# SHENG-YEN CHOU

@ shengyenchou@cuhk.edu.hk    in Sheng-Yen Chou    ⌨ FrankCCCCC    % Personal Website    % Blog    % Google Scholar

**Research Interests**

I'm interested in **trustworthy (generative) AI** from **optimization constraint, kernel analysis, uncertainty quantification, reinforcement learning (RL), human feedback (RLHF)**, etc. My goal is to **launch an authentication agency startup to speed up the deployment of reliable autonomous intelligence** in the world.

# EDUCATION

» **Incoming Ph.D. Student in Computer Science**, *Cornell University*          New York, USA, 2024 – 2029 (Expected)
Admitted to Cornell CS Ph.D. program.
» **B.S. Computer Science**, *National Tsing Hua University (NTHU)*          Hsinchu, Taiwan, 2017 – 2022
Changed major from Power Mechanical Engineering in 2019, **Last 60 GPA: 4.1/4.3, Last 2 Year GPA: 4.15/4.3**
Member of (Official) Leadership Development Program at NTHU

# RESEARCH EXPERIENCE

» **Research Assistant**, *The Chinese University of Hong Kong*          Hong Kong, Jul 2022 – Present
- Supervised by Prof. Tsung-Yi Ho and Dr. Pin-Yu Chen
- Proposed a **new backdoor attack on diffusion model (DM)**, called **BadDiffusion on CVPR 2023** with **workshop best paper award** and a universal advanced framework: **VillanDiffusion on NeurIPS 2023** with **workshop oral** (both are **first author**).
- **Invented a new defense: Elijah to secure DMs** and **accelerate backdoor data distillation with 20 to 50 times than SOTA method** with Prof. Chia-Mu Yu. Investigated in watermarking diffusion models with Prof. Sinno Jialin Pan

» **Research Assistant**, *National Ting Hua University*          Taiwan, Feb 2020 – Jun 2022
- Supervised by Prof. Shan-Hung Wu
- **Machine Learning**
  - Derived a new **single level objective from the bilevel optimization problem of GANs** to stabilize and **speed up the training process** based on the **NTKs**. Got **better image quality** with **10x less training time than SOTA GANs and small dataset**.
- **Distributed Database**
  - Implemented the **prototype of the "DependencyAnalyzer" component for the open-source database: VanillaCore**.
  - Built an auto-tuning distributed DB with ML. **Reproduced experiments of the paper: MB2 in the distributed DB: ElaSQL.**
- AAAI'21 Paper Review
- CS565600: Deep Learning Teaching Assistant

# WORKING EXPERIENCE

» **Founder**, *Green Pepper Delivery*          Taiwan, Sep 2018 - Dec 2018
- Built a delivery service with reusable containers with 10000 USD funding from Ching Piao (a B cooperation) and the university.
- I led a team of 8 people to run the delivery service and **served more than 100 orders every day.**

» **IT Internship / Expatriate Software Engineer / Project Manager**, *Ching Piao*          Taiwan, Aug 2018 – Jul 2019
- Built an online rental service running for Pingtung County Government and communicated with partner IT company.
- **Acquired more than 1000 membership within 1.5 months and achieved 200 daily usages.**
- **Created an APP(Ching Piao Rental Service POS) and doubled the speed of the service.**

» **Business Development / Software Engineer Internship**, *Vexanium*          Indonesia, Jul 2019 - Aug 2019
- Conducted market research on Taiwan and Indonesia and analyzed VexGift user data to create DAPP market strategy.
- Developed a plugin for authors to upload article and shared profits to them.

» **Founder / Software Engineer / Business Development**, *LEAFHOPPER.IO*          Taiwan, May 2019 - Dec 2020
- **Created immutable traceability for Dong Ding Oolong tea** with **66000 USD funding from the Lu-Gu township government** and **invited to hold a public tender**.
- I led a team of 7 people and built a tea traceability system with **Postgres, React and, Ethereum (under construction).**

