MENG JIANG, PhD
Associate Professor

University of Notre Dame
Department of Computer Science & Engineering
384 Fitzpatrick Hall
Notre Dame, IN 46556
574-631-7454
mjiang2@nd.edu
www.meng-jiang.com

EDUCATION
Doctor of Philosophy in Computer Science (2015)
Bachelor of Engineering in Computer Science (2010)
Tsinghua University, China

PROFESSIONAL APPOINTMENTS
University of Notre Dame

Associate Professor / **Frank M. Freimann Collegiate Professor** of Engineering *(w/ tenure)*, Computer Science and Engineering (*2023-Present*)
Director of Foundation Models and Applications Lab at Lucy Family Institute (*2025-Present*)
Program Chair of ND-IBM Tech Ethics Lab (*2025-Present*)
Assistant Professor, Computer Science and Engineering (*2017-2023*)
University of Illinois at Urbana-Champaign
Postdoctoral Research Associate, Computer Science (*2015-2017*)

RESEARCH INTERESTS
- **Foundation Model Development:** Enhancing knowledge, reasoning, and instruction-following capabilities using knowledge augmentation, self-verification, reflection, and instruction tuning.
- **Foundation Model Applications:** Novel AI technologies that learn from (semi-)structured data, tons of document/literature texts and images, to automate personalized service, scientific discovery, and healthcare.

PROFESSIONAL ACCOMPLISHMENTS & SERVICE
- **167 peer-reviewed publications** (*148 in independent career*) **h-index = 52, i-10 index = 134, Total citations = 10,372**. Recent highlights from last three years:

  Knowledge-augmented Foundation Models:
  o W. Yu, [...], **M. Jiang**. "Dict-BERT: Enhancing Language Model Pre-training with Dictionary" In Meetings of Association for Computational Linguistics (ACL), **2022**.
  o W. Yu, [...], **M. Jiang**. "Diversifying Content Generation for Commonsense Reasoning with Mixture of Knowledge Graph Experts" In Meetings of Association for Computational Linguistics (ACL), **2022**.
  o M. Yu, [...], **M. Jiang**. "Pre-training Language Models for Comparative Reasoning" In Empirical Methods on Natural Language Processing (EMNLP), **2022**.
  o W. Yu, [...], **M. Jiang**. "Retrieval Augmentation for Commonsense Reasoning: A Unified Approach" In Empirical Methods on Natural Language Processing (EMNLP), **2022**.
  o W. Yu, [...], **M. Jiang**. "Generate rather than Retrieve: Large Language Models are Strong Context Generators" In International Conference on Learning Representations (ICLR), **2023**.
  o N. Ziems, [...], **M. Jiang**. "Large Language Models are Built-in Autoregressive Search Engines" In Findings of Association for Computational Linguistics (ACL), **2023**.

  Reasoning with Foundation Models:
  o Z. Wu, [...], **M. Jiang**. "Instructing Large Language Models to Identify and Ignore Irrelevant Conditions" In North American Chapter of Association for Computational Linguistics (NAACL), **2024**.
  o Z. Wu, [...], **M. Jiang**. "Get an A in Math: Progressive Rectification Prompting" In Association for the Advancement of Artificial Intelligence (AAAI), **2024**.
  o Z. Wu, [...], **M. Jiang**. "Large Language Models Can Self-Correct with Key Condition Verification" In Empirical Methods on Natural Language Processing (EMNLP), **2024**.
  o Z. Zhu, [...], **M. Jiang**. "MultiChartQA: Benchmarking Vision-Language Models on Multi-Chart Problems" In North American Chapter of Association for Computational Linguistics (NAACL), **2025**.
  o Y. Lu, [...], **M. Jiang**. "Optimizing Decomposition for Optimal Claim Verification" In Meetings of Association for Computational Linguistics (ACL), **2025**.
  o Z. Wu, [...], **M. Jiang**. "Enhancing Mathematical Reasoning in LLMs by Stepwise Correction" In Meetings of Association for Computational Linguistics (ACL), **2025**.

  Instruction-tuned Foundation Models:
  o Z. Zhang, [...], **M. Jiang**. "Auto-Instruct: Automatic Instruction Generation and Ranking for Black-Box Language Models" In Empirical Methods on Natural Language Processing (EMNLP), **2023**.
  o Z. Zhang, [...], **M. Jiang**. "Learn Beyond the Answer: Training Language Models with Reflection for Mathematical Reasoning" In Empirical Methods on Natural Language Processing (EMNLP), **2024**.
  o N. Ziems, [...], **M. Jiang**. "TOWER: Tree Organized Weighting for Evaluating Complex Instructions" In Empirical Methods on Natural Language Processing (EMNLP), **2024**.
  o Z. Zhang, [...], **M. Jiang**. "PLUG: Leveraging Pivot Language in Cross-Lingual Instruction Tuning" In Meetings of Association for Computational Linguistics (ACL), **2024**.

- Z. Zhang, [...], **M. Jiang**. "IHEval: Evaluating Language Models on Following the Instruction Hierarchy" In North American Chapter of Association for Computational Linguistics (NAACL), **2025**.

