

# Zifeng Wang

🏠 Home Page | 🎓 Google Scholar | 🌐 GitHub | ✉ zifengw2@illinois.edu

## EDUCATION BACKGROUND

---

### University of Illinois, Urbana-Champaign

PhD student, Computer Science, The Grainger College of Engineering  
Research Interest: AI for Healthcare & Clinical Trial; Advised by: Prof. Jimeng Sun

Illinois, US

Sept. 2021-Present

### Tsinghua University

MS, Data Science, Tsinghua-Berkeley Shenzhen Institute (TBSI)  
Thesis: Information Bottleneck for Representation Learning: New Vision  
Co-advised by: Prof. Shao-Lun Huang, TBSI and Prof. Khalid M. Mosalam, UC-Berkeley

Shenzhen, China

Sept. 2018-Jun. 2021

### Tongji University

B.Eng., Structural Engineering, School of Civil Engineering  
Advised by: Prof. Suzhen Li

Shanghai, China

Sept. 2014-Jun. 2018

## RESEARCH TOPICS

---

I am working on AI for clinical trials & healthcare, including several topics like:

- **Prediction.** Enable flexible individual/trial outcome predictions.
- **Generation.** Generate synthetic EHR/trial data to boost general health AI applications.
- **Optimization.** Clinical trial protocol referential retrieval and optimization.
- **Explanation.** Provide explainable and actionable counterfactuals to improve clinical trials.

## PAPERS

---

### ◇ Preprints & Working Papers:

- **Z Wang**, C Gao, L Glass and J Sun. Artificial Intelligence for In Silico Clinical Trials: A Review. Under submission to **Nature Communications**.
- **Z Wang**, Z Wu, D Agarwal and J Sun. MedCLIP: Contrastive Learning from Unpaired Medical Images and Text. Under submission to **EMNLP'22**.
- **Z Wang** and J Sun. PromptEHR: Conditional Electronic Healthcare Records Generation with Prompt Learning. Under submission to **EMNLP'22**.
- **Z Wang** and J Sun. Trial2Vec: Zero-Shot Clinical Trial Document Similarity Search using Self-Supervision. Under submission to **EMNLP'22**.
- D Agarwal, **Z Wang** and J Sun. MedPrompt: Multimodal Prompt Tuning for Medical Vision-Language Models. Under submission to **ICLR'23**.

### ◇ Conferences:

- **Z Wang** and J Sun. TransTab: Learning Transferable Tabular Transformers Across Tables. **NeurIPS'22**.
- **Z Wang**, R Wen, X Chen, S-L Huang, N Zhang, and Y Zheng. Finding Influential Instances for Distantly Supervised Relation Extraction. **COLING'22 (Oral)**.
- **Z Wang** and J Sun. SurvTRACE: Transformers for Survival Analysis with Competing Events. **ACM-BCB'22**.
- **Z Wang**, S-L Huang, E. E. Kuruoglu, J Sun, X Chen, and Y Zheng. PAC-Bayes Information Bottleneck. **ICLR'22 (Spotlight, 176/3391)**.
- **Z Wang**, Y Yang, R Wen, X Chen, S-L Huang, and Y Zheng. Lifelong Learning Disease Diagnosis on Clinical Notes. **PAKDD'21 (Best Student Paper, 1/768)**. [video]
- **Z Wang**, R Wen, X Chen, S Cao, S-L Huang, B Qian, and Y Zheng. Online Disease Self-diagnosis with Inductive Heterogeneous Graph Convolutional Networks. **WWW'21**. [video]
- **Z Wang**, X Chen, R Wen, S-L Huang, E. E. Kuruoglu, and Y Zheng. Information Theoretic Counterfactual Learning from Missing-Not-At-Random Feedback. **NeurIPS'20**. [poster]
- **Z Wang**, H Zhu, Z Dong, X He, and S-L Huang. Less Is Better: Unweighted Data Subsampling via Influence Function. **AAAI'20**. [poster]

### ◇ Journals:

- **Z Wang**, Y Zhang, K. M. Mosalam, Y Gao, and S-L Huang. Deep Semantic Segmentation for Visual Understanding on Construction Sites. **Computer-Aided Civil And Infrastructure Engineering**, 2021.
- **Z Wang** and S Li. Data-driven Risk Assessment on Urban Pipeline Network Based on a Cluster Model. **Reliability Engineering and System Safety**, 2020, 196: 106781.

## SOFTWARE

---

- **TransTab**: Transferable Transformers for Tabular Learning and Prediction. [Doc] [Github]
- **PyTrial**: A Comprehensive Python Package on AI for In Silico Clinical Trial Optimization.

## PROFESSIONAL EXPERIENCE

---

**Amplitude**, Research Intern May 2022-Aug. 2022  
Topics: Multi-treatment Causal Inference  
Supervisor: Cao Xiao

**Tencent Jarvis Lab**, Research Intern Dec. 2019-Jun. 2021  
Topics: Information-theoretic DL; Lifelong Learning Diagnostic AI; Graph DL for Clinical Diagnosis  
Supervisor: Yefeng Zheng

**Noah's Ark Lab**, Research Intern Apr. 2019-Oct. 2019  
Topics: Learning from Noisy Data; Unbiased Recommendation;  
Supervisors: Zhenhua Dong, Xiuqiang He

## PROFESSIONAL SERVICE

---

- PC Member/Reviewer for NeurIPS'22, EMNLP'22, AACL'22, IJCAI'22, ICIP'21, ICASSP'21.

## TEACHING

---

- TA, CS 598 Deep Learning for Healthcare, Prof. Jimeng Sun *Spring, 2022*
- TA, Optimization Models and Applications, Prof. Laurent El Ghaoui *Summer, 2020*
- TA, Bayesian Learning and Data Analysis, Prof. Ercan E. Kuruoglu *Spring, 2020*
- TA, Learning from Data, Prof. Shao-Lun Huang and Prof. Yang Li *Fall, 2019*

## AWARDS

---

- Yunni & Maxine Pao Memorial Fellowship (3 at UIUC each year, \$5000) *Feb 2022*
- Outstanding graduate student of Tsinghua University (2/168) *June 2021*
- Best Student Research Runner-up of 13rd PhD Student Symposium of Bay Area *June 2021*
- Best Student Paper Award of PAKDD'21 (1/768) *May 2021*
- National Graduate Student Scholarship at Tsinghua University (3/229) *Oct. 2020*
- Best Student Research Runner-up of 1st TBSI Workshop On Data Science *Dec. 2019*
- Outstanding graduate student (4/40), graduate thesis (3/168) of Tongji University *Jun. 2018*
- Merit student scholarship of Tongji University *2015/2016/2017*
- Meritorious winner (1st class prize,  $\approx 7\%$ ) in USA Mathematical Contest in Modeling *Apr. 2017*

## SKILLS & CERTIFICATION

---

- English: TOEFL (105), IELTS (7.0), CET-6 (615),
- IT: Linux, Python, C++ and Python packages including Pytorch, Tensorflow, Numpy, Scipy, Pandas, Sklearn, etc.
- Hobbies: Bamboo flute, Hulusi.

(Updated on September 20, 2022.)