

# Jingqing Zhang

---

+44 0 7422 942641 | [jzhang@pangaeadata.ai](mailto:jzhang@pangaeadata.ai) | [jingqingz.github.io](http://jingqingz.github.io)

## Education

### PHD | 2017.10 - 2022.02 | IMPERIAL COLLEGE LONDON | UK

- Department of Computing, Data Science Institute, EPSRC CDT in HiPEDS
- Research interests: Natural Language Processing, Deep Learning, Machine Learning, Data Mining
- Supervisor: Prof. Yike Guo

### M.RES | 2016.10 - 2017.09 | IMPERIAL COLLEGE LONDON | UK

- Department of Computing, Degree with **Distinction**
- Courses include: Mathematical for Inference and Machine Learning, Computational Optimisation, Dynamical System and Deep Learning

### B.ENG | 2012.08 - 2016.07 | TSINGHUA UNIVERSITY | CHINA

- Department of Computer Science and Technology, **Overall GPA: 91/100, Ranking: 9/116**
- Courses include: Computer Architecture, Data Mining, Artificial Neural Networks, Data Structures, Discrete Mathematics, Numerical Analysis, Programming, Probability and Statistics, Linear Algebra, Calculus, Physics for Scientists and Engineers

## Work Experience

### HEAD OF AI | 2019.12 - PRESENT | PANGAEA DATA | LONDON, UK

- Lead the research, development and productization of cutting-edge natural language processing and machine learning algorithms in life sciences.

### RESEARCH INTERN | 2019.09 - 2019.12 | GOOGLE RESEARCH, BRAIN TEAM | CA, USA

- Propose PEGASUS, a pre-training model tailored for abstractive text summarization, which achieves state-of-the-art performance on all 12 downstream datasets we consider.

## Books

### DEEP REINFORCEMENT LEARNING: FUNDAMENTALS, RESEARCH AND APPLICATIONS

- Hao Dong, Zihan Ding, Shanghang Zhang, **Jingqing Zhang**, et al. Springer Nature, 2020.  
<https://deepreinforcementlearningbook.org/>

### DEEP LEARNING USING TENSORLAYER

- Hao Dong, Yike Guo, Guang Yang, **Jingqing Zhang**, et al. Book in Chinese. Publishing House of Electronics Industry, 2018, ISBN: 9787121326226.

## Publications

### **UNSUPERVISED NUMERICAL REASONING TO EXTRACT PHENOTYPES FROM CLINICAL TEXT BY LEVERAGING EXTERNAL KNOWLEDGE**

- Ashwani Tanwar\*, **Jingqing Zhang\***, Julia Ive, Vibhor Gupta, Yike Guo. arXiv preprint arXiv:2204.10202.

### **SELF-SUPERVISED DETECTION OF CONTEXTUAL SYNONYMS IN A MULTI-CLASS SETTING: PHENOTYPE ANNOTATION USE CASE**

- **Jingqing Zhang**, Luis Bolanos, Tong Li, Ashwani Tanwar, Guilherme Freire, Xian Yang, Julia Ive, Vibhor Gupta, Yike Guo. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021).

### **CLINICAL UTILITY OF THE AUTOMATIC PHENOTYPE ANNOTATION IN UNSTRUCTURED CLINICAL NOTES: ICU USE CASES**

- **Jingqing Zhang**, Luis Bolanos, Ashwani Tanwar, Albert Sokol, Julia Ive, Vibhor Gupta, Yike Guo. arXiv preprint arXiv:2107.11665.

### **PEGASUS: PRE-TRAINING WITH EXTRACTED GAP-SENTENCES FOR ABSTRACTIVE SUMMARIZATION**

- **Jingqing Zhang\***, Yao Zhao\*, Mohammad Saleh, Peter J. Liu. Thirty-seventh International Conference on Machine Learning (ICML 2020).

### **UNSUPERVISED ANNOTATION OF PHENOTYPIC ABNORMALITIES VIA SEMANTIC LATENT REPRESENTATIONS ON ELECTRONIC HEALTH RECORDS**

- **Jingqing Zhang**, Xiaoyu Zhang, Kai Sun, Xian Yang, Chengliang Dai, Yike Guo. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).

