Dan Kim

US Citizen | +1 (201) 970 – 8008 | dankim1@seas.upenn.edu | Website | Linkedin | Github

EDUCATION

University of Pennsylvania, School of Engineering and Applied Science | Philadelphia, PA *May 2026*
Bachelor of Science in Engineering in Computer Science, Concentration in AI, Minors in Math and Data Science **GPA:** 3.8/4.0

Relevant Coursework: Artificial Intelligence, Database and Info Systems, Big Data Analytics, Data Structures & Algorithms, Computer Systems, Programming Languages & Techniques, Linear Algebra, Probability & Statistics, Business: Product Design

Activities: Research, CS Tutor and TA, Wharton Data Analytics Club, Wharton Korean Business Society, Develop for Good, Penn Apps

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

TECHNICAL SKILLS

Languages: Java, Python, Typescript, JavaScript, Swift, C, C++, OCaml, SQL; **Web Development:** Node.js, React.js, Redux, HTML5, CSS, MySQL, PostgreSQL; **Data Science:** Pandas, spaCy, PySpark, PyTorch; **Tools:** Git, AWS, Azure, Docker, Linux, Apache Airflow, CI/CD

PROFESSIONAL EXPERIENCE

Computational Social Science Lab at Penn **Philadelphia, PA**
Research Software Engineer – Full-Stack, Data Science, NLP *Jun 2024 – Present*

- **Expanded user database on main experiment platform** by redesigning frontend to enable multilingual support in 10 languages, using **Typescript** and **react-i18next** for seamless UI adaptation based on users' language preferences.
- **Enhanced 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 for improved research insights.
- **Reduced experiment scaling efforts** by implementing a pipeline that leverages the GPT 4-o model to generate statements and normalize them using **OpenAI** and **Python.**
- **Ideated and engineered automated text processing and translation workflows** with **Github Actions, Amazon Translate**, and **Pandas** to process/translate files in new languages and resolve machine translation inconsistencies.
- **Enhanced user flow and experience** by resolving critical bugs including UI styling, errors in feedback and statement submissions, and faulty sign-in/sign-out issues.
- **Increased testing coverage** by integrating 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 *Jun 2023 – Aug 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**.
- **Led frontend development** by creating core pages with backend integration through **Axios-based API calls**.
- **Architected RESTful API** in backend to facilitate seamless **CRUD** functionality for **185+** employee records.
- **Ensured role-based data access** across **8** distinct roles by integrating user sign-in/sign-out UI with **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*

- **Collaborated in a team of 3 to build a journal app** where users can rate and log their gratitude and visualize their progress through weekly, quarterly, and yearly reports.
- **Led backend development,** integrating **Core Data** for data persistence and a **Network Manager** to handle GPT requests and responses, boosting user retention and enabling personalized **AI summaries** of collected entries.

Multimedia Q&A Assistant - RAG Application | *Project Link* *Python, Langchain, Beautiful Soup, OpenAI*

- **Engineered a Q&A Retrieval-Augmented Application (RAG)** that allows users to extract information from various media sources such as PDFs, audio files, YouTube videos, and web pages.
- **Implemented an NLP pipeline** leveraging document parsing, web scraping, embeddings, and indexing for efficient retrieval.

Machine Learning for Fraud Detection in Credit Card Transactions | *Project Link* *Pandas, Scikit-learn, Seaborn*

- **Developed a machine learning pipeline** for credit card fraud detection using **feature engineering**, **PCA** for dimensionality reduction, and **SMOTE** to address imbalanced data, achieving **98% accuracy** and **90% recall** with an **XGBoost** model.