

# Guang Yang

Ph.D. Student (3<sup>rd</sup> Year),  
Paul G. Allen School of Computer Science & Engineering,  
University of Washington.

---

CONTACT INFORMATION	Paul G. Allen Center, Box 352350, 185 E Stevens Way NE, Seattle, WA 98195 <a href="mailto:gyang1@cs.washington.edu">gyang1@cs.washington.edu</a> , <a href="mailto:guangyangnlp@gmail.com">guangyangnlp@gmail.com</a> <a href="https://guang-yng.github.io">https://guang-yng.github.io</a> .
RESEARCH INTERESTS	Multimodal Learning, AI for Music, Vision-Language Models, Natural Language Processing
EDUCATION	<b>University of Washington</b> , Seattle, WA  Ph.D., (Expected 2028) Sep. 2023 – Present Paul G. Allen School of Computer Science & Engineering <ul style="list-style-type: none"><li>• Advisor: Prof. Noah Smith</li><li>• Research Area: Multimodal Learning, AI for Music, NLP</li></ul> <b>Tsinghua University</b> , Beijing, China  B.S., Yao Class, Sep. 2019 – Jul. 2023 Institute for Interdisciplinary Information Sciences <ul style="list-style-type: none"><li>• Major: <i>Computer Science and Technology</i></li><li>• GPA: 3.92/4.00</li><li>• Awards:<ul style="list-style-type: none"><li>• 2023 Tsinghua Outstanding Graduates</li><li>• 2020/2022 Comprehensive Excellence Scholarship</li></ul></li></ul>
RESEARCH EXPERIENCE	<b>Research Assistant</b> Apr. 2024 – Present  Advisor: Noah A. Smith Paul G. Allen School of Computer Science & Engineering, University of Washington Ongoing Projects: <ul style="list-style-type: none"><li>• <i>MMMG: a Comprehensive and Reliable Evaluation Suite for Multitask Multimodal Generation (Preprint)</i> Developed <b>MMMG</b>, a large-scale benchmark for <b>multimodal generation evaluation</b> spanning image, audio, and interleaved modalities. We designed 49 tasks and 937 instructions to assess reasoning and controllability of multimodal generative models, achieving <b>94.3% human-metric agreement</b>. Benchmarking 24 leading models revealed significant performance gaps—especially in multimodal reasoning and audio generation—highlighting key directions for future research.</li></ul> Completed Projects:

- *LEGATO: Large-scale End-to-end Generalizable Approach to Typeset OMR* (*Accepted by NeurIPS 2025 Workshop AI4Music*)  
Developed **Legato**, the first large-scale **end-to-end pretrained model for optical music recognition (OMR)** that converts full-page music scores into ABC notation. By combining a pretrained vision encoder with an ABC decoder trained on over 214K images, Legato generalizes across diverse score types and achieves up to **68% and 47.6% error reduction** on key benchmarks, surpassing previous state-of-the-art models.
- *Toward a More Complete OMR Solution* (*Accepted by ISMIR 2024*)

<b>Research Assistant</b>	Apr. 2021 – Apr. 2022 Sep. 2022 – Jul. 2023
---------------------------	--

Advisor: Zhiyuan Liu

Tsinghua University

Project Involved:

- *Parameter-efficient Fine-tuning of Large-scale Pre-trained Language Models* (*Nature Machine Intelligence*, Accepted)

This paper has been accepted by ***Nature Machine Intelligence*** and is featured as the front cover article for the March 2023 issue.

<https://www.nature.com/natmachintell/volumes/5/issues/3>

<b>Student Intern</b>	Feb. 2022 – Aug. 2022
-----------------------	-----------------------

Advisor: Heng Ji

University of Illinois at Urbana-Champaign

Project Involved:

- *Video Event Extraction via Tracking Visual States of Arguments* (AAAI 2023, Accepted)

## PUBLICATIONS & PREPRINTS

1. [NeurIPS 2025 AI4Music Workshop] “LEGATO: Large-scale End-to-end Generalizable Approach to Typeset OMR”, **Guang Yang**, Victoria Ebert, Nazif Tamer, Brian Siyuan Zheng, Luiza Pozzobon, and Noah A. Smith. [arXiv:2506.19065](https://arxiv.org/abs/2506.19065)
2. [Preprint] “MMMG: a Comprehensive and Reliable Evaluation Suite for Multitask Multimodal Generation”, Jihan Yao, Yushi Hu, Yujie Yi, Bin Han, Shangbin Feng, **Guang Yang**, Bingbing Wen, Ranjay Krishna, Lucy Lu Wang, Yulia Tsvetkov, Noah A. Smith, and Banghua Zhu. [arXiv:2505.17613](https://arxiv.org/abs/2505.17613)
3. [ISMIR 2024] “Toward a More Complete OMR Solution”, **Guang Yang**, Muru Zhang, Lin Qiu, Yanming Wan, and Noah A. Smith. [doi:10.5281/zenodo.14877483](https://doi.org/10.5281/zenodo.14877483)
4. [AAAI 2023] “Video Event Extraction via Tracking Visual States of Arguments”, **Guang Yang**, Manling Li, Jiajie Zhang, Xudong Lin, Shih-Fu Chang, and Heng Ji. [doi:10.1609/aaai.v37i3.25418](https://doi.org/10.1609/aaai.v37i3.25418)

5. [*Nature Machine Intelligence* (2023), **Cover Article**] “Parameter-efficient Fine-tuning of Large-scale Pre-trained Language Models”, Ning Ding\*, Yujia Qin\*, **Guang Yang**, Fuchao Wei, Zonghan Yang, Yusheng Su, Shengding Hu, Yulin Chen, Chi-Min Chan, Weize Chen, Jing Yi, Weilin Zhao, Xiaozhi Wang, Zhiyuan Liu, Hai-Tao Zheng, Jianfei Chen, Yang Liu, Jie Tang, Juanzi Li, and Maosong Sun. doi:[10.1038/s42256-023-00626-4](https://doi.org/10.1038/s42256-023-00626-4)