

Zhoujun Cheng

blankcheng@sjtu.edu.cn • Homepage • Github

EDUCATION

Shanghai Jiao Tong University, Shanghai, China

- M.S. in Computer Science Sep 2021 – *Present*
 - Supervisor: Prof. Fan Cheng
 - Research Interest: Natural Language Processing, especially on structured data modeling
- B.S. in Computer Science (IEEE honor class) Sep 2017 – Jun 2021
 - Overall GPA: 3.96 / 4.30, Rank: 8 / 109

EXPERIENCE

Microsoft Research Asia DKI Group, Beijing, China

- Research Intern, working on tabular data Dec 2020 – *Present*
 - Built a hierarchical table dataset for Table QA and NLG
 - Pretrained tabular models with spreadsheet formulas
 - Supervisor: Haoyu Dong

Alibaba TmallSearch, Hangzhou, China

- Engineering Intern, TmallShop Search Jul 2020 – Sep 2020
 - Implemented a recommendation module for Double Eleven Carnival
 - Supervisor: Jin He

MVIG Group, Shanghai Jiao Tong University, Shanghai, China

- Research Intern, working on 3D object matching Jan 2019 – Feb 2020
 - Collected a 3D keypoint dataset to help fine-grained object correspondence
 - Supervisor: Prof. Cewu Lu

PUBLICATIONS

- [1] Haoyu Dong, **Zhoujun Cheng**, Xinyi He, Mengyu Zhou, Anda Zhou, Fan Zhou, Ao Liu, Shi Han, Dongmei Zhang, “Table Pre-training: A Survey on Model Architectures, Pretraining Objectives, and Downstream Tasks”, **Preprint, January 2022**
- [2] **Zhoujun Cheng***, Haoyu Dong*, Zhiruo Wang*, Ran Jia, Jiaqi Guo, Yan Gao, Shi Han, Jian-Guang Lou, Dongmei Zhang, “HiTab: A Hierarchical Table Dataset for Question Answering and Natural Language Generation”, **ACL 2022**
- [3] **Zhoujun Cheng***, Haoyu Dong*, Ran Jia, Pengfei Wu, Shi Han, Fan Cheng, Dongmei Zhang, “FORTAP: Using Formulae for Numerical-Reasoning-Aware Table Pretraining”, **ACL 2022**
- [4] Yang You, Yujing Lou*, Chengkun Li*, **Zhoujun Cheng**, Liangwei Li, Lizhuang Ma, Cewu Lu, Weiming Wang, “KeypointNet: A Large-scale 3D Keypoint Dataset Aggregated from Numerous Human Annotations”, **CVPR 2020**
- [5] Yujing Lou*, Yang You*, Chengkun Li*, **Zhoujun Cheng**, Liangwei Li, Lizhuang Ma, Weiming Wang, Cewu Lu, “Human Correspondence Consensus for 3D Object Semantic Understanding”, **ECCV 2020**

AWARDS & SCHOLARSHIPS

- National Scholarship (2%) 2018
- MSRA Stars of Tomorrow 2021
- Shanghai Outstanding Graduates 2021
- SJTU Zhiyuan Honor Scholarship (5%) 2018 – 2020

SKILLS

- Programming Languages: Python > C++ > Others
- Deep Learning Tools: Pytorch > Huggingface > AllenNLP > Fairseq \approx Tensorflow
- Languages: Mandarin > English \approx Shanghai dialect