Dongwei Jiang

jiangdongwei0@gmail.com 443-453-7610 Google Scholar Website

Currently working on LLM reasoning, agents and self-improvement. Six years of prior industrial research experience in speech processing and NLP.

EDUCATION

JOHNS HOPKINS UNIVERSITY	**M.Eng. in Computer Science**	Aug 2023 - Dec 2024
PEKING UNIVERSITY	**B.S. in Geographic Information System**	Sep 2011 - Jun 2016

WORK EXPERIENCE

AMAZON Santa Clara, Jan 2025 - Present

Applied Scientist II. Working on reinforcement learning for agentic models and auxiliary guide systems

- **Generalized RL Agents:** Investigating how to train agents that generalize across diverse task environments rather than being specialized for single domains. Training agentic models using GRPO on multiple AgentBench environments simultaneously, with reward mechanisms designed to encourage beneficial behaviors like self-reflection and exploration, achieving 15-20% performance improvement on unseen tasks
- **Guide Model:** Investigating how to enhance LLMs for specific agentic tasks without expensive retraining. Training smaller task-specific guide models using RL with task-level outcomes as rewards to provide contextual guidance for tool selection and process orchestration, allowing adaptation while keeping the core model unchanged
- **Security Agentic Testing:** Building end-to-end RL training pipeline for autonomous security agents to discover and exploit vulnerabilities. Developing assertion-based evaluation framework combining LLMaJ with deterministic security evidence validation (e.g., file creation, system access) and synthetic data generation by programmatically inserting vulnerabilities into test applications, achieving performance improvements on both internal web applications and external benchmarks like XBOW and CVEBench

SHOPEE Beijing, Nov 2021 - Nov 2022

Senior Research Scientist. Tech lead for a team of four. Responsible for low-resource ASR systems and deployment across multiple product verticals

- **Low Resource ASR Optimization:** Implemented weakly supervised training pipeline (predating Whisper) using YouTube data crawling for base model improvement and confidence-filtered pseudo-labeling on unlabeled target domain data, achieving 15-20% performance gains over baseline models
- **Low Resource ASR Annotation:** Established annotation best practices and quality control processes, achieving 25% improvement in data quality and consistency across markets
- **Product Support:** Deployed ASR systems for live stream clipping, content search, and customer service across markets. Models contributed to 23,000 additional daily orders by improving product video supply and content understanding

YUANFUDAO Beijing, Dec 2020 - Oct 2021

Senior Research Scientist. Tech lead for a team of six. Responsible for research and application of End-to-End ASR and Talking Face Generation

- **Educational Application ASR:** Implemented application-specific techniques including code switching for English courses, children speech recognition, and deep contextual biasing for proper nouns in academic subjects
- **Product Support:** Deployed ASR systems for quality assurance, achieving 82% recall while reducing manual review workload by 97%, saving 900k RMB monthly. Also supported teacher-parent communication analysis with 93% intent classification accuracy
- **Talking Face Generation:** Implemented Wav2lip models to automatically generate lip-sync videos of English teachers reading words, reducing manual video production work for children's educational content

DIDICHUXING Beijing, Jan 2017 - Dec 2020

Senior Research Scientist. Tech lead for a team of five. Responsible for the development and deployment of novel ASR and TTS technologies

- **End-to-End ASR Development:** Implemented and deployed Speech Transformer with Attention + CTC multi-task learning and word-level WFST, achieving 10-15% WER reduction. Successfully replaced all legacy hybrid models across internal products, being among the first companies in China to deploy end-to-end ASR at scale
- **Streaming ASR Systems:** Developed streaming end-to-end speech recognition solutions including neural transducer, RNN-T, and Mocha for low-latency applications

- **Text-to-Speech Systems:** Implemented end-to-end TTS pipeline with Tacotron2 and FastSpeech models, including multi-speaker synthesis and style transfer for personalized voice generation
- **Open Source Leadership:** Led development of Athena speech processing framework, achieving state-of-the-art performance on public benchmarks and providing production-ready implementations of ASR, TTS, and self-supervised learning algorithms. The project has gained significant community adoption with 900+ GitHub stars and 200+ PRs

