

# Yuwei Yin (Joey) 阴昱为

[Homepage](#) || [Scholar](#) || [GitHub](#) || [LinkedIn](#) || Email: seckexyin@gmail.com || Address: Central, Hong Kong

## WORK EXPERIENCE

- **The University of Hong Kong (HKU)** Pokfulam, Hong Kong  
*Research Assistant (Full-time, Onsite, Paid)* Sep. 2022 – Present
  - **Affiliation:** HKU NLP Group, Department of Computer Science, Faculty of Engineering
  - **Duty:** Conducting research on multimodal natural language processing and FinTech.
- **Microsoft (MSFT)** Haidian, Beijing, China  
*Research Intern (Full-time, Onsite, Paid)* Nov. 2020 – Sep. 2021
  - **Affiliation:** Natural Language Computing (NLC) Group, Microsoft Research Asia (MSRA)
  - **Duty:** Conducted research on multilingual neural machine translation.

## EDUCATION

- **Beihang University (BUAA)** Haidian, Beijing, China  
*Ph.D. (Dropped-out)* in Computer Science at State Key Lab of Software Dev. Env. Sep. 2019 – Sep. 2021  
*M.Sc. (Transferred)* in Computer Technology at School of Computer Sci. & Eng. Sep. 2018 – Jun. 2019  
GPA: 92/100; Rank 3%
- **Beijing Information Science and Technology University (BISTU)** Chaoyang, Beijing, China  
*B.Eng.* in Computer Science and Technology at School of Computer Science Sep. 2014 – Jun. 2018  
GPA: 93/100; Rank 1<sup>st</sup> of 202

## PUBLICATIONS

[Google Scholar](#) || [OpenReview](#) || [Semantic Scholar](#) || [ResearchGate](#) || [AMiner](#) || [ORCID](#) || [DBLP](#) || [ACL Anthology](#)

**Research Interests:** Natural Language Processing, Multilingualism & Multimodality, Reasoning & Interpretability

(\* indicates equal contribution)

## PREPRINTS

- [P2] **M3IT: A Large-scale Dataset towards Multimodal Multilingual Instruction Tuning**  
Lei Li, Yuwei Yin, Shicheng Li, Liang Chen, Peiyi Wang, Shuhuai Ren, Mukai Li, Yazheng Yang, Jingjing Xu, Xu Sun, Lingpeng Kong, Qi Liu  
Under review || [arXiv 2306.04387](#) || [Data](#) Jun 2023
- [P1] **TTIDA: Controllable Generative Data Augmentation via Text-to-Text and Text-to-Image Models**  
Yuwei Yin, Jean Kaddour, Xiang Zhang, Yixin Nie, Zhenguang Liu, Lingpeng Kong, Qi Liu  
Under review || [arXiv 2304.08821](#) Apr 2023

