

# Xuan Ouyang

✉ xouyang7@wisc.edu | 6083208044

---

## Education

### University of Wisconsin–Madison

Graduation Date: 2026/012 (Expected)

Bachelor of Science in Computer Science (honor)

GPA: 3.87/4.00 (Good academic standing)

Course Work: Deep Learning in Computer Vision, Data-Centric ML, Real Analysis, Probability Theory, Foundation Models, Machine Learning, Advanced Natural Language Processing

Wisconsin, USA

---

## Research Interests

I am broadly interested in Multimodal Large Language Models and Generative AI, with a focus on building systems that can perceive, reason, and generate across modalities. My recent research explores data-centric approaches to improve model robustness and efficiency, including dataset pruning, pretraining data selection.(Meltrim) and Novel method for multimodal generation tasks(MSSB). Currently, I am particularly interested in advancing Data-centric AI and Multimodal LLMs as foundations for more capable and trustworthy intelligent agents. More information can be found on my website: <https://yancyou.github.io/>.

---

## Research Experience

### University of Wisconsin–Madison NLP Lab

2025/08 - now

Research Assistant

Advised by: Prof. Junjie Hu

- Developed cross-lingual audio representations for Hmong in a low-resource setting by fine-tuning XLS-R (Fairseq/HuggingFace) and experimenting with audio tokenizers
- Designed a contrastive objective on cross-speaker/tone same-word pairs with targeted augmentation, evaluated via cross-speaker/tone retrieval (Top-1/Top-5) and downstream tasks. Create bias-aware, transferable speech representations for under-resourced languages, including a pilot healthcare deployment to assess pronunciation accuracy for hearing-impaired Hmong speakers.

Madison, US

### University of Wisconsin–Madison

2025/05 - 2025/09

Member of REU team (Research for Undergraduate students)

Advised by: Prof. Hanbaek Lyu

- Designed and implemented a multimodal Schrödinger Bridge (SB) framework for instruction-guided image editing and cross-domain image generation, formulating the task as a multi-marginal optimal transport problem.
- The approach highlights a broader vision of leveraging Schrödinger Bridges as a foundation for cross-modal generation, pushing beyond diffusion and flow-matching methods by providing a mathematically grounded yet flexible framework for multimodal alignment.

Madison, US

### Shanghai Jiao Tong University Epic Lab

2025/01 - 2025/05

Research Intern

Advised by: Prof. Linfeng Zhang

- Designed and implemented a coarse-to-fine dataset-pruning framework for speech recognition (MelTrim), building compact, high-utility coresets via acoustic clustering and frame-level utility pruning; showed consistent accuracy gains with less data and compute across emotion recognition and speaker identification benchmarks.
- Broadly advances data-centric SR by adapting principled pruning to acoustic sequences—making ASR pipelines more efficient and robust and providing a general recipe for large-scale speech datasets.

Madison, US

### University College London

2024/08 - 2024/12

Research Assistant

Advised by Dr. Yujian Gan

- Designed and evaluated Text-to-SQL systems that move beyond the single-query assumption by curating Spider-S, a benchmark with paired single- and multi-query solutions and a flexible execution-based evaluator; developed a lightweight router to decide between single vs. multiple SQL generation for greater robustness and accuracy. Try to address the mismatch between natural language descriptions and the corresponding SQL queries.

London, UK

---

Publications

- **From Single-SQL to Multi-SQL: Bridging Real-World Needs with the SpiderS Benchmark.**  
*Xuan Ouyang\**, *Yujian Gan\**, *Yuxi Lin\**, *Zhilin Zhang*, *Jinxia Xie*, *Vassilis Routsis†*. **AAAI Reviewing, 2026.**
- **MelTrim: Coarse-to-Fine Data Pruning for Speech Classification.**  
*Shaobo Wang\**, *Xuan Ouyang\**, *Xintong Li\**, *Tianle Niu*, *Zhengkun Ge*, *Yue Min*, *Xiaoqian Liu*, *Hankun Wang*, *Linfeng Zhang†*. **ACL ARR October Reviewing, 2026.**
- **Multi-Marginal f-Divergence Schrödinger Bridges: Towards a Unifying Framework for Generation and Distillation.**  
*Xuan Ouyang*, *Ishaan Kharbanda*, *Yudi Li*, *Rahul Choudhary*, *Hanbaek Lyu†* **ICLR Reviewing, 2026**

---

Working Experience

**Machine Learning Research Assistant of Marler Lab**

2025/01 - 2025/05

University of Wisconsin–Madison

- Collaborated with graduate students and faculty to optimize machine learning models for processing and synchronizing complex behavioral data, including semantic and syntactic analyses of mouse vocalizations.
- Developed and maintained robust computational pipelines for analyzing rodent vocalizations and movements using cutting-edge tools such as DeepSqueak and Sleep.

Madison, US

---

Selected Awards

I am the head of the Mathematical Modeling Association of Nanjing University of Science and Technology. I have participated in many mathematical modeling competitions in the past and have rich experience in machine learning. I am very happy to be selected as a **Dean's List** student at UW–Madison for every semester.

**Dean's List in UW–Madison**

2024/09 - 2025/05

- High-achieving students in UW–Madison.

**MCM/ICM (3-people Group)**

2023/12 - 2024/02

- Honorable Mention, Second Prize in World Class.

**May Day Mathematical Modeling Contest (3-people Group)**

2023/05 - 2023/05

- First Prize in Jiangsu Province. (0.1%)

---

Technical Skills

- **Languages:** Python, C/C++, java, Shell, LaTeX, Julia
- **Frameworks and Tools:** Pytorch, Docker, Sklearn, Numpy, Pandas
- **AI:** Natural language Processing (llama-2, ChatGLM-3, CPM-Bee) | Computer Vision (YOLO Series, OpenCV)