

# Linqing Liu

90 High Holborn, London WC1V 6LJ, UK  
linqing.liu.20@ucl.ac.uk | likicode@gmail.com  
http://likicode.com

## EDUCATION

---

### University College London

Ph.D. in *Computer Science*

London, UK

Sept.2020 - June.2024 (*Expected*)

- Supervisors: Pontus Stenetorp, Sebastian Riedel
- Area: Natural Language Processing

### University of Waterloo

MMath in *Computer Science (thesis)*

Waterloo, Canada

Sept.2018 - July.2020

- Supervisor: Prof. Jimmy Lin
- Related Courses: *Theory of Deep Learning, Data-Intensive Distributed Computing, Optimization for Data Science, IoT & Intelligent Connectivity*

### Tongji University

B.Eng. in *Software Engineering*

Shanghai, China

Sept.2013 - June.2017

- Related Courses: *Data Mining & Analysis, Data Warehouse Technology, Object-Oriented Programming, Data Structures, Operating System, Compiler Principle, Computer Architecture, Software Engineering*

## PUBLICATIONS

---

1. **Challenges in Generalization in Open Domain Question Answering**  
[Linqing Liu](#), Patrick Lewis, Sebastian Riedel, Pontus Stenetorp. Arxiv preprint 2021
2. **PAQ: 65 Million Probably-Asked Questions and What You Can Do With Them**  
Patrick Lewis, Yuxiang Wu, [Linqing Liu](#), Pasquale Minervini, Heinrich Kuttler, Aleksandra Piktus, Pontus Stenetorp, Sebastian Riedel. Arxiv Preprint, 2021
3. **Controllable Abstractive Dialogue Summarization with Sketch Supervision**  
Chien-Sheng Wu\*, [Linqing Liu](#)\*, Wenhao Liu, Pontus Stenetorp, Caiming Xiong. (\*Equal Contribution) **ACL-IJCNLP Findings 2021**
4. **MKD: a Multi-Task Knowledge Distillation Approach for Pretrained Language Models**  
[Linqing Liu](#), Huan Wang, Jimmy Lin, Richard Socher and Caiming Xiong. Arxiv Preprint 2020
5. **Incorporating Contextual and Syntactic Structures Improves Semantic Similarity Modeling**  
[Linqing Liu](#), Wei Yang, Jinfeng Rao, Raphael Tang, Jimmy Lin. **EMNLP 2019**
6. **Bridging the Gap between Relevance Matching and Semantic Matching with Hierarchical Co-Attention Network**  
Jinfeng Rao, [Linqing Liu](#), Yi Tay, Wei Yang, Peng Shi, Jimmy Lin. **EMNLP 2019**
7. **Distilling Task-Specific Knowledge from BERT into Simple Neural Networks**  
Raphael Tang\*, Yao Lu\*, [Linqing Liu](#)\*, Lili Mou, Olga Vechtomova, Jimmy Lin. (\* Equal Contribution) Arxiv preprint 2019
8. **Generative Adversarial Network for Abstractive Text Summarization**  
[Linqing Liu](#), Yao Lu, Min Yang, Qiang Qu, Jia Zhu. **AAAI abstract 2018**.
9. **Detecting Smart Spammers On Social Network: A Topic Model Approach** [Linqing Liu](#), Yao Lu, Ye Luo, Renxian Zhang, Jianwei Lu, Laurent Itti. **NAACL-HLT Student Session 2016**

## HONOURS & AWARDS

---

- 20' **EfficientQA NeurIPS competition**, win 2 tracks with the UCL-FAIR team
- 18' **David R. Cheriton Graduate Scholarship**, University of Waterloo, based on academic excellence
- 18' **AAAI Student Scholarship**, AAAI
- 16' **Google Anita Borg Memorial Scholarship: Asia-Pacific**, Google Inc (only 9 recipients in China)
- 16' **Top 3 of Beauty of Programming Competition 2016**, Microsoft Inc (Top 0.3%)

## WORK EXPERIENCES

---

### Salesforce Research

**Palo Alto, USA**

Research Intern

*Sept. 2019 - Aug. 2020*

- Supervisor: Caiming Xiong
- *Controllable Abstractive Dialogue Summarization*: We propose a two-stage generation strategy that generates a preliminary summary sketch serving as the basis for the final summary, an additional length-controllable generation method specifically for dialogue summarization. Our model achieves state-of-the-art performance on the largest dialogue summarization corpus.
- *Multi-Task Knowledge Distillation Approach for Pretrained Language Models*: We propose to distill the student model from different tasks jointly. We evaluate our approach on two different student model architectures. Compared with other approaches, the LSTM based student with bi-attention mechanism achieves better performance under the same computational constraints, another Transformer-based student outperforms other KD methods without using external training data.

### Chinese Academy of Sciences (CAS)

**Shenzhen, China**

Research Intern

*July. 2017 - Dec. 2017*

- Supervisor: Min Yang
- *Generative Adversarial Network for Abstractive Text Summarization*: We propose an adversarial process for abstractive text summarization, in which we simultaneously train a generative model G (as an agent of reinforcement learning, which takes the raw text as input and predicts the abstractive summarization) and a discriminative model D which attempts to distinguish the generated summary from the ground truth summary.

## PROJECTS

---

### Efficient Open-Domain Question Answering

**EfficientQA Competition**

*Sept.2020 - Dec.2020*

- We generate and index a large corpus of question-answer pairs and retrieve the most similar question to the input question. This approach requires low parameter-count models and less space to store qa-pairs. With additional system compression techniques, we win two tracks: best accuracy while system size under 500MiB and smallest system to get 25% accuracy.

### Text-to-Image Information Retrieval from Large-Scale Web Archives

**JCDL demo**

*Nov.2018 - Jan.2019*

- The first integration of deep learning models with a toolkit for exploring web archives (Archives Unleashed Toolkit) to support content-based image analysis at scale. By broadcasting the pretrained model to all Spark executors in parallel, we are able to extract images of common objects from a 4TB web archive of GeoCities, which we then compile into browsable collages.

## TECHNICAL STRENGTH

---

**Proficiency:** Python (PyTorch, NumPy, spaCy, HuggingFace Transformers), L<sup>A</sup>T<sub>E</sub>X, Shell, SLURM.

**Prior Experience:** HTML/CSS, Javascript, Tensorflow, Matlab, R, Scala, MapReduce, Spark, Cocos2d.