## CONFERENCE PROCEEDINGS

- [C7] **GanLM: Encoder-Decoder Pre-training with an Auxiliary Discriminator**  
Jian Yang, Shuming Ma, Li Dong, Shaohan Huang, Haoyang Huang, Yuwei Yin, Dongdong Zhang, Liqun Yang, Zhoujun Li and Furu Wei  
Accepted in Conference [ACL 2023](#) || [arXiv 2212.10218](#) || [Code](#) Jul 2023
- [C6] **HanoiT: Enhancing Context-aware Translation via Selective Context**  
Jian Yang, Yuwei Yin, Shuming Ma, Liqun Yang, Hongcheng Guo, Haoyang Huang, Dongdong Zhang, Yutao Zeng, Zhoujun Li, Furu Wei  
Published in Conference [DASFAA 2023](#) || [Paper](#) || [arXiv 2301.06825](#) Apr 2023
- [C5] **CROP: Zero-shot Cross-lingual Named Entity Recognition with Multilingual Labeled Sequence Translation**  
Jian Yang, Shaohan Huang, Shuming Ma, Yuwei Yin, Li Dong, Dongdong Zhang, Hongcheng Guo, Zhoujun Li, Furu Wei  
Published in Conference [EMNLP 2022 Findings](#) || [Paper](#) || [arXiv 2210.07022](#) || [Code](#) Dec 2022
- [C4] **UM4: Unified Multilingual Multiple Teacher-Student Model for Zero-Resource Neural Machine Translation**  
Jian Yang\*, Yuwei Yin\*, Shuming Ma, Dongdong Zhang, Shuangzhi Wu, Hongcheng Guo, Zhoujun Li, Furu Wei  
Published in Conference [IJCAI-ECAI 2022](#) || [Paper](#) || [arXiv 2207.04900](#) || [Code](#) Jul 2022
- [C3] **HLT-MT: High-resource Language-specific Training for Multilingual Neural Machine Translation**  
Jian Yang, Yuwei Yin, Shuming Ma, Dongdong Zhang, Zhoujun Li, Furu Wei  
Published in Conference [IJCAI-ECAI 2022](#) || [Paper](#) || [arXiv 2207.04906](#) || [Code](#) Jul 2022
- [C2] **Improving Multilingual Neural Machine Translation with Auxiliary Source Languages**  
Weijia Xu, Yuwei Yin, Shuming Ma, Dongdong Zhang, Haoyang Huang  
Published in Conference [EMNLP 2021 Findings](#) || [Paper](#) || [Code](#) Nov 2021
- [C1] **Multilingual Agreement for Multilingual Neural Machine Translation**  
Jian Yang, Yuwei Yin, Shuming Ma, Haoyang Huang, Dongdong Zhang, Zhoujun Li, Furu Wei  
Published in Conference [ACL-IJCNLP 2021](#) || [Paper](#) || [Code](#) Aug 2021

## JOURNAL ARTICLES

- [J1] *GTrans: Grouping and Fusing Transformer Layers for Neural Machine Translation*

Jian Yang, Yuwei Yin, Shuming Ma, Haoyang Huang, Dongdong Zhang, Furu Wei, Zhoujun Li

Published in Journal [IEEE TASLP](#) || [Paper](#) || [arXiv 2207.14467](#) || [Code](#)

Nov 2022

## ACADEMIC SERVICE

---

### PEER REVIEW

- Conference Proceedings
  - The 61st Annual Meeting of the Association for Computational Linguistics ([ACL 2023](#)) 2023
  - The 2023 Conference on Empirical Methods in Natural Language Processing ([EMNLP 2023](#)) 2023
  - The Fortieth International Conference on Machine Learning ([ICML 2023](#)) 2023
- Journal Articles
  - ACM Transactions on Asian and Low-Resource Language Information Processing ([ACM TALLIP](#)) 2023
  - IEEE Transactions on Audio, Speech, and Language Processing ([IEEE TASLP](#)) 2022

## HONORS

---

### SCHOLARSHIPS

- National Scholarship (Awarded by The Ministry of Education of China) Nov. 2016 & Nov. 2017
- Excellent Academic Scholarship (Awarded by BISTU) Jan. 2016 & Jan. 2017 & Jan. 2018
- Research and Innovative Scholarship (Awarded by BISTU) Jan. 2016 & Jan. 2017
- Social Contribution Scholarship (Awarded by BISTU) (Volunteering in Public Service) Jan. 2017 & Jan. 2018

### AWARDS

- Award of Excellence in Starts of Tomorrow Internship Program (Awarded by MSRA) Sep. 2021
- Beijing Outstanding Graduate Award (Awarded by Beijing Municipal Commission of Education) Jul. 2018
- Outstanding Graduate Award (Awarded by BISTU) Jul. 2018
- Diligent Study & Research Pioneer Award (Awarded by BISTU) Dec. 2017

### PRIZES

- 1<sup>st</sup> Prize in The 32nd National Physics Competition, Beijing District (Awarded by Beijing Physics Society) Dec. 2015
- 1<sup>st</sup> Prize in The 9th Mathematics Competition (Awarded by BISTU) Jun. 2015