

# TAO MENG

(310) · 890 · 6057 ◊ mtsomethree@outlook.com

<https://mtsomethree.github.io/>

## EDUCATION

---

**University of California, Los Angeles**

*September 2018 - June 2024*

Ph.D in Computer Science

*UCLANLP Group; w/Kai-Wei Chang*

**Tsinghua University**

*September 2014 - July 2018*

B.S. in Institute for Interdisciplinary Information Science (a.k.a Andrew Chi-Chih Yao's class)

*CS-Major GPA: 88.5/100*

## RESEARCH INTERESTS

---

1. *Constraints integration and learning for structural data*
2. *Controllable generation via constrained decoding and inference*

## RESEARCH EXPERIENCES

---

**Microsoft AI**

*January 2026 - Present*

*Member of Technical Staff @ SuperIntelligence; w/Dr. Weizhu Chen*

*Remote*

- Integrate OpenAI data synthesize and training recipe to Microsoft self-sufficient model.
- Design training and evaluation pipeline for midtraining.

**Zoom Video Communications, Inc**

*July 2024 - December 2025*

*Senior AI Scientist @ AI Incubation; w/Dr. Pengcheng He*

*Remote*

- Take the lead in designing and managing data pipelines for pretraining small language models (SLM) in Zooms workspace, overseeing tasks such as data selection, filtering, rewriting, and ensuring data quality.
- Designed and evaluated routing strategies across specialized models (Zoom SLMs) and APIs, integrating uncertainty estimation and costlatency trade-offs into production inference flows.

**Amazon.com**

*September 2023 - December 2023*

*Applied Scientist Intern @ Trustworthy NLP; w/Dr. Rahul Gupta*

*Los Angeles*

- Finetune large language models based on controllable generation.

**Meta.Inc**

*June 2022 - September 2022*

*Research Scientist Intern @ AI Integrity; w/Dr. Sinong Wang*

*Seattle*

- Detoxifying large pretraining language models with controllable generation.

**Amazon.com**

*June 2020 - September 2020*

*Applied Scientist Intern @ Alexa NLP; w/Dr. Shervin Malmasi, Dr. Anjie Fang* *Seattle (Remote)*

- Improving Named Entity Recognition (NER) with gazetteers.

## HONORS, AWARDS, AND SCHOLARSHIPS

---

Amazon Fellowship	2022
Xuetang Class Scholarship, Tsinghua University	2014-2018
Gold medal, 28th Chinese Mathematics Olympics	2012
Silver medal, 29th, 30th National Olympics of Informatics	2012, 2013

## SELECTED PUBLICATION

---

1. Wenxuan Zhou, Shujian Zhang, Lingxiao Zhao, **Tao Meng**. T-reg: Preference Optimization with Token-level Reward Regularization, *Association for Computational Linguistics (ACL 2025)*
2. Honghua Zhang, Liunian Harold Li, **Tao Meng**, Kai-Wei Chang and Guy Van den Broeck. On the Paradox of Learning to Reason from Data, *International Joint Conferences on Artificial Intelligence (IJCAI 2023)*.
3. **Tao Meng\***, Sidi Lu\*, Nanyun Peng and Kai-Wei Chang. Controllable Text Generation with Neurally-Decomposed Oracle, *Conference on Neural Information Processing Systems (NeurIPS 2022)*.
4. Sidi Lu, **Tao Meng**, Nanyun Peng. nsNet: An Efficient Framework for Bridging Sequential and Parallel Decoding in Insertion-based Text Generation, *Conference on Neural Information Processing Systems (NeurIPS 2022)*.
5. **Tao Meng**, Anjie Fang, Oleg Rokhlenko and Shervin Malmasi. GEMNET: Effective Gated Gazetteer Representations for Recognizing Complex Entities in Low-context Input, *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021)*.
6. **Tao Meng**, Kai-Wei Chang. An Integer Linear Programming Framework for Mining Constraints from Data, *International Conference on Machine Learning (ICML 2021)*
7. Da Yin, **Tao Meng**, Kai-Wei Chang. SentiBERT: An Effective, Transferable and Interpretable Architecture for Compositional Sentiment Semantics, *Association for Computational Linguistics (ACL 2020)*.
8. Fan Yin, Quanyu Long, **Tao Meng**, Kai-Wei Chang. On the Robustness of Language Encoders against Grammatical Errors, *Association for Computational Linguistics (ACL 2020)*.
9. Shengyu Jia\*, **Tao Meng\***, Jieyu Zhao, Kai-Wei Chang. Mitigating Gender Bias Amplification in Distribution by Posterior Regularization, *Association for Computational Linguistics (ACL 2020)*.
10. **Tao Meng**, Nanyun Peng, Kai-Wei Chang. Target Language-Aware Constrained Inference for Cross-lingual Dependency Parsing, *Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)*.

## PROFESSIONAL EXPERIENCE

---

1. Program Committee / Reviewer, SoCalNLP2019-2021, 2023, EMNLP 2020-2022, 2024, ACL 2020-2022, NAACL 2021-2022, ARR 2021-2024, ICML 2022-2023, NeurIPS 2022-2023

## TECHNICAL STRENGTHS

---

<b>Programming Languages</b>	C, C++, C#, Python, Java, Matlab
<b>Natural Languages</b>	Mandarin, English
<b>Strengths</b>	Mathematics, Algorithm