

# ZECHENG TANG

✉ zecheng.tang@foxmail.com · 🌐 <https://zetangforward.github.io/>

## 🎓 EDUCATION & EXPERIENCE

**StepFun**, Step-Video Generation Team May 2024 – Oct 2024

*Research Intern* Mentor: Chenfei Wu and Nan Duan

**Microsoft Research Asia (MSRA)**, NLC Group, NUWA Team Jan 2023 – Feb 2024

*Research Intern* Mentor: Chenfei Wu and Nan Duan

**Soochow University**, Computer Science (CS) Sep 2022 – Present

*Ph.D. Student* Advisor: Assoc. Prof. Juntao Li and Prof. Min Zhang

**Soochow University**, Software Engineering (SE) Sep 2018 – Jun 2022

*Bachelor student* Outstanding & Honorary Graduate Student (Rank 1 / 380), Huawei Scholarship (Top 5%)

## 📖 SELECTED PUBLICATIONS

\* DENOTES EQUAL CONTRIBUTION

*To see all my publications, please visit my [Homepage](#) or [Google Scholar](#).*

- (SCIS, CCF-A, Journal) OpenBA: An Open-Sourced 15B Bilingual Asymmetric Seq2Seq Model Pre-trained from Scratch | Weijie Liu, [Zecheng Tang](#), Juntao Li, Kehai Chen, Min Zhang
- (EMNLP 2024) CMD: a framework for Context-aware Model self-Detoxification | [Zecheng Tang](#), Keyan Zhou, Juntao Li, Yuyang Ding, Pinzheng Wang, Bowen Yan, Renjie Hua, Min Zhang
- (ICML 2024) StrokeNUWA: Tokenizing Strokes for Vector Graphic Synthesis | [Zecheng Tang](#), Chenfei Wu, Zekai Zhang, Mingheng Ni, Shengming Yin, Yu Liu, Zhengyuan Yang, Lijuan Wang, Zicheng Liu, Juntao Li, Nan Duan
- (ICLR 2024) LayoutNUWA: Revealing the Hidden Layout Expertise of Large Language Models | [Zecheng Tang](#), Chenfei Wu, Juntao Li, Nan Duan
- (WSDM 2024) Chinese Spelling Check, Datasets, Chinese Mobile Search Engine | Yue Wang, Zilong Zheng, [Zecheng Tang](#), Juntao Li, Zhihui Liu, Kunlong Chen, Jinxiong Chang, Qishen Zhang, Zhongyi Liu, Min Zhang
- (ACL 2023) Open-ended Long Text Generation via Masked Language Modeling | Xiaobo Liang\*, [Zecheng Tang\\*](#), Juntao Li, Min Zhang
- (EMNLP2022) Improving Temporal Generalization of Pre-trained Language Models with Lexical Semantic Change | Zhaochen Su\*, [Zecheng Tang\\*](#), Xinyan Guan, Lijun Wu, Min Zhang, Juntao Li

## 📖 SELECTED PREPRINT PUBLICATIONS

\* DENOTES EQUAL CONTRIBUTION

*Not updated due to limited space, please visit my [Homepage](#) for new publications/projects.*

- (Aug 2024 Preprint) MemLong: Memory-Augmented Retrieval for Long Text Modeling | Weijie Liu, [Zecheng-Tang\\*](#), Juntao Li, Kehai Chen, Min Zhang
- (Mar 2023 Preprint) Visual ChatGPT: Talking, Drawing and Editing with Visual Foundation Models | Chenfei Wu, Shengming Yin, Weizhen Qi, Xiaodong Wang, [Zecheng Tang](#), Nan Duan

## ⚙️ SELECTED PROJECTS

- **TaskMatrix (Visual ChatGPT)**: Connecting ChatGPT and a series of Visual Foundation Models to enable sending and receiving images during chatting. | [Github Stars: 35K+](#)
- **OpenBA-series**: The open-sourcing of 15B bilingual asymmetric seq2seq model for its entire training process, data sources, collection, and construction. | [Total Github Stars: 94 + 23](#)