

# CHEN YU-HAO

AI Data Scientist / Engineer



Email: a29015822@gmail.com | Phone: 0930-130-959 | Location: New Taipei City

## 🎯 Core Philosophy & Professional Drivers

Seeking an **AI Engineer** role to leverage hands-on experience in **LLM End-to-End Implementation** and a Systems-Level Thinking approach.

- \*\*Systems Thinker:\*\* Believes the true battlefield has shifted to **System Design, Operational Strategy, and Business Model Creativity**.
- \*\*Pragmatic Execution:\*\* Rigorously vets ideas based on user needs, cost, and feasibility (e.g., evaluating RAG vs. Fine-tuning benefits).
- \*\*Rapid Growth & Responsibility:\*\* Driven by a strong sense of responsibility, translating to a work ethic that is "not afraid of hard work and embraces new challenges head-on." The only fear is stagnation.
- \*\*Technical Translator:\*\* Served as the "translator" between technical/AI strategy and business/cross-functional teams (PMs, UX).
- \*\*Responsible AI:\*\* Committed to creating AI solutions that are technically robust and focus on **"creating genuine value for users"**.

## 💻 Core Technical Stack: LLM End-to-End Practice

Category	Key Technologies & Frameworks
----------	-------------------------------

**LLM Application**	<a href="#">LangChain</a> (Prototyping), <a href="#">LangGraph</a> (Specialized in Agentic Workflows)
**Model Optimization/Alignment**	<a href="#">RAG</a> , <a href="#">Fine-tuning</a> , <a href="#">RLHF</a> (PPO, GRPO, DPO)
**AI System Observability (LLMOps)**	<a href="#">LangFuse</a> (Tracing, Debugging, Monitoring performance and cost)
**Programming/DB/Backend**	<a href="#">Python</a> (Primary), <a href="#">SQL</a> , <a href="#">PgVector</a> , <a href="#">Milvus</a> , <a href="#">ChromaDB</a> , <a href="#">FastAPI</a> , <a href="#">Django</a>
**DL/ML/CV**	<a href="#">PyTorch</a> , <a href="#">Numpy</a> , <a href="#">OpenCV</a> , CNN, ResNet, YOLO, <a href="#">Machine Learning</a> , <a href="#">Deep Learning</a>

## Professional Experience

---

### AI Data Scientist

*Zhan Shi Hua Technology (展市華科技有限公司) | 2025/9 – Present*

- Developing \*\*AI-Agent\*\* and \*\*RAG\*\* related technologies.
- Key Technical Focus: [#AI](#)    [#PyTorch](#)    [#LangChain](#)    [#LangGraph](#)    [#VLLM](#) .

### Associate Engineer (AI Engineer - R&D Alternative Service)

*Industrial Technology Research Institute (ITRI, 財團法人工業技術研究院) | 2023/12 – 2025/8*

- Core member responsible for \*\*AI strategy, research, and product implementation\*\*.
- Executed \*\*GAI model calibration\*\* using advanced RLHF strategies (Fine-tuning, GRPO, PPO, DPO).
- Designed and implemented GAI applications leveraging [RAG](#), [AI-Agent](#), [AI-Workflow](#), and [Prompt Engineering](#).
- Provided technical leadership, offering AI design methodologies and improvement suggestions to department colleagues.

### AI Algorithm Engineer

*Evervision Electronics Co., Ltd. (億力光電股份有限公司) | 2023/3 – 2023/10*

- Proposed and conducted Demo experiments for the company's future development projects.

## Education

---

### **Master of Science in Electrical Engineering**

*National Taiwan University of Science and Technology (NTUST)*

### **Bachelor of Science in Physics**

*Soochow University (東吳大學)*

## Key Projects (Selected)

---

### **ITRI: Advanced GAI Nursing Summary System (POC)**

*2024/6 – 2024/12*

- \*\*Data Pipeline:\*\* Designed and implemented PDF processing, including high-efficiency \*\*de-identification\*\* to ensure data security.
- \*\*AI Core:\*\* Applied ChatGPT-4o or other LLMs for semantic understanding and information extraction from nursing assessment texts.
- \*\*Feature Engineering:\*\* Developed Prompt Engineering to automatically map text content to standardized nursing features (e.g., activity, pain score).
- \*\*Optimization:\*\* Established standardized testing to validate AI results against nurse reviews, iteratively optimizing LLM prompts and matching models.

### **ITRI: Smart Customer Service RAG System Development**

*2025/2 – 2025/8*

- Designed and implemented a multi-domain Q&A system based on the \*\*RAG architecture\*\*.
- Employed the \*\*Milvus\*\* vector database for semantic retrieval, improving answer accuracy.
- Established an \*\*LLM routing mechanism\*\* to automatically determine question type and select the optimal data source.

## ★ Additional Details

- \*\*Desired Titles:\*\* AI Engineer, Algorithm Engineer, Data Scientist, Data Analyst.
- \*\*Target Industries:\*\* Software/Internet, Telecom, Computer/Consumer Electronics Manufacturing, Semiconductor.
- \*\*Remote Work:\*\* Open to remote work.
- \*\*Language Proficiency:\*\* Mandarin (Fluent), English (Intermediate), Taiwanese (Intermediate).

*"If you are looking for an engineer with hands-on LLM experience, pragmatic systems thinking, and who is driven by a strong sense of responsibility and a passion for growth, I would be very excited to speak with you."*