster, Google BigQuery, Polars
- **DevOps**: GitHub, GitHub Actions, Terraform
- **Infrastructure**: Docker, GCP Cloud Run, GCP Cloud SQL, GCP Compute Engine, GCP HA VPN, GCP Cloud Router, GCP Cloud Load Balancing, GCP Cloud Storage, Azure Enterprise Application, AWS VPC, AWS S3

ResearchAI *July 2023 – Aug 2023*

- Various AI chat applications ↗ developed for enterprise clients, focusing on retrieving information from sensitive and PII documents.
- **Backend**: Python, FastAPI, PostgreSQL, Weaviate, PineconeDB
- **Frontend**: TypeScript, React, Next.js, T3, tRPC, Prisma, Auth.js, Zustand, TailwindCSS, Shadcn UI
- **DataOps**: Dagster, NumPy, Pandas
- **AI**: Azure OpenAI, Langchain, OpenCV
- **DevOps**: GitHub, GitHub Actions, Terraform

- ○ **Infrastructure**: Docker, Azure Kubernetes Service

SDG Impact Tool *Mar 2023 – June 2023*
- ○ Web application ↗ developed for Gold Standard to digitize their process of reporting, validating, verifying, and tracking the contribution of project activities to the UN Sustainable Development Goals.
- ○ **Backend**: Python, Django, PostgreSQL, Node.js, Express
- ○ **Frontend**: TypeScript, Vite, React, Redux, TailwindCSS, Material UI
- ○ **DevOps**: GitHub, GitHub Actions, GCP Cloud Build, Terraform
- ○ **Infrastructure**: GCP Cloud Run, GCP Cloud SQL, Auth0

Single Customer View *Feb 2023 – Mar 2023*
- ○ A reusable single customer view web app.
- ○ **Backend**: Python, FastAPI, SQLAlchemy, Snowflake
- ○ **Frontend**: TypeScript, Vite, React, Redux, Material UI, TailwindCSS
- ○ **DevOps**: GitHub, GitHub Actions, Terraform
- ○ **Infrastructure**: Docker, AWS EC2, AWS ECS, Cloudflare

Telecom Geospatial Analytics Platform *Jan 2023 – Feb 2023*
- ○ An internal geospatial analytics platform.
- ○ **Backend**: Python, Django, PostgreSQL, Elasticsearch
- ○ **Frontend**: TypeScript, Vite, React, Redux, TailwindCSS, Chakra UI
- ○ **DevOps**: GitHub, Jenkins, Terraform
- ○ **Infrastructure**: Docker, GCP Cloud Run, GCP Cloud SQL, GCP Compute Engine

Stevn Books *Nov 2022 – Jan 2023*
- ○ Web app ↗ for inventory management and profit calculator.
- ○ **Backend**: Python, Django, PostgreSQL
- ○ **Frontend**: JavaScript, React, Webpack, Material UI, Swagger UI
- ○ **Infrastructure**: AWS Cognito, DigitalOcean, Jenkins, Docker

AI Playground *Mar 2022 – Nov 2022*
- ○ A unified frontend to allow regular users to experiment with deployed ML models being developed in-house.
- ○ **Backend**: Python, Django
- ○ **Frontend**: TypeScript, Vite, React, Redux, Material UI, Ant Design
- ○ **Infrastructure**: AWS ECS, AWS Cognito, Sentry

Search Console *July 2021 – Nov 2022*
- ○ A web app to allow business users to search for specific data in the internal data lakehouse.
- ○ **Backend**: Go, Chi, Redis
- ○ **Frontend**: TypeScript, React, Webpack, Redux, Material UI, GraphQL
- ○ **Infrastructure**: AWS ECS, AWS Cognito

Gateway 3 *Jan 2021 – Mar 2021*
- ○ A B2B SaaS geospatial analytics platform.
- ○ **Backend**: Python, Django, PostgreSQL
- ○ **Frontend**: JavaScript, React, Webpack, Konva, Material UI

Supply Chain Performance Indicator *Jan 2021 – June 2021*
- ○ A supply chain performance indicator survey tool developed for UK-based market research firm Ideas & Action.
- ○ **Backend**: Python, Django, PostgreSQL, Redis
- ○ **Frontend**: JavaScript, React, Webpack, Material UI
- ○ **Infrastructure**: AWS EC2, AWS RDS, NGINX

PrimerDriver *Jan 2020 – Feb 2020*
- ○ CLI ↗ and web application ↗ for automating the design of mutagenic PCR primers for site-directed mutagenesis.
- ○ **Backend**: Python, Flask, BioPython, Sphinx
- ○ **Frontend**: JavaScript, React, Webpack, Material Design Bootstrap, Click
- ○ **Infrastructure**: Heroku

Hannibot *June 2020 – July 2021*
- ○ A Discord bot ↗ that keeps track of upcoming K-pop releases, and crossposts Twitter and VLIVE media to K-pop-oriented Discord guilds (GitHub ↗).

