Qiao-Yun (Katharina) Cheng

qiaoyc2@uw.edu • (206)823-6913 • http://www.linkedin.com/in/qiaoyc

EDUCATION

University of Washington   **Seattle, WA**
Bachelor of Science in Informatics, Concentration: Data Science   June 2026
Annual Dean's List (2022 – 2025)   Cumulative GPA: 3.81/ 4.0
Relevant Coursework: Databases & Data Modeling | Client-Side Development | Server-Side Development | Introduction to Search & Recommendation Systems | Text Mining | Advanced Methods in Data Science | Data Structure & Algorithms | Database Design & Management | Introduction to Machine Learning | Introduction to Artificial Intelligence

SKILLS

Technical: R, Python, Java, SQL, HTML, CSS, D3, JavaScript, Swift, Tableau, Figma, DevOps, Azure, MongoDB
Languages: Bilingual (Mandarin, English)

WORK EXPERIENCE

University of Washington   **Seattle, WA**
Teaching Assistant   Sep 2024 - Jun 2025
- Conduct lab sessions for INFO 103, facilitate discussions on social media history, privacy, consent, and free speech
- Grade assignments and provide feedback to students on their performance, ensuring clarity and accuracy
- Support students applying concepts to automation projects like Bluesky, Reddit, or Discord bot

RELATIVE EXPERIENCE

Study Buddy   **Seattle, WA**
Project Lead & Full-Stack Developer   Oct 2025 - Dec 2025
- Led a 4-person team and produced the system architecture + schema + endpoint documentation for a UW-only study-partner platform with verified login and profiles
- Built homepage + invitation workflow and implemented session middleware to resolve current user identity
- Co-built real-time chat (WebSockets + database-backed chat history) for matched users

GameLensDB (NoSQL Video Game Metadata & Trends)   **Seattle, WA**
Data Engineer (Web Scraping & MongoDB)   May 2025 - Jun 2025
- Scraped MobyGames with Selenium + BeautifulSoup, exporting JSON and loading 1,100+ titles into MongoDB
- Standardized 8 key attributes across 1.1k+ MongoDB docs, improving filterability and downstream analyses
- Designed aggregation pipelines to surface genre/platform trends, developer/publisher activity, and score vs. business model, packaged into concise takeaways

Evaluating Comment-Post Relatedness in the Climate Reddit Community   **Seattle, WA**
LLM Prompting Engineering (Data Science and Modeling)   Nov 2024 - Dec 2024
- Implement a binary text classification model to evaluate the relatedness of comments to post titles in the Climate subreddit, using Logistic Regression, BERT embeddings, and LLM prompting
- Trained and fine-tuned large language models (chatgpt3.5) with both zero-shot and few-shot prompting to predict the relevance of comments in climate-related Reddit posts
- Crawled the subreddit and structured data around 600 Reddit posts, facilitated the training of models

Recipe Finder: Recipe Discovery and Recommendation Platform   **Seattle, WA**
Search and Recommendation System Designer & Front-End Developer   Nov 2024 - Dec 2024
- Designed a platform with information retrieval capabilities, leveraging information crawled from the internet
- Utilize a mapping function to suggest related queries and improve user experience over time
- Integrated PyTerrier and the BM25 ranking model to ensure efficient and relevant recipe retrieval

Plate Pal: Client-Side Website Development Project   **Seattle, WA**
Project Leader & Front-End Developer   Jun 2024 - Aug 2024
- Developed a website to calculate daily calorie intake and generate personalized meal plans
- Designed responsive interfaces and implemented dynamic functionality for user interactions
- Integrated Google Firebase for real-time database management