

# Can Karakoc

[cankarakoc@berkeley.edu](mailto:cankarakoc@berkeley.edu) | [linkedin/can-karakoc](https://www.linkedin.com/in/can-karakoc/) | [github/can-karakoc](https://github.com/can-karakoc) | [www.cankarakoc.com](http://www.cankarakoc.com) | **Work Authorization:** Green Card

UC Berkeley CS & Data Science graduate with end-to-end full-stack, data, and ML engineering experience, grounded in a design-driven, human-centered approach. What excites me most is the intersection of AI, biology, and product, where my engineering, design, and biology interests come together.

## TECHNICAL SKILLS

**Languages:** Python, JavaScript/TypeScript, SQL, R

**Software Engineering:** Full-stack web development, backend and API design, databases, testing, deployment

**Data & Machine Learning:** Data pipelines and analysis, applied ML, deep learning, embeddings and vector search, unsupervised methods

**Computational Biology:** Structural and sequence analysis, single-cell and genomic data, protein language models, bioinformatics tooling

**Design & Product:** Human-centered design, user research, prototyping, product and UX design

## EXPERIENCE

**Software Engineer (Contract) - Cendra** | May 2026 - July 2026 | Remote

- Built an end-to-end GTM automation pipeline solo on a Supabase data layer, a scheduled Fly.io cron worker, and a Next.js dashboard on Vercel, processing ~120k scraped short-term-rental vendor records through validation, enrichment, and CRM sync.
- Designed a GPT-4o-mini agent that classifies each vendor's site as an active short-term-rental operator using mined UI heuristics and dead-site detection, with a human-in-the-loop approval table the agent learns from, validated against 80 hand-labeled operators.
- Engineered a cost-optimized, 8-provider enrichment waterfall for full-name resolution that hands off to Clay for decision-maker contact data, adding guardrails like fallback-name detection and broken company-name repair to lift hit rate on small businesses.
- Built a Smartlead agent that mines guest-communication complaints from Google reviews for personalized outreach hooks, and shipped an operator dashboard with funnel analytics, per-stage monitoring, and provider credit and hit-rate tracking with automatic blocking.

**Design & Content Manager - UC Berkeley Development Engineering** | June 2024 - Aug 2025 | Berkeley, CA

- Owned design and content for the Master of Development Engineering department's web presence, conducting user research with students and redesigning deprecated pages around their pain points and needs through a human-centered approach.
- Designed Figma mockups and prototypes, presenting to operational staff and the department Dean for feedback and approval, then built and shipped them in WordPress and Elementor while developing reusable, CMS-driven design structures.
- Improved accessibility, interactivity, and visual liveliness across existing pages, and produced marketing material for department seminars and talks.

**Data Analysis Intern - Velocity Inc.** | May 2023 - Aug 2023 | Istanbul, Turkey

- Built an unsupervised learning pipeline for an alternative credit-scoring product that segmented users from proprietary behavioral data, engineering temporal and behavioral features such as commute routes, inferred home location, dwell time, and device ownership.
- Applied PCA for dimensionality reduction and compared multiple clustering methods with systematic hyperparameter tuning, evaluating segments with internal validation metrics and longitudinal stability analysis across weekly cohorts.
- Designed interactive Tableau dashboards with heatmaps and geospatial views that surfaced segment profiles and behavioral patterns, informing how the product scored users beyond conventional financial history.

## PROJECTS

**Protein I/O - Structural Biology Analysis Workbench** | Personal Project | [App](#) | [GitHub](#)

- Built a full-stack computational structural-biology platform for analyzing protein structures, pairing a Mol\* 3D viewer in a Next.js and TypeScript frontend with a Python and FastAPI backend, deployed on Vercel and Render.
- Implemented a wide range of published structural-biology methods in-house with NumPy, SciPy, Gemmi, and RDKit, including interaction classification, Shrake-Rupley buried surface area, IDDT and DockQ model-quality scoring, LIGSITE binding-pocket detection, ligand validation, and antibody CDR numbering across four schemes.
- Designed a stateless, local-first backend that computes every metric from coordinates on CPU with no external binaries, with fetch-and-analyze integration for RCSB PDB, AlphaFold DB, UniProt, and ChEMBL.

**Single-Cell RNA-seq Analysis & Cell-Type Classification** | Personal Project | [GitHub](#)

- Built an end-to-end single-cell RNA-seq pipeline on the 10X Genomics 3k PBMC dataset with Scanpy, taking raw counts through quality control, normalization, highly-variable-gene selection, PCA, and Leiden clustering, then annotating clusters into immune cell types using canonical marker genes.
- Compared logistic regression, random forest, a PyTorch neural network, and an autoencoder-embedding model for immune cell-type classification, evaluating each on accuracy and macro-F1 to weigh the trade-offs between interpretability and model complexity.

**Protein Embedding Search** | Personal Project | [App](#) | [GitHub](#)

- Built a full-stack semantic protein search tool that embeds sequences with the ESM2 protein language model and serves nearest-neighbor matches over an adaptive FAISS index that switches between exact and approximate search, with UniProt fetch, sequence validation, and a cached embedding store.

## EDUCATION

**University of California, Berkeley**

B.A. Computer Science; B.A. Data Science | Aug 2021 - May 2025 | GPA: 3.8/4.0

**Emphasis:** Computational Methods in Molecular & Genomic Biology

**Relevant Courses:** Introduction to Machine Learning for Computational Biology; Introduction to Computational Molecular and Cell Biology; Machine Learning; Designing, Visualizing, and Understanding Deep Neural Networks; Data Engineering; Optimization Models.

## LEADERSHIP & DESIGN

- Design Consultant - Berkeley Innovation** | Jan 2023-May 2024 - Delivered product redesigns for real clients at UC Berkeley's first student-run human-centered design consultancy through user research, competitive analysis, and prototyping, including a redesign of Playbook's Publish feature with advanced editing and real-time collaboration tools.
- Teaching Assistant - Human-Centered Design DeCal** | May-Aug 2023 - Led weekly sessions for 60+ students on design methodology, research synthesis, and prototyping.