

Szu-Ju Chen

(724) 807-5424 | szjuichen@gmail.com | szjuichen.github.io | github.com/szjuichen

Education

- Carnegie Mellon University**, M.S. in Intelligent Information System at SCS Aug. 2026 – Dec. 2027
- **Courses:** Advanced Natural Language Processing, AI Agents, Introduction of Question Answering
- National Taiwan University**, B.S. in Computer Science and Information Engineering Sep. 2021 – Jun. 2025
- **Courses:** Machine Learning, Applied Deep Learning, Computer Vision, Information Retrieval and Extraction
 - GPA: 4.15/4.3; Dean's List Award (Ranked 1st in Spring 2025 & Fall 2024); Fu Bell Scholarship (Fall 2024 & Fall 2021)

Work Experience

- Software Engineer**, Google – New Taipei City, Taiwan Aug. 2025 – Aug. 2026
- Delivered Pixel Clock Timer production mechanism and user interface to set up timer with keypads and preset buttons
 - Constructed Pixel Calculator end-to-end evaluation dashboard with 15+ automated metric scripts across 2M+ user data
 - Implemented in-app parser evaluating AI-generated Python formulas for complex expression handling in Pixel Calculator
- Machine Learning Scientist Intern**, Appier – Taipei City, Taiwan Sep. 2024 – May 2025
- Resolved one-epoch overfitting issues in CTR prediction architectures, boosting accuracy by 5%
 - Fine-tuned production ranking models and optimized feature engineering pipelines across large-scale advertising data
- Software Engineer Intern**, Google – New Taipei City, Taiwan Jun. 2024 – Sep. 2024
- Built an Android package for updatable onboarding flows, enabling dynamic feature updates and immediate bug fixes
 - Designed modular data transfer workflows and refactored components into reusable libraries; passing 7/7 internal tests
- Student Training Engineering Program Intern**, Google – New Taipei City, Taiwan Jun. 2023 – Sep. 2023
- Integrated multiple network FTP speed testing features into internal modem testing Android app
 - Delivered high-precision network performance analytics under 5G environments, fully adopted by internal teams

Publications

- Retrieving the Right Law: Enhancing Legal Search with Style Translation**, SIGIR2025 Short Paper
- Szu-Ju Chen, Jing Jin, Sheng-Lun Wei, Chien-Hung Chen, Hsin-Hsi Chen* | 10.1145/3726302.3730246
- Designed style translation module and retrieval algorithms for legal question answering using related Taiwanese laws
 - Trained a Bi-Encoder model and fine-tuned Taiwan-LLaMA-3B using original user queries and LLM-rewritten queries

Research Experience

- Multimodal Speech Processing Lab**, Dr. Carlos Busso – Carnegie Mellon University Aug. 2026 – Present
- Designed agentic multimodal frameworks to identify emotional responses and risk profiles from social media content
 - Integrated 5 data streams: facial expressions, audio, acoustic signals, comments, and lyrics to cross-modal representations
- MACHI Program in DATA Lab**, Dr. Xia Ben Hu – Rice University Feb. 2025 – Mar. 2025
- Applied dynamic routing algorithms to multi-agent visual reasoning, cutting latency by 50% and raising accuracy by 7%
 - Awarded 2nd Place out of all participants in the 2025 Rice MACHI Poster Competition
- Information Retrieval Lab**, Dr. Pu Jen Cheng – National Taiwan University, Taiwan Sep. 2023 – Jan. 2025
- Engineered a pipeline integrating LightGCN trained on millions of user-track data with LLM-generated expansions
 - Built a hybrid recommender system for genre and keyword queries, earning the Department Research Excellence Award
 - Investigated autoregressive image retrieval utilizing VQ-VAE indexing and contrastive learning representations

Projects

- Visual Grounding & Image Inpainting Pipeline** Dec. 2024
- Designed a computer vision pipeline integrating visual grounding with image inpainting for indoor object replacement
 - Fine-tuned and benchmarked diffusion-based models against PowerPaint and BrushNet on a custom-curated dataset
- Konnect – Cross-Platform Fan Interaction App** Sep. 2023 - Jun. 2024
- Led development of a full-stack React Native app integrated with Google Maps, GCP backend, and real-time data sync
 - Achieved MVP stage and passed initial testing with user posting, activity tracking, and content sharing features
- Chinese Grammatical Error Correction with Seq2Seq Models** Apr. 2023 - Jun. 2023
- Benchmarked BART, mT5, and Taiwan-LLaMA models to evaluate zero-shot and fine-tuned correction performance
 - Finetuned google/mt5-small on open-source corpora, evaluating with self-designed Exact-Match and M^2 metrics

Skills

Programming Languages: C++, C, Python, Java, JavaScript, Kotlin, SQL

Machine Learning: LLMs, recommendation systems, information retrieval, computer vision, RAG, multi-agent

App Development: Android apps, React Native, GCP, Firebase, Git