Last modified: April 17, 2026

Pei-Fu (Fred) Guo

r12922217@csie.ntu.edu.tw Homepage Google Scholar +886-919790276

Education

- **National Taiwan University (NTU)** *2024/2 - Present*
  M.S. in Computer Science & Information Engineering (Current GPA : 3.84/4.3)
  – MSLab (Advisor: Shou-De Lin)
  – Relevant Course : Natural Language Processing (A+), Applied Deep Learning (A+)
- **National Taiwan University (NTU)** *2019/9 - 2023/12*
  B.S. in Economics (Overall GPA : 3.89/4.3)
  – Relevant Course : Statistical Learning and Deep Learning (A+), Machine Learning (A)

Publications (97 citations by Apr 2026).

[1] **Pei-Fu Guo**, Ya-An Tsai, Chun-Chia Hsu, Kai-Xin Chen, Yun-Da Tsai, Kai-Wei Chang, Nanyun Peng, Mi-Yen Yeh, and Shou-De Lin. Beyond Facts: Benchmarking Distributional Reading Comprehension in Large Language Models. *ACL 2026 Findings.* **[PDF]**

[2] **Pei-Fu Guo**, Yun-Da Tsai, Chun-Chia Hsu, Kai-Xin Chen, Ya-An Tsai, Kai-Wei Chang, Nanyun Peng, Mi-Yen Yeh, and Shou-De Lin. LiveCLKTBench: Towards Reliable Evaluation of Cross-Lingual Knowledge Transfer in Multilingual LLMs. *ACL 2026 Main Conference.* **[PDF]**

[3] **Pei-Fu Guo**, Yun-Da Tsai, and Shou-De Lin. Why is the LLM unsure? Profiling the Causes of LLM Uncertainty for Adaptive Model and Uncertainty Metric Selection. **[PDF]**

[4] **Pei-Fu Guo**, Yun-Da Tsai, and Shou-De Lin. Benchmarking Uncertainty Metrics for LLM Target-Aware Search. *EMNLP 2025 Findings.* **[PDF]**

[5] Yun-Da Tsai*, Ting-Yu Yen*, **Pei-Fu Guo**, Zhe-Yan Li, and Shou-De Lin. Text-centric Alignment for Bridging Test-time Unseen Modality. *EMNLP 2025 Findings.* **[PDF]**

[6] **Pei-Fu Guo***, Ying-Hsuan Chen*, Yun-Da Tsai, and Shou-De Lin. Towards Optimizing with Large Language Models. *KDD 2024 Knowledge-Infused Learning Workshop.* **[PDF]**

Research Projects

- **Text2DistBench: Benchmarking Distributional Reasoning in LLMs [1]** **Automated Evaluation**
  *Joint Research, MSLab & **UCLA-NLP** (Advisors: **Kai-Wei Chang**, **Nanyun Peng**, Mi-Yen Yeh, Shou-De Lin)*
  – Developed *Text2DistBench*, an automated reading comprehension benchmark for evaluating LLMs' ability to infer distributional knowledge from natural language.
- **LiveCLKTBench: Reliable Evaluation of Cross-Lingual Knowledge Transfer [2]** **Automated Evaluation**
  *Joint Research, MSLab & **UCLA-NLP** (Advisors: **Kai-Wei Chang**, **Nanyun Peng**, Mi-Yen Yeh, Shou-De Lin)*
  – Developed *LiveCLKTBench*, an automated pipeline generating contamination-free QA pairs from time-sensitive sources to rigorously evaluate LLMs' cross-lingual knowledge transfer capability.
- **Decomposing LLM Uncertainty for Adaptive Metric Selection [3]** **Uncertainty Estimation**
  Graduate Research, MSLab (Advisor: Shou-De Lin)
  – Proposed a sequential prompting framework to decompose distinct sources of model uncertainty by quantifying output divergence at each stage, demonstrating that source-specific metrics enhance response failure prediction.
- **Uncertainty Metrics for LLM Inference-Time Search [4]** **Uncertainty Estimation**
  Graduate Research, MSLab (Advisor: Shou-De Lin)
  – Investigated uncertainty metrics for inference-time tree search, revealing that they steer search towards response diversity, suggesting their potential to enhance exploration.
- **Text-Centric Alignment for Test-Time Unseen Modality [5]** **Test-time Generalization**
  Undergraduate Research, MSLab (Advisor: Shou-De Lin)
  – Proposed a text-centric alignment framework that maps unseen modalities into a unified text space, addressing modality mismatch without parameter updates.
- **Towards Optimizing with Large Language Models [6]** **Test-time Generalization**
  Undergraduate Research, MSLab (Advisor: Shou-De Lin)
  – Evaluated model adaptability to novel task structures using unseen synthetic combinatorial optimization problems, revealing that performance degrades significantly as dimensionality increases.

Work Experience

- **APPIER – LLM Research Scientist Intern** **Conversational Sales Agents** *2024/7 – 2025/3*
  AI SaaS company specializing in AI-driven marketing and sales automation solutions. Taipei, Taiwan
  – Developed sales agent that integrates retrieval, dialogue management, and recommendation modules for personalized user interaction in online shopping.
  – Implemented context-aware intent detection by fusing multi-turn dialogue history with user behavior signals.
  – Constructed large-scale product knowledge graph to enhance LLM reasoning with structured domain context.
  – Designed automated evaluation framework to benchmark recommendation accuracy and intent understanding.

Competitions & Awards

- **TSMC IT CareerHack: Digital Transform x Generative AI** **Domain-Knowledge Chatbot** *2024/1*
  Finalist (Top 6 teams selected)
  – Competition topic: Factory domain knowledge chatbot implementation.
  – Developed a retrieval-augmented generation (RAG) system with adaptive LLM caching and dynamic model switching to reduce inference cost and support multimodal input.

- **Ministry of Interior Hackathon: Data Innovation Application Competition** **Geospatial ML** *2023/7*
  Finalist (Top 8 teams selected)
  – Applied machine learning and statistical analysis to derive indicators for pedestrian road safety in Taipei City.
  – Developed a web-based visualization platform presenting city-wide safety maps and analytics results.

Academic Services

Reviewers: ACL ARR 2025, ACML 2024