
# Wang Yizhong

eastonwyz@gmail.com
yizhong.space

**September 15, 1993**
Mobile: 86-18818212946
github.com/eastonwang


## Education

- **Shanghai Jiao Tong University** Shanghai
  *B.S. in Computer Science* *2012.9 - 2016.7*
  - IEEE Pilot Class, School of Electronic Information and Electrical Engineering
  - GPA: 88/100
  - CET4: 626 CET6: 575
- **Peking University** Beijing
  *M.S. in Computer Science* *Starts from 2016.9*
  - Institute of Computational Linguistics, under the supervision of Prof. Sujian Li
  - Be recommended to admission without exams

## Work Experience

- **CooTeck Inc. - TouchPal Keyboard (Input Method Engine Group)** Shanghai
  *Software development intern* *2015.11 - 2016.3*
  - Responsible for building language models for TouchPal's multilingual keyboard.
  - Development of tools for corpus crawling, model evaluation and user profiling.

## Awards

- **Meritorious Winner (First Prize) of 2015 Mathematical Contest in Modeling**
- **2014 Chun-Tsung Scholar (Named after Tsung-Dao Lee)**
- **2013/2014/2015 Merit Scholarship of Shanghai Jiao Tong Univ.**
- **2013/2015 Xindong Scholarship**
- **2013 Three-good Student of Shanghai Jiaotong University**

## Skills

- **Development:** Python (preferred) / Java / JavaScript / PHP / C / Shell
- **Research:** Natural Language Processing / Machine Learning / Text Mining
- **Toolkits:** NLTK / Numpy / Sci-learn / Gensim / Bootstrap / Laravel

## Featured Projects

- **Mining Cultural Differences between Terms from Multilingual Text** Ongoing
  *http://adapt.seiee.sjtu.edu.cn/cdminer* *Python, Java*
  - This project aims to detect cultural difference with respect to terms in English and Chinese. We focus on concepts that are shared among languages but have very different associations and uses, such as "dragon" or "ghost".
  - We try to find the best methods to represent words, including co-occurrence matrix, SVD and Word2Vec. And we explored the relationship between vector spaces of different languages.
  - I'm the leader of this project. Two papers have been submitted to ACL 2016.
- **Sequence-based Natural Language Dependency Parser** 2014.5 - 2015.3
  *http://adapt.seiee.sjtu.edu.cn/dependency-parser* *Java*
  - We proposed a novel sequence-based framework to improve the performance of data-driven dependency parser for multilingual and achieved 90.2% accuracy with $O(n^2)$ time complexity.
  - Machine learning techniques, such as linear model, support vector machine and learning to rank, are employed to achieve our goal.
  - I implemented the main algorithm and build the demo website for this application.
- **Academic Search Engine** 2014.5 - 2015.3
  *http://acemap.sjtu.edu.cn* *C++, Python, Solr, etc.*
  - This project aims to set up an academic search engine which helps users find the most suitable papers under a topic and visualize the topic evolution tree starting from a particular paper.
  - I implemented automatic paper clustering with "terminology extraction" and "LDA model".
  - I visualized the paper topics and keyword cloud of authors with D3.js

## Featured Curricula

- Software Engineering (EI333)
- Computer System Architecture (EI332)
- Operating Systems (EI338)
- Computation Theory (EI340)
- Database System Technology (SE305)
- Machine Leaning (CS385)
- Natural Language Processing (CS382)
- Multicore Architecture and Parallel Programming (CS427)

## Other Activities:

- Work as a class monitor for 2 years till now
- Work as the minister of the amateur astronomers association in Shanghai Jiao Tong Univ.
- Strong interests in basketball, badminton and photographing