

# Jirui Dai

Phone: 1+ 4438355175 & Email: [Jdai27@jh.edu](mailto:Jdai27@jh.edu)

My interests lie at the intersection of Medical AI and Reinforcement Learning, focusing on (1) building LLM/MLLM/Agent systems for clinical practice that transform expert knowledge and real-world medical signals into reliable, usable tools, and (2) using RL for foundation models to expand capabilities beyond alignment toward broader, more robust solution spaces.

## EDUCATION

<b>Nanchang Hangkong University(NCHU)</b> Bachelor of Engineer in Software Engineering	Nanchang, China 09/2020–06/2024
<b>Nanjing University of Chinese Medicine (NJUCN)</b> Research Assistant	Nanjing, China 01/2025–07/2025
<b>Johns Hopkins University(JHU)</b> Master of Science in Engineering in Computer Science	Baltimore, America 08/2025–06/2027

## RESEARCH EXPERIENCE

### PART 1: Medical AI & Clinical Foundation Models

#### Project One: Reinforcement Learning for Robust Clinical Reasoning Transfer (Traditional Chinese Medicine & Biomedical Pathways)

- Assistant Researcher (collaboration) supervised by PhD. Zhi Liu, postdoctoral researcher at Nanjing University of Chinese Medicine & Stanford Visiting Scholar* Nov. 2025 — Present
- . Developing a GRPO-style on-policy reinforcement learning + on-policy distillation pipeline to test whether LLMs can acquire new compositional reasoning skills that transfer to out of distributions.
  - . Building chain-of-thought compositional data that unifies Wenbing theory, herbal pharmacology, and biomedical pathway knowledge into a single reasoning space.
  - . Designing a new benchmarking paradigm (“Medmulti-Agent as a Judge”) to reduce single-judge bias and improve robustness under distribution shift.
  - . Current status: dataset structuring and evaluation infrastructure in progress.

#### Project Two: DERM-3R:Multimodal Dermatology Vision-Language Model for Psoriasis (Single-Image & Multi-Image Reasoning)

- Research Collaborator with University of Washington & The Gulou Hospital of Traditional Chinese Medicine of Beijing* Aug. 2025 — Jan. 2026
- . Developing a vision-language model for psoriasis that supports both single-image recognition and longitudinal multi-image reasoning.
  - . Applying supervised fine-tuning(SFT) + LoRA (parameter-efficient adaptation) with a curriculum-style multi-stage training strategy to better align dermatology images with clinical text and structured reasoning outputs.. Single-image: generate lesion description and pathogenesis cues; multi-image: generate progression summary, syndrome diagnosis, treatment principle, formula selection, and prescription draft.
  - . Benchmarking with RAG-backed (retrieval-augmented, evidence-grounded) Multi-LLMs-as-a-Judge evaluation to reduce evaluator bias under distribution shift on clinical reasoning tasks.
  - . Final outcome: released the DERM-3R resource-efficient multimodal agent framework and public project page, packaging the model, data pipeline, and evaluation protocol for real-world dermatology workflows.

#### Project Three: Med-Shicheng:From Master Physicians to Clinical Agents via LLM Inheritance (traditonal chinese medicine expertise standardization & scaling)

- Research Assistant supervised by PhD. Zhi Liu,a postdoctoral researcher at Nanjing University of Chinese Medicine & Stanford Scholar* Dec. 2024 — Nov.2025
- . Built a structured training framework for Traditional Chinese Medicine “master-to-agent” inheritance on Qwen2.5-1.5B-Base; curated 3M+ structured medical samples from textbooks, clinical records, and expert knowledge.
  - . Proposed RAG-SFT (retrieval-augmented fine-tuning): injected retrieved expert evidence into chain-of-reasoning data using GTE embeddings (text retrieval vectors), improving dialectical reasoning accuracy by 32%.
  - . Engineered an end-to-end training pipeline: CPT (continued pretraining) → Cold-Start SFT (supervised warm-up) → GRPO (RL-style policy optimization with rejection sampling) → RAG-SFT → KTO alignment (preference-style tuning).
  - . Constructed an expert knowledge base capturing lifetime clinical insights from five renowned Traditional Chinese Medicine physicians and enabled one model to learn and switch among multiple masters’ styles.

- . Final outcome: achieved top-tier performance among lightweight models, with overall evaluation trends comparable to strong general LLMs on multi-task clinical reasoning.

**Project Four: Training and Evaluation of a Traditional Chinese Medicine and Western Medicine Vertical Domain Model**

*Researcher supervised by PhD Jiaxi Yang, a postdoctoral researcher at Columbia University*

*Jun.2024-Dec.2024*

- . Preprocessed the open-source datasets to compile two high-quality datasets of 300,000 entries, each suitable for incremental pre-training and supervised fine-tuning, and manually annotated a third dataset with 10,000 samples for preference alignment
- . Trained Llama-3.1-8B-zh using LLaMA-Factory on AutoDL; implemented LoRA-based continued pretraining and SFT, reducing training loss from  $14.10 \rightarrow 5.47$  (CPT) and  $7.27 \rightarrow 3.43$  (SFT). Attempted preference alignment with DPO (Direct Preference Optimization) to improve instruction-following and response quality.
- . Post-mortem / outcome: the project did not yield a competitive domain model due to limited data scale/quality and base-model data overlap (the selected open corpora had likely been seen during base pretraining), leading to weak measurable gains. The effort provided hands-on experience in dataset quality control, decontamination awareness, and end-to-end training/evaluation workflows for domain adaptation.

