Yuxuan (Wayne) Wang

https://waynewangyuxuan.github.io/personal_site/ || waw009@ucsd.edu || 9295838556

EDUCATION SUMMARY:

University of California, San Diego, Jacobs School of Engineering, San Diego, CA   September 2025 - June 2027
Master of Science, Computer Science

New York University, Tandon School of Engineering, Brooklyn, NY   September 2021 - May 2025
Bachelor of Science**,** Computer Science & Math

TECHNICAL SKILLS:

C/C++ || **Java** || **Python** || **Linux** (Kernel, Shell/Bash) || **Multi-Threaded/Process Programming** (Threads, Memory Management, IPC) || **Agent/Workflow Development** (LangGraph, n8n, Dify) || **Information Retrieval** (query processing, dense/sparse re-ranking) || **Big Data** (Kafka, Flink, Spark) ||

EXPERIENCES:

Software Engineering Intern - ByteDance   June 2025 – September 2025
Global Monetization Product and Technology - TikTok Ads Diagnosis
(Java, Python, SQL, Kafka, Hive, Spark, Big Data, TikTok Ads)

- Built the **Copilot knowledge & retrieval layer**: Feishu-doc ETL → Elasticsearch with hybrid text+vector recall (ES-KNN), plus a conversation store and access-control via bot/user tokens.
- Productionized entry points & plugins: exposed Copilot via AdDebug Web Chat and Lark Bot through an RPC→HTTP gateway; added a **plugin framework to call HTTP/RPC/X-Table/Auto Diagnoser**.
- Prototyped **"Chart AI" analysis for funnel/benchmark** charts and shipped an anomaly analysis base (periodicity shifts, spikes, drop-to-zero, trend jumps) used in Copilot explanations.
- Led **migration of 20+ ads pipelines (Kafka→Hive/HDFS; ClickHouse→Doris)**: refactored HiveSQL/schema, added query-parity & p50/p95 latency harnesses; shipped **Java adapters to gate multi-DC cutovers with automated QA.**
- Designed Intent-vs-Actual Budget Delta and **standardized budget utilization across stages**; re-architected **2** primary and upgraded **13** downstream pipelines (**~450 TB backfill**) so spend gaps are transparent to users in delivery reporting.

Research - New York University   June 2024 – May 2025
Evaluation of Graph-Based Vocabulary Mismatch Solution in Information Retrieval
Supervised by Professor Torsten Suel @ New York University, Tandon School of Engineering
(Information Retrieval, Search Engine, Query Processing, Database, Linux, HPC, Python)

- Conducted a systematic study on impact of **seed quality graph-based expansion** in LADR (Lexically-Accelerated Dense Retrieval); designed a **retrieval pipeline supporting multiple sparse retrieval models** as initial seeds sets.
- Implemented a graph expansion module for candidate documents using KNN/HNSW; incorporated a vector dimension masking mechanism (PRFDIME) that **constructs semantic masks via pseudo-relevance feedback** to enhance query representation and improve reranking effectiveness.
- Demonstrated that reranking **using only the top-3000 graph-expanded passages** via Bi-Encoder/Cross-Encoder models. Achieve comparable performance to full-corpus reranking, improving recall and reducing computational cost.

Data Engineering Intern - CITIC Poly Fund (Guangzhou)   June 2023 - August 2023
(NLP-based News ETL, Web Scraping, Python, AWS, NoSQL)

- Built an end-to-end investment-intelligence platform for the chip/EDA sector—ingesting **Wind + multi-site crawlers** into a Kafka→Lambda/DynamoDB/S3 pipeline with de-dup, structured field extraction (amount/round/investors), and company-profile enrichment—and **delivered an insights layer** (event/keyword mining, BERT sentiment, T5 summarization, weekly trends) with a Streamlit dashboard that replaced manual spreadsheets with a department-wide weekly brief.

PROJECTS:

- **Notate** — AI-native capture app that listens for trigger text (e.g., !! buy milk !!) and **auto-converts snippets into structured actions** (contacts, calendar, reminders, Map, Email).
- **LifeTrace** — Automated time analytics that unifies calendar + notes, refreshes daily, and visualizes multi-week → multi-year **trends with AI-driven categorization**.
  ShadowDash — C++ DSL that constructs Ninja build DAGs in-memory; validated on zlib/LLVM; **~40%–2× faster than .ninja parsing**.
  ContextWIN (paper) — Whittle-index **Mixture-of-Experts** for contextual RMAB with theoretical guarantees.