- **Backend**: Python, Flask, PostgreSQL, Discord API, Twitter API, VLIVE API
- **Frontend**: JavaScript, Vite, React, Bootstrap
- **Infrastructure**: Heroku, Vercel

Homelab *May 2023 – present*
- Personal homelab running on a high-availability Kubernetes cluster hosted on-premises.
- **Infrastructure**: Linux, Proxmox, Kubernetes, K3s, Ansible, Raspberry Pi
- **Networking**: pfSense, Traefik, Bind9, Pi-hole

Banyuhay *Oct 2024 – present*
- A geospatial web app ↗ to allow users within Metro Manila to locate nearby restrooms with bidets (GitHub ↗).
- **Backend**: Python, FastAPI, PostgreSQL, PostGIS
- **Frontend**: TypeScript, Vite, SvelteKit, LibreMaps, Protomaps
- **Infrastructure**: Docker, Kubernetes, Helm, Cloudflare
- **Data**: OpenStreetMaps

KenGPT *Feb 2023*
- Yet another ChatGPT clone ↗ (GitHub ↗.
- **Backend/Frontend**: TypeScript, React, Next.js, TailwindCSS, PostgreSQL
- **Infrastructure**: Supabase, Vercel
- **AI**: OpenAI, Anthropic

LSFM *Apr 2022*
- A clone ↗ of the LE SSERAFIM Digital Souvenir system (GitHub ↗).
- **Frontend**: TypeScript, Vite, React, TailwindCSS, Shadcn UI, WebAssembly, ffmpeg
- **Infrastructure**: Vercel

NCOVenience *Mar 2020 – June 2020*
- COVID-19 PH Cases Tracker Map (GitHub ↗).
- **Backend**: Python, Flask, NumPy, Pandas, BeautifulSoup
- **Frontend**: JavaScript, React, Webpack, Material Design Bootstrap, Chart.js, Mapbox
- **Infrastructure**: Heroku

Certifications

Azure AI Engineer Associate *May 2024*
- Microsoft ↗

Professional Cloud Architect *July 2021*
- Google Cloud ↗

Publications

Compressively sampled speech: How good is the recovery? Sept 2020

K. V. Domingo, M. N. Soriano

Proceedings of the Samahang Pisika ng Pilipinas 38 SPP-2020-4C-04 ↗

Frequency domain reconstruction of stochastically sampled signals based on compressive sensing May 2019

K. V. Domingo, M. N. Soriano

Proceedings of the Samahang Pisika ng Pilipinas 37 SPP-2019-PB-38 ↗

Technologies

Backend: Python (Django, FastAPI, Flask), Go (Fiber), Node.js (Express, Hono, Bun)

Frontend: TypeScript (React, Next.js, Svelte, Astro), UI (Shadcn, TailwindCSS, Material UI)

Databases: OLTP (PostgreSQL, SQLite, Redis), OLAP (DuckDB, Trino), Warehouse (BigQuery, Snowflake) Vector (Chroma, Weaviate)

Cloud Platforms: Google Cloud Platform, AWS, Azure, Vercel

DevSecOps: Git, Docker, Kubernetes, Helm, Snyk, ArgoCD, GitHub, GitHub Actions, GitHub Advanced Security, Azure DevOps

Data: Dagster, NumPy, (Geo)Pandas, Polars, Delta Lake, Apache Spark

AI/ML: Ollama, LangChain, TensorFlow, Keras, Scikit-learn, OpenCV

Infrastructure: Linux (Debian, Ubuntu), IaC (Terraform, Ansible), Monitoring (Prometheus, Grafana), Load Balancing (Traefik, NGINX)

Business Intelligence/Dashboarding: Apache Superset, Metabase

Miscellaneous: Microsoft Excel, Google Sheets, LaTeX

AI/ML: Ollama, LangChain, TensorFlow, Keras, Scikit-learn, OpenCV

Infrastructure: Linux (Debian, Ubuntu), IaC (Terraform, Ansible), Monitoring (Prometheus, Grafana), Load Balancing (Traefik, NGINX)

Business Intelligence/Dashboarding: Apache Superset, Metabase

Miscellaneous: Microsoft Excel, Google Sheets, LaTeX