

# Dan Kim

US Citizen | +1 (201) 970 – 8008 | dankim1@seas.upenn.edu | dankim444.github.io | LinkedIn | Github

## EDUCATION

University of Pennsylvania, School of Engineering and Applied Science | Philadelphia, PA

Aug 2022 - May 2026

Bachelor of Science in Engineering in Computer Science, Concentration in AI, Minors in Math and Data Science

GPA: 3.77/4.0

**Relevant Coursework:** Artificial Intelligence, Database & Information Systems, Big Data Analytics, Data Structures & Algorithms, Machine Learning, Computer Systems, Linear Algebra, Probability & Statistics, Business: Product Design

**Activities:** Wharton Undergraduate Data Analytics Club, Wharton Korean Business Society, Penn Apps, CS Tutor and TA, Research

**Awards:** Penn Undergraduate Research Mentorship Scholar, Global Research and Internship Program Funding Award, Dean's List

## TECHNICAL SKILLS

**Languages:** Java, Python, Typescript, JavaScript, Swift, C, C++, OCaml, SQL; **Web Development:** Node.js, React.js, Flask, HTML/CSS, MySQL, MongoDB; **Data Science:** Pandas, NumPy, PySpark, PyTorch; **Tools:** AWS, Azure, Docker, Git, Linux, Apache Airflow, JIRA

## PROFESSIONAL EXPERIENCE

Computational Social Science Lab at Penn

Philadelphia, PA

Research Software Engineer - Full-Stack, Data Science, NLP

June 2024 - Present

- **Expanded user database on main experiment platform by 225%** by redesigning frontend to enable multilingual support in 10 languages, using **Typescript** and **react-i18next** for seamless UI adaptation based on users' language preferences.
- **Improved research insights** by enhancing backend for tracking and analysis of language-specific data with **Node.js**, **MySQL**, and **Sequelize**, optimizing retrieval and storage of multilingual statements and user responses.
- **Reduced experiment scaling efforts by 65%** by implementing a pipeline that trains a large language model (LLM) on extracted media data, generates statements, and standardizes them using **OpenAI** and **Python**.
- **Ideated and engineered automated text processing and translation workflow** with **Github Actions**, **Amazon Translate**, and **Pandas** to process/translate statement files in new languages and resolve machine translation inconsistencies.
- **Enhanced user flow and experience** by resolving critical bugs including errors in feedback and statement submissions, UI styling, and faulty sign-in/sign-out issues.
- **Contributed to increased testing coverage from 15% to 95%** by implementing automated tests for user flows, end-to-end scenarios, and localization checks with **Cypress**, ensuring system reliability and minimizing errors.
- **Promoted proper code maintenance** with best practices (**communication**, **code reviews**, **documentation**, **PRs**, etc.).

Ruta N Medellín

Medellín, Colombia

Software Engineer Intern - Full-Stack

June 2023 – August 2023

- **Reduced admin time to manage employee data by ~75%** by designing and developing a full-stack employee management web application with **Figma**, **React.js**, **Node.js**, **Express.js**, and **MongoDB**.
- **Developed entire frontend** by implementing pages for registration, uploading payroll/documents, and email handling.
- **Architected RESTful API in backend** to facilitate seamless **CRUD** functionality for **185+** employee records.
- **Ensured role-based data access across 10+ distinct roles** by integrating user sign-in/sign-out and **JWT** authentication.
- **Collaborated in an agile manner** with cross-functional team sprints, participating in **100%** of daily scrums.

## PROJECTS

Gratify – iOS Mobile Application | [Project Link](#)

Swift, SwiftUI, OpenAI API

- Built and deployed a digital journal app, supporting 40 users on their journeys toward daily gratitude and self-improvement.
- Integrated major features like daily reminders, profile settings and management, intuitive and minimalistic UI, daily entry user flow, and a personalized AI assistant that summarizes weekly and monthly highlights based on user entries.

Q/A Study Buddy and Notetaker - RAG Application | [Project Link](#)

Python, Flask, Langchain, ChromaDB, OpenAI

- Developed a RAG application designed to assist busy college students in studying more effectively.
- Implemented a Q/A study-bot that queries and answers questions from a ChromaDB vector knowledge database and generates and exports personalized notes into pdf formats.

CNNs and ML Techniques for Fraud Detection in Financial Transactions | [Project Link](#)

Pandas, Scikit-learn, PyTorch

- **Trained a neural network with 96% accuracy** on **50,000+** bank statements to classify and predict fraudulent transactions by utilizing key data science practices such as **data pre-processing**, **data visualization**, **feature engineering**, and **modeling**.