### **INTEGRATED MULTI-OMICS ANALYSIS USING VARIATIONAL AUTOENCODERS: APPLICATION TO PAN-CANCER CLASSIFICATION**

- Xiaoyu Zhang, **Jingqing Zhang**, Kai Sun, Xian Yang, Chengliang Dai, Yike Guo. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).

### **INTEGRATING SEMANTIC KNOWLEDGE TO TACKLE ZERO-SHOT TEXT CLASSIFICATION**

- **Jingqing Zhang\***, Piyawat Lertvittayakumjorn\*, Yike Guo. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), 2019. **Oral**.

### **DEEP SEQUENCE LEARNING WITH AUXILIARY INFORMATION FOR TRAFFIC PREDICTION**

- Binbing Liao\*, **Jingqing Zhang\***, Chao Wu, Douglas McIlwraith, Tong Chen, Shengwen Yang, Yike Guo, and Fei Wu. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2018.

## **DEST-RESNET: A DEEP SPATIOTEMPORAL RESIDUAL NETWORK FOR HOTSPOT TRAFFIC SPEED PREDICTION**

- Binbing Liao, **Jingqing Zhang**, Ming Cai, Siliang Tang, Yifan Gao, Chao Wu, Shengwen Yang, Wenwu Zhu, Yike Guo, Fei Wu. In Proceedings of the 2018 ACM on Multimedia Conference. ACM, 2018.

## **THE DEEP POINCARÉ MAP: A NOVEL APPROACH FOR LEFT VENTRICLE SEGMENTATION**

- Yuanhan Mo, Fangde Liu, Douglas McIlwraith, Guang Yang, **Jingqing Zhang**, Taigang He, and Yike Guo. The 21st International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI, 2018.

## **I2T2I: LEARNING TEXT TO IMAGE SYNTHESIS WITH TEXTUAL DATA AUGMENTATION**

- Hao Dong, **Jingqing Zhang**, Douglas McIlwraith, Yike Guo. International Conference on Image Processing, ICIP, 2017.

## **TENSORDB: DATABASE INFRASTRUCTURE FOR CONTINUOUS MACHINE LEARNING**

- Fangde Liu, Axel Oehmichen, **Jingqing Zhang**, Kai Sun, Hao Dong, Yuanhan Mo, and Yik Guo. International Conference Artificial Intelligence, ICAI, 2017.

\* Equal contribution.

## **Projects**

### **PEGASUS**

- ~1.3k stars on GitHub. Cutting-edge pre-trained abstractive summarization models.  
[github.com/google-research/pegasus](https://github.com/google-research/pegasus)

### **TENSORLAYER 2.0**

- ~7k stars on GitHub. Leading the development of TensorLayer 2.0, a novel TensorFlow-based deep learning and reinforcement learning library designed for researchers and engineers. TensorLayer was awarded the 2017 Best Open Source Software by the ACM Multimedia Society.  
[github.com/tensorlayer/tensorlayer](https://github.com/tensorlayer/tensorlayer)

### **KDDCUP 2019 AUTOML TRACK**

- A solution for the KDDCUP 2019 AutoML Track using LightGBM.

## **Patents**

### **TRAINING TEXT SUMMARIZATION NEURAL NETWORKS WITH AN EXTRACTED SEGMENTS PREDICTION OBJECTIVE**

- United States application or PCT international application number 16/869,419

## **Professional Activities**

- Reviewer of conferences: ACL, EMNLP, NAACL, ACL Rolling Review, MICCAI
- Reviewer of journals: TKDE, Cognitive Systems Research, Bioinformatics, Artificial Intelligence Review
- Programme Committee member: BioNLP

## Awards & Scholarships

- PhD Scholarship supported by LexisNexis HPCC Systems Academic Program 2016-2020
- National Scholarship of China (top 1%) in 2015
- Tsinghua & Tung OOCL Scholarship (top 10%) in 2014
- Tsinghua & Evergrand Scholarship (top 10%) in 2013

## Skills

- Computer skills: Python, C/C++/C#, Java, JavaScript, HTML, CSS, Linux
- Machine Learning Frameworks: TensorFlow 1.0 & 2.0, PyTorch, TensorLayer, Scikit-learn
- GitHub: <https://github.com/JingqingZ>

Updated: 2022-04