

Archer Zhang

Email: ez806@nyu.edu | Tel: 646-906-7909 | Website: <https://archertakesitez.github.io/>

SKILLS

Programming Languages: Python, Java, SQL, HTML, CSS, JavaScript, TypeScript, Shell, R

Expertise: Data Structures, Algorithms, Full-stack Development, Cloud Computing, AWS, Databases, Linux, Unit Testing, Git, Distributed Systems, Docker, Kubernetes, CI/CD, System Design, Test Automation, Operating Systems, AI, LLMs

EDUCATION

New York University, New York, NY

Sep. 2023 – May. 2025

Master of Science in Data Science

GPA: 3.833/4.0

Coursework: Programming for Data Science | Cloud Computing | Big Data | Machine Learning | Database Systems

Brandeis University, Waltham, MA

Aug. 2019 – May. 2023

Bachelor of Science in Computer Science, Bachelor of Arts in Economics, Minor in Mathematics

GPA: 3.749/4.0

Coursework: Advanced Java | Data Structure | Operating Systems | Algorithms | Deep Learning | Natural Language Processing

PROFESSIONAL EXPERIENCE

Kineviz Inc.

San Francisco, CA (remote)

Software Development Engineer Intern

Jun. 2024 – Dec. 2024

- Incorporated **LLMs (Large Language Models)** into a **chat GUI** via **Python** and **JavaScript**
- Developed a **Cypher** query parser with ANTLR in Python, incorporating Named Entity Recognition (**NER**) and **vector search** to validate and enhance Cypher queries, enabling advanced search capabilities in **Neo4j** graph databases
- Wrote automated tests for Kineviz's graph visualization tool (GraphXR) using **Cypress**
- Built a website leveraging **FastAPI** that employs **LLMs** to translate plain texts into user-defined **JSON** schemas

YUSUR Technology Co., Ltd

Beijing, China

Software Development Engineer Intern

May. 2021 – Aug. 2021

- Utilized **Apache Spark** and implemented **Shell** and **Python** scripts to process the Iris dataset on **distributed and parallel systems**, benchmarking the performance of our self-developed Kernel Processing Unit (KPU)
- Built **data pipelines** in **Python** to preprocess and clean data using **Prefect** and **Mage**
- Visualized our findings with **D3.js**, **Matplotlib**, and **Seaborn**, and made presentations using these data visualizations

State Key Laboratory of Software Development Environment, Beihang University

Beijing, China

Software Development Engineer Intern

Dec. 2020 – Feb. 2021

- Queried datasets from **SQL** databases (**MySQL**, **PostgreSQL**) and **NoSQL** database (**MongoDB**)

ACADEMIC PROJECTS

PaperGist: A Cost-Efficient Cloud-Native Research Paper Summarization Platform

Spring 2025

- Built a serverless **cloud-native** platform on **AWS** using Lambda, API Gateway, EC2 (for inference), SQS, EventBridge, S3, and DynamoDB with content-based deduplication to enable scalable, cost-efficient research paper summarization

Deck-to-CPT: AI-Driven Reimbursement Code Discovery for HealthTech Start-Ups

Fall 2024

- Built a **Python**-based web application with **HTML/CSS** for processing PDF pitch decks and returning relevant Current Procedural Terminology (CPT) codes using **AI**
- Automated **web scraping** to retrieve detailed CPT code information from healthcare websites
- Applied Named Entity Recognition (**NER**) to extract key information from PDFs and utilized Retrieval-Augmented Generation (**RAG**) for accurate CPT code recommendations

Fixplainer: Failure Explainer for Multiple Object Tracking (MOT)

Spring 2024

- Built a **GUI** tool to generate SHAP-based visual explanations of tracking failures in MOT video frames using YOLOv8

PUBLICATION

GraphBERT: Bridging Graph and Text for Malicious Behavior Detection on Social Media. Published by ICDM-2022.

<https://archertakesitez.github.io/static/assets/papers/GraphBERT.pdf>

Fair Graph Representation Learning via Diverse Mixture-of-Experts. Published by The ACM Web Conference 2023.

https://archertakesitez.github.io/static/assets/papers/Representation_learning.pdf