

# Alp Efe Karalar

karalar.alpefe@gmail.com • (814) 996-8979 • [alpkaralar.com](http://alpkaralar.com) • [linkedin.com/in/alp-efe-karalar/](https://linkedin.com/in/alp-efe-karalar/) • [github.com/EfeKaralar](https://github.com/EfeKaralar)

## EDUCATION

---

**The Pennsylvania State University** | *B.S. in Computer Engineering* | **University Park, PA**

Dec 2024

**Relevant Courses:** OS Design, Comm. Networks, Computer Architecture, Data Structures & Algorithms, Neural Networks

**Senior Capstone:** Advanced Vehicle Team, Perception Department - Developed real-time computer vision models for autonomous vehicle navigation, achieving 3rd place overall in SAE AutoDrive Challenge II Competition

## TECHNICAL SKILLS

---

**Languages:** Python, C/C++, JavaScript, TypeScript, Java, SQL, Bash, R, MATLAB, PowerBI

**Technologies:** Linux (Arch, Ubuntu, Debian), PyTorch, n8n, AWS, Docker, PostgreSQL, Git, Nginx, systemd

**Concepts:** Networking (TCP/IP, HTTP/HTTPS, DNS), Neural Networks, CI/CD, REST APIs, embedded systems

## EXPERIENCE

---

**The Pennsylvania State University** | *Research Assistant* | **University Park, PA**

Dec 2025 - Present

- Build Python-based data matching pipelines to link >10,000 firms across a billion job postings using entity resolution techniques
- Process job posting datasets to extract hiring patterns, temporal trends, and employer-position relationships for research analysis
- Deploy fuzzy matching algorithms to resolve entity duplicates and ensure accurate company identification across large datasets
- Scale data infrastructure to handle decades of longitudinal employment data for labor market analysis in the finance sector

**DementiAnalytics** | *Software Engineer* | **University Park, PA**

Feb 2025 - Present

- Construct end-to-end ML-powered platform for dementia diagnosis, building React/Expo mobile application and Flask backend API
- Establish REST API integrating speech analysis models with patient data, processing audio biomarkers across 5-stage ML pipeline
- Implement HIPAA-compliant data handling architecture and authentication system for sensitive medical research data
- Design multiple web applications for research data visualization, model performance monitoring, and public-facing project websites

**Bakkal Co.** | *Software Engineer Intern* | **Pleasanton, CA**

Jun 2023 - Nov 2023

- Engineered comprehensive testing infrastructure for React Native mobile application supporting iOS and Android platforms
- Identified and documented critical bugs during testing cycles, working with senior engineers to implement fixes before release
- Integrated test suite with CI/CD pipeline, enabling automated test execution on pull requests

**The Pennsylvania State University** | *Learning Assistant (CMPEN270: Digital Design)* | **University Park, PA**

Jan 2022 - May 2024

- Troubleshoot technical issues for 250+ students per semester, creating documentation and video tutorials to improve student success
- Managed course operations, delivered lectures, and created course content while leading a team of learning assistants

## PROJECTS

---

[SevanBOT - RAG Pipeline for Writer Persona](#) | Python, FastAPI, Qdrant, Embeddings, LLM APIs, BM25

- Orchestrated a scrape→Markdown→semantic chunk pipeline turning ~1,300 articles into ~7,100 retrieval-ready contextual chunks
- Built a modular retrieval stack: dense (Qdrant + OpenAI), sparse (BM25), and hybrid fusion (RRF/weighted), with a CLI comparator
- Integrated contextual retrieval +RAG answering with Claude prompts and conversation memory behind a FastAPI chat backend

[TafsirBot – AI Quranic Commentary Assistant](#) | Python, RAG, Qdrant, n8n, OpenAI Embeddings, Claude API, Docker

- Built ayah-scoped RAG ingestion pipeline (clean → chunk → embed → upsert) over several Tafsir corpus with citable scripture metadata
- Implemented hybrid BM42 sparse + dense vector retrieval over Qdrant with serverside RRF fusion for verse reference & semantic queries
- Designed n8n RAG workflow with intent classification, Ayah resolution, and citation-enforced responses across Telegram, web, and X

[Linux Kernel Mailing List \(LKML\) Dashboard - AI-Powered Kernel Development Tool](#) | Python, Flask, React, JS, SQLite

- Engineered platform processing 600+ daily LKML emails through automated parsing, thread reconstruction, and AI summaries
- Designed Flask REST API with 10 endpoints and SQLite FT55 full-text search, achieving sub-second performance

[Linux RAID Storage System \(mdadm\)](#) | C, Memory Management, LRU Cache, Systems Programming

- Developed Linux mdadm-inspired RAID emulator integrating 16 disks into a virtual block device with mount, read, write operations
- Implemented write-through LRU caching mechanism, achieving 87.9% cache hit rate and 30% reduction in average read latency

[Research Summarizer - AI Academic Paper Analysis](#) | FastAPI, React, PostgreSQL, Docker, Nginx

- Deployed web application integrating arXiv API, PyPDF2, and multiple LLM providers for paper summarization
- Implemented FastAPI backend with SQLAlchemy ORM supporting PostgreSQL and summary caching for performance optimization

[Custom LISP Interpreter](#) | C, Parser Combinators, REPL, Language Design/Implementation

- Built a fully functional Lisp interpreter featuring parser combinators, a recursive parser, REPL, S-expressions, lambdas, and closures