Personalization Foundation Models and Applications:
- Z. Tan, [...], **M. Jiang**. "Democratizing Large Language Models via Personalized Parameter-Efficient Fine-Tuning" In Empirical Methods on Natural Language Processing (EMNLP), **2024**.
- Z. Tan, [...], **M. Jiang**. "Personalized Pieces: Efficient Personalized Large Language Models through Collaborative Efforts" In Empirical Methods on Natural Language Processing (EMNLP), **2024**.
- B. Nguyen, [...], **M. Jiang**. "Reference-based Metrics Disprove Themselves in Question Generation" In Findings of Empirical Methods on Natural Language Processing (EMNLP), **2024**.
- G. Liu, [...], **M. Jiang**. "Learning Attribute as Explicit Relation for Sequential Recommendation" In SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), **2025**.
- Z. Tan, [...], **M. Jiang**. "Aligning Large Language Models with Implicit Preferences from User-Generated Content" In Meetings of Association for Computational Linguistics (ACL), **2025**.
- B. Nguyen, [...], **M. Jiang**. "QG-SMS: Enhancing Test Item Analysis via Student Modeling and Simulation" In Meetings of Association for Computational Linguistics (ACL), **2025**.

Foundation Models in AI for Science:
- G. Liu, [...], **M. Jiang**. "Semi-Supervised Graph Imbalanced Regression" In SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), **2022**.
- G. Liu, [...], **M. Jiang**. "Data-Centric Learning from Unlabeled Graphs with Diffusion Model" In Neural Information Processing Systems (NeurIPS), **2023**.
- G. Liu, [...], **M. Jiang**. "Graph Diffusion Transformer for Multi-Conditional Molecular Generation" In Neural Information Processing Systems (NeurIPS), **2024**.
- G. Liu, [...], **M. Jiang**. "Multimodal Large Language Models for Inverse Molecular Design with Retrosynthetic Planning" In International Conference on Learning Representations (ICLR), **2025**.
- G. Liu, [...], **M. Jiang**. "Learning Molecular Representation in a Cell" In International Conference on Learning Representations (ICLR), **2025**.

Foundation Model Safety:
- Z. Liu, [...], **M. Jiang**. "Towards Safer Large Language Models through Machine Unlearning" In Findings of Association for Computational Linguistics (ACL), pp. 1817-1829, **2024**.
- Z. Liu, [...], **M. Jiang**. "Protecting Privacy in Multimodal Large Language Models with MLLMU-Bench" In North American Chapter of Association for Computational Linguistics (NAACL), **2025**.
- Z. Liu, [...], **M. Jiang**. "Modality-Aware Neuron Pruning for Unlearning in Multimodal Large Language Models" In Meetings of Association for Computational Linguistics (ACL), **2025**.
- Z. Liu, [...], **M. Jiang**. "Disentangling Biased Knowledge from Reasoning in Large Language Models via Machine Unlearning" In Meetings of Association for Computational Linguistics (ACL), **2025**.

Foundation Models with Taxonomies:
- Q. Zeng, [...], **M. Jiang**. "Enhancing Taxonomy Completion with Concept Generation via Fusing Relational Representations" In Conference on Knowledge Discovery and Data Mining (KDD), **2021**.
- Q. Zeng, [...], **M. Jiang**. "Chain-of-Layer: Iteratively Prompting Large Language Models for Taxonomy Induction from Limited Examples" In ACM International Conference on Information and Knowledge Management (CIKM), **2024**.
- Q. Zeng, [...], **M. Jiang**. "CodeTaxo: Enhancing Taxonomy Expansion with Limited Examples via Code Language Prompts" In Findings of the Association for Computational Linguistics, **2025**.

- **3 books**: (1) <Modeling Polymers with Neural Networks> by American Chemical Society, (2) <Deep Learning for Polymer Discovery: Foundation and Advances> by Synthesis Lectures on Data Mining and Knowledge Discovery, Springer, (3) <Knowledge-augmented Methods for Natural Language Processing> by Springer.
- **4 pending/issued patents**; 1 patent licensed for development by IBM; **MJ is an advisor**.
- **39 invited external talks/seminars/keynotes** during independent career (since 2017)
- **7.9M in external grant awards** in support of MJ's research lab (since 2017)
- **Tutor/organizer of 15 tutorials** in international conferences (KDD, ACL, EMNLP, WWW, WSDM, etc.)
- **Chair of 10 workshops** in international conferences (KDD, ACL, AAAI, etc.)
- Organizer of Midwest Speech and Language Days (2025) and Open Polymer Challenge on Kaggle (2025)
- **Awards**: EMNLP Outstanding Paper Award (2023), **NSF CAREER Award** (2022), ACM SIGSOFT (ICSE) Distinguished Paper Award (2021), ISDSA Annual Meeting Best Paper Award (2020), Notre Dame International Faculty Research Award (2019), ACM SIGKDD Best Papers of KDD (2014)
- **Elected** IEEE Senior Member (2023) and ACM Senior Member (2023)
- Editor of IEEE Bulletin of the Technical Committee on Data Engineering, December 2024, Vol. 48 No. 4
- Designed new courses on topic of "*Computational Behavior Modeling*" and "*Large Language Models*".
- **Supervision of 17 PhD students, 3 Master students, 1 postdoctoral fellow, 30 undergraduates, 7 high school students, 2 visiting PhD students, and 21 visiting undergraduates**