RESEARCH EXPERIENCE

Mentors: Prof. Daniel Khashabi, Benjamin Van Durme JHU, Sep 2023 - May 2025

Research Assistant. (1) ***Self-[In]Correct:*** *Investigating why LLMs struggle with intrinsic self-improvement - discovering they are worse at discriminating among self-generated alternatives than at initial generation (2)* ***Feedback Friction:*** *Investigating whether LLMs can self-improve to perfection with perfect external feedback - finding they exhibit resistance to incorporating feedback even under near-ideal conditions (3)* ***RATIONALYST:*** *Distilling rationales from unlabelled data to provide process feedback for general reasoning (4)* ***RDTE:*** *Formulating a consistent and theoretically grounded method for annotating decompositional entailment datasets*

Mentor: Prof. Greg Durrett UT Austin (remote), Jul 2024 - Dec 2024

Research Intern. ***To CoT or Not to CoT:*** *Investigating the effectiveness of CoT prompting across 100+ papers and 20 datasets. Discovered CoT benefits mainly math/symbolic reasoning tasks*

Mentor: Prof. Shay B. Cohen University of Edinburgh (remote), Apr 2023 - Dec 2023

Research Intern. ***LeanReasoner:*** *Developed novel approach combining LLMs with Lean theorem prover to solve logical reasoning problems. Used LLMs to translate natural language into formal logic, then applied symbolic reasoning for rigorous proof generation. Discovered that pre-training on mathematical theorem proving data significantly improves logical reasoning*

Mentor: Dr. Xiangang Li DiDi, Jan 2017 - Dec 2020

Research Scientist. (1) ***Masked Predictive Coding (MPC):*** *Developed BERT-inspired self-supervised pre-training method for Transformer-based speech recognition, achieving performance improvements through masked frame reconstruction on unlabeled speech data. (2)* ***Speech SimCLR:*** *Pioneered contrastive learning approach for speech representation by combining data augmentation (pitch shift, speed perturbation, noise) with NT-Xent loss and reconstruction objectives*

SELECTED PUBLICATIONS

[1] **Dongwei Jiang**, Alvin Zhang, Andrew Wang, Nicholas Andrews, Daniel Khashabi. ***Feedback Friction: LLMs Struggle to Fully Incorporate External Feedback*** , in submission to NeurIPS, 2025

[2] **Dongwei Jiang**, Guoxuan Wang, Yining Lu, Andrew Wang, Jingyu Zhang, Chuyu Liu, Benjamin Van Durme, Daniel Khashabi. ***RATIONALYST: Pre-training Process-Supervision for Improving Reasoning*** , in ACL, 2025 [**Featured in Lilian Weng's blog**]

[3] Zayne Sprague, Fangcong Yin, Juan Diego Rodriguez, **Dongwei Jiang**, Manya Wadhwa, Prasann Singhal, Xinyu Zhao, Xi Ye, Kyle Mahowald, Greg Durrett. ***To CoT or not to CoT? Chain-of-thought helps mainly on math and symbolic reasoning*** , in ICLR, 2025

[4] **Dongwei Jiang**, Jingyu Zhang, Orion Weller, Nathaniel Weir, Benjamin Van Durme, Daniel Khashabi. ***SELF-[IN]CORRECT: LLMs Struggle with Refining Self-Generated Responses***, in AAAI, 2025

[5] Nathaniel Weir, Kate Sanders, Orion Weller, Shreya Sharma, **Dongwei Jiang**, Zhengping Jiang, Bhavana Dalvi Mishra, Oyvind Tafjord, Peter Jansen, Peter Clark, Benjamin Van Durme. ***Enhancing Systematic Decompositional Natural Language Inference Using Informal Logic***, in EMNLP, 2024

[6] **Dongwei Jiang**, Marcio Fonseca, Shay B. Cohen. ***LeanReasoner: Boosting Complex Logical Reasoning with Lean***, in NAACL, 2024

[7] **Dongwei Jiang**, Wubo Li, Miao Cao, Wei Zou, Xiangang Li. ***Speech simclr: Combining contrastive and reconstruction objective for self-supervised speech representation learning***, in InterSpeech 2021

[8] **Dongwei Jiang**, Wubo Li, Ruixiong Zhang, Miao Cao, Ne Luo, Yang Han, Wei Zou, Kun Han, Xiangang Li. ***A further study of unsupervised pretraining for transformer based speech recognition***, in ICASSP 2021

[9] **Dongwei Jiang**, Xiaoning Lei, Wubo Li, Ne Luo, Yuxuan Hu, Wei Zou, Xiangang Li. ***Improving transformer-based speech recognition using unsupervised pre-training***, arxiv, 2019