

# Bowei Zhang

☎ (+86) 18101389470

✉ bw\_zhang@pku.edu.cn

in [linkedin.com/in/imbowei](https://www.linkedin.com/in/imbowei)

📍 No. 128, Zhongguancun North Street, Haidian District, Beijing

📄 Zip Code : 100080

📄 Online Resume

## Software Engineer ( NLP | ML )

My research on semantic parsing and Data mining that involve a series of issues : probabilistic graphical models, transfer learning, multitask learning and neural networks. Currently, I am interesting in various approaches to deep learning (GAN, Reinforcement Learning, etc.) and their applications to natural language processing.

## 🎓 Education

|                   |                 |  |                                       |
|-------------------|-----------------|--|---------------------------------------|
| 2017.09 – Present | <b>Master</b>   | Data Science, Natural Language Processing<br><i>Language Computing &amp; Web Mining Group, Institute of Computer Science and Technology, Peking University</i><br><i>The MOE Key Laboratory of Computational Linguistics, Peking University</i><br><i>Center for Data Science, Peking University</i> | <b>Peking University (PKU)</b>        |
| 2013.09 – 2017.06 | <b>Bachelor</b> | Intelligence Science And Technology<br>GPA : 92.11/100   Ranking : 3/59  | <b>Central South University (CSU)</b> |

## 🔗 Projects & Experiences

|                    |   |                      |
|--------------------|---|----------------------|
| Present<br>2018.04 | <b>Semi-automatically Extending Mandarin VerbNet Annotations</b>  | <b>Python</b>        |
|                    | <ul style="list-style-type: none"><li>➢ <i>Semi-automatic Labeling</i> : Applying automatic semantic role labeling to improve annotation efficiency and accuracy for Mandarin VerbNet.</li><li>➢ <i>Semantic Role Labeling</i> : Automatic annotation method for semantic roles based on sequence labeling.</li><li>➢ <i>Self-Attention</i> : Self-attention alleviates the loss of semantic expression in long sentences.</li></ul>  |                      |
|                    | <span>Semantic Role Labeling</span> <span>Semi-automatic Labeling</span> <span>Self-Attention</span>  |                      |
| 2018.07<br>2018.06 | <b>Shared Task : Text Traceability Technology Evaluation (SMP2018-ETST)</b>   | <b>Ranking First</b> |
|                    | <ul style="list-style-type: none"><li>➢ <i>Paraphrase Detection</i> : Designed a new paraphrase detection algorithm to determine whether a sentence is adapted from another sentence.</li><li>➢ <i>Inverted Index</i> : Pre-screening candidate sentence pairs by establishing an inverted index by weighted tf-idf.</li><li>➢ <i>Ensemble Learning</i> : Improving the accuracy of paraphrase detection with ensemble learning.</li></ul>  | <b>Python</b>        |
|                    | <span>Paraphrase Detection</span> <span>Language Model</span> <span>Unsupervised Learning</span> <span>Inverted Index</span>  |                      |
| 2018.03<br>2017.12 | <b>Leveraging Paraphrases to Enhance Semantic Dependency Parsing</b>  | <b>Python</b>        |
|                    | <ul style="list-style-type: none"><li>➢ <i>Paraphrase Extraction</i> : Designed a new paraphrase extraction algorithm and building PKU Paraphrase Bank, a Chinese paraphrase corpus consisting of over 290,000 sentence pairs.</li><li>➢ <i>Semantic Consistency</i> : Exploring the implicit semantic consistency of paraphrase sentence pairs.</li><li>➢ <i>Semi-supervised Learning</i> : Extending the training data set using a semi-supervised learning method; improved semantic parsing accuracy is obtained.</li></ul>         |                      |
|                    | <span>Paraphrase Extraction</span> <span>Semantic Consistency</span> <span>Semi-supervised Learning</span>  |                      |
| 2017.05<br>2017.03 | <b>Anomaly Detection Tool Based On Extended Latent Dirichlet Allocation (LDA)</b>   | <b>Java/Js/SQL</b>   |
|                    | <ul style="list-style-type: none"><li>➢ <i>Topic Modeling</i> : Constructing an extended multi-view LDA model to establish associations between different categories of features in the same instance.</li><li>➢ <i>Anomaly Detection</i> : Five methods of similarity measurement are analyzed and the best method is selected.</li><li>➢ <i>Result Presentation</i> : Web-based data visualization. It is possible to provide an overview of the data as a whole and to pay attention to the details of the data as needed.</li></ul> |                      |
|                    | <span>Extended LDA</span> <span>Anomaly Detection</span> <span>Back-end Development</span> <span>Web Design</span> <span>Data Visualization</span>  |                      |
| 2015.08<br>2015.03 | <b>Interactive Drum Robot Based On Wearable Computing</b>   | <b>C++/PCB</b>       |
|                    | <ul style="list-style-type: none"><li>➢ <i>Hardware Building</i> : Design the robot body skeleton structure according to the principle of mechanics;</li><li>➢ <i>Circuit Design</i> : Drum stick circuit design, Printed Circuit Board(PCB) drawing;</li><li>➢ <i>Algorithm Design</i> : Design and implementation of knocking algorithm and intelligent soundtrack algorithm;</li></ul>   |                      |
|                    | <span>Wearable Computing</span> <span>Intelligent Instrument</span> <span>Internet of Things</span> <span>Cloud Computing</span>  |                      |

## Campus Life

---

- › Chuang Yue Space Of Central South University ( first maker space, Star Club of CSU) **Vice President** 2015—2016  
On behalf of Central South University to participate in the Hunan Science and Technology Association's first "Chuangke-Hunan" technology products and project exhibition;
- › Intelligent Car Association of Central South University **Vice President** 2014—2015  
Organizing 2016 Freescale Smart Cars National Finals; conducting a series of pre-competition training courses and promotional activities;
- › College Football Team (more than 3,600 people in the whole college) **Player** 2013—2015  
2013 school-level freshman semi-finals; 2015 school-level football league champion;