

## Data Mining Group, University of Illinois at Urbana-Champaign

December 24, 2017

RM 2132, SIEBEL CENTER FOR COMPUTER SCIENCE  
201 N. GOODWIN AVENUE  
URBANA, IL 61801, USA

### Application for Visiting Researcher

Dear Dr. Jiawei Han,

I am Frank Fangzheng Xu, a second-year master student at Prof. Kenny Zhu's ADAPT lab at Shanghai Jiao Tong University. I am writing to apply for visiting research internship in your group starting from next March for 6-12 months (tentative). I have been visiting Data Mining Group in UIUC from Jun. 2017 to Dec. 2017, and it has been a great time for me doing research. Prior to that, I received my bachelor's degree majoring computer science at IEEE Honored Class in SJTU on 2016 and achieved Best Thesis Award as top 1% in the university. I am a big fan of information extraction, text mining and natural language processing.

Interestingly, I found that DMG group's "turning massive, unstructured text corpora into machine-actionable structures" coincides with our group's "transform unstructured text data into structured knowledge and use such knowledge to better understand text". It is one of those directions that interests me the most. So I confidently believe that I could make a really good addition to your group, following such research motto.

The motivation to return to DMG next spring all started with my last visit to your group. During the past 6 months at UIUC and Data Mining Group, I really had a great time while collaborating extensively with group members. I find information extraction and enriching knowledge from massive corpora is really interesting and exciting. I have collaborated with Ellen on the paper accepted by WSDM 2018 and currently I am actively extending the framework for KDD/ACL 2018. Liyuan, Jingbo and I have improved sequence labeling tasks and the work got accepted by AAAI 2018, and a improved version submitted to NAACL 2018 as well. We are looking deeper into the problem for future work, which may greatly help biomedical text related problems. Besides, Chao Zhang, Qi Zhu and I are also working on open event/relation extraction from different corpus from different domains, with an aim for KDD 2018. I am also leading a project together with Yu Zhang, Chao Zhang on multi-sense embedding for text and graph. Ahmed and I are actively preparing the submission to VLDB 2018 about subword mining and embedding. With so many interesting collaborations already being conducted between us, and promising future directions, I think it would be best for me to work face to face again with all the smart and dedicated folks in DMG.

In summary, I really believe that I have the great passion for doing research and is self-motivated in so many attractive problems to solve! I have already been excited, just on the mere thought of carrying out all these proposed in your group with all these talented and nice people during the upcoming year! Thanks for your time and looking forward to your consideration.

Sincerely,

**Frank F. Xu**

*Attached: Curriculum Vitae*