**PART 2: General ML & NLP Engineering**

**Project Five: Comparison of Federated Learning Algorithms for Predicting Results Based on the Fashion-MNIST Dataset in Non-IID Data Environments**

*Research Leader supervised by Prof. Soummya Kar from Carnegie Mellon University*

*Mar.2024-May 2024*

- . Compared four FL algorithms (FedAvg, FedSGD, SCAFFOLD, FedProx) under simulated non-IID client data distributions.
- . Built an end-to-end experimental pipeline (data partitioning, CNN baseline, evaluation metrics) and ran controlled ablations across methods.
- . Implemented and optimized FedProx and SCAFFOLD, and wrote distributed scripts to parallelize runs across multiple machines.
- . Key finding: FedProx consistently outperformed others across metrics ( $\text{FedProx} > \text{SCAFFOLD} > \text{FedAvg} > \text{FedSGD}$ ); results formed the basis of a published paper.

**Project Six: LMSYS - Chatbot Arena Human Preference Predictions (Kaggle Competition)**

*Research Leader of a Four-person Team*

*Jun.2024- Aug.2024*

- . Split 20% of a competition dataset(user interactions from the ChatBot Arena) as a training validation set
- . Trained two LLMs, i.e., gemma-2-9b and llama-3.1-8b, using the optimal configurations determined through adjusting the parameters (learning rate, frozen layers, prompt lengths, etc.)
- . Leveraged the validation set to assess the two models' performance
- . Utilized ensemble learning techniques to assign weights to the outputs of two models and then combine these outputs through weighted summation
- . Achievement: Upon evaluating the logarithmic loss between the predicted probabilities and the actual values, our team achieved a silver medal in the competition.

**Internship: ByteDance (TikTok) — NLP Algorithm Intern**

*Beijing, China Sept. 2024-Dec.2024*

- . Built and cleaned a TikTok video-comment dataset to 100K samples; implemented preprocessing to ensure a consistent and structured corpus.
- . Conducted sentiment analysis (SnowNLP baseline) and summarized related literature; presented findings in team seminars.
- . Reproduced and benchmarked two conference NER models: annotated data in BIOES format, implemented training/evaluation, and compared via F1 and loss on dev/test sets.
- . Improved NER performance through stronger pretrained backbones, attention tweaks, data augmentation, mixed-precision training, and dropout regularization.

**Project Seven: Text-CNN for Social Hot Topic Classification(Undergraduate Thesis)**

*The 1st Author, Dec.2023- Mar.2024*

- Collected and annotated news text from major sources; built a cleaned dataset for supervised topic classification.
- Implemented Text-CNN and tuned architecture (kernel sizes, layer depth) with dropout to reduce overfitting.
- Achieved strong validation/test performance and enabled near real-time topic categorization in experiments.

## PUBLICATION

1. Ziwen Chen\*; Zhendong Wang\*; Chongjing Wang\*; ...; Jirui Dai†; Changyong Luo†; Xiameng Gai†; Haibing Lan†; Zhi Liu†; et al. “DERM-3R: A Resource-Efficient Multimodal Agents Framework for Dermatologic Diagnosis and Treatment in Real-World Clinical Settings.”

Under review at npj Digital Health. Patent pending.

(\* equal contribution, co-first authors; † co-corresponding authors)

Project Website: <https://github.com/NJUCM-BJUCM-TCM-AI/DERM-3R>

2. Changyong, Luo\*; Jirui, Dai\*; Zhendong, Wang\*; et al. “From Physician Expertise to Clinical Agents: Preserving, Standardizing, and Scaling Physicians’ Medical Expertise with Lightweight LLM.”

Manuscript submitted to Nature Medicine. Patent pending. (\* equal contribution, Co-first author)

Project Website: [https://njucm-bjucm-tcm-ai.github.io/Med-Shicheng\\_project\\_website/](https://njucm-bjucm-tcm-ai.github.io/Med-Shicheng_project_website/)

3. Jirui Dai. “Comparative analysis of federated learning algorithms under non-IID data.” Applied and Computational Engineering (2024) DOI: 10.54254/2755-2721/86/20241581

[https://www.researchgate.net/publication/382753798\\_Comparative\\_analysis\\_of\\_federated\\_learning\\_algorithms\\_under\\_non-IID\\_data](https://www.researchgate.net/publication/382753798_Comparative_analysis_of_federated_learning_algorithms_under_non-IID_data)

## PROFESSIONAL SKILLS

Standardized Tests: TOEFL:103, GRE:332 (verbal:162; quantitative:170; analytical writing:4.5)

Programming Language Skills: C(2 yrs), C++(2 yrs), Java(4 yrs), Python(3 yrs), etc.

Models & Architectures: LLM, VLM, MLLM, Agents, Dense, MoE

Training Frameworks: LLaMA-Factory, OpenRLHF, EasyR1, Hugging Face TRL, VeRL, SLiME, ROLL

Inference Serving: vLLM, SGLang

## AWARDS, SCHOLARSHIPS & LEADERSHIP

- Won the Silver Medal in the LMSYS - Chatbot Arena Human Preference Predictions (Top 2%) *Aug.2024*
- Obtained the Third Prize in the Lanqiao Cup National Software and Information Technology Professional Talent Competition (Top10%) *May.2022*
- Awarded the Third-Class Scholarship at NCHU (Top7%, Three Times) *Mar.2022&2023&2024*
- Monitor of Class 15 of 2020 at School of Software, NCHU *Sept.2021-Jun.2024*

I enjoy both active collaboration and quiet, deep thinking. Working across medicine and AI, I often serve as the bridge—explaining technical decisions in plain language so clinicians can confidently use and critique the system. I place a strong emphasis on engineering execution: I build quickly, instrument experiments, and validate ideas fast—because shortening the loop from hypothesis to evidence matters more than having endless ideas.