# PUBLICATION

## 👥 Accepted by Conference Proceedings

- Ming-Yu Chung **Sheng-Yen Chou**, Chia-Mu Yu-Pin-Yu Chen Sy-Yen Kuo Tsung-Yi Ho (2023). "Rethinking Backdoor Attacks on Dataset Distillation: A Kernel Method Perspective". In: International Conference on Learning Representations (**ICLR 2024**).
- **Sheng-Yen Chou**, Pin-Yu Chen and Tsung-Yi Ho (2022). "How to Backdoor Diffusion Models?" In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR 2023**), **ICLR 2023 BANDS Workshop Best Paper Award**.
- **Sheng-Yen Chou**, Pin-Yu Chen and Tsung-Yi Ho (2023). "VillanDiffusion: A Unified Backdoor Attack Framework for Diffusion Models". In: Advances in Neural Information Processing Systems (**NeurIPS 2023**), **NeurIPS 2023 BUGS Workshop Oral**.
- Shengwei An **Sheng-Yen Chou**, Kaiyuan Zhang-Qiuling Xu-Guanhong Tao Guangyu Shen-Siyuan Cheng Shiqing Ma Pin-Yu Chen Tsung-Yi Ho Xiangyu Zhang (2023). "Elijah: Eliminating Backdoors Injected in Diffusion Models via Distribution Shift". In: The Association for the Advancement of Artificial Intelligence (**AAAI 2024**), **NeurIPS 2023 BUGS Workshop**.

### Pre-Print

- Yu-Rong Zhang Ruei-Yang Su, **Sheng Yen Chou** and Shan-Hung Wu (2021). *Single-level Adversarial Data Synthesis based on Neural Tangent Kernels*.
- Yu-Shan Lin Ping-Yu Chen, Yu-Xuan Lin **Sheng Yen Chou** Wei-Yu Lin Chao-Wei Lin and Shan-Hung Wu (2021). *Cost-Effective Joint Data Fusion and Transaction Routing for Deterministic Database Systems*.

## HONORS

| | |
|---|---|
| **6th NTHU Garage** (Enrolled our startup: LEAFHOPPER.IO) | Taiwan, 2019 |
| **NTUST Micro Accelerator** (Enrolled our startup: LEAFHOPPER.IO) | Taiwan, 2019 |
| **3rd place of 7th ENTREPRENEUR DAYS** (Won by our startup: LEAFHOPPER.IO) | Taiwan, 2019 |
| **Academic Achievement** (Top 5% students in class with highest GPA) | Taiwan, 2022 |

## OTHER EXPERIENCE

» **Program Member,** *Leadership Development Program at National Tsing Hua University*       Taiwan, 2018 - 2021
- An **official leadership cultivation program based on Project-Based Learning**, sponsored by Mr. Sandy Chau and NTHU alumni. Students will take 9 credits over 3 years and conduct three projects to **boost their leadership skills and impact society**.

» **Consultants,** *Teamie*       Taiwan, 2022
- Built a member-matching platform for everyone who wants to launch side projects.

## SKILLS

| | |
|---|---|
| **Programming Language** | C++, C, Python, Java, JavaScript, TypeScript, Matlab, React, Flutter |
| **Business** | Project/Product Management, Market Research |
| **Language** | Mandarin (Native), Taiwainese (Native), English (Fluent) |
| **Certification** | TOEFL iBT MyBest: 103 |

## PERSONAL PROJECTS

### Implementation of 2V2PL

- Implemented the **2V2PL concurrency protocol on VanillaDB with Java** and **improved the throughput by 5 times than S2PL**.

### EfficientDet

- An **EfficientDet implementation in TF2.0** based on the paper EfficientDet: Scalable and Efficient Object Detection on CVPR'20.

### DRL Collection

- A collection of the implement of classical DRL algorithms with Tensorflow 2.0, including **A3C, A2C, DDQN, and REINFORCE.**

### ML Collection

- Implemeted and derived ML algorithms in Python, including **SVM and VBGMM.**

### Blocked Floyd Warshall With CUDA

- Solved all pair shortest path problem with **Blocked Floyd Warshall algorithm and** parallel on **multi-GPU with CUDA in C++.**

### Mandelbrot Set Generator

- Generated the photo of Mandelbrot set with **MPI, Pthread, OMP, and vectorization in C++.**

### PoW algorithm of bitcoin protocol

- Implemented the **PoW algorithm of Bitcoin in Python.**

### Traceability Platform

- Built up a traceability system and a transparent selling platform with **React, NodeJS, and Postgres**

### Scalable Runner

- A **distributed task executor** with multi-GPU support in Python.

### Gomoku AI

- A Gomoku AI based on **threat space search, Negamax, and MCTS** with C++.