

# YAN XU

Email: [yxucb@connect.ust.hk](mailto:yxucb@connect.ust.hk) ◇ Tel: [+86 14715496316](tel:+8614715496316)

## EDUCATION

---

**The Hong Kong University of Science and Technology**

Ph.D. in ECE

Hong Kong, China

*Sept. 2018 - Present*

**Xi'an Jiaotong University**

B.Sc. in Automation (Rank: 3/177)

Shaanxi, China

*Sept. 2014 - Jun. 2018*

## WORK EXPERIENCE

---

**Applied Scientist Intern**

Alexa AI, Amazon

Remote

*Nov. 2021 - Feb. 2022*

- Worked individually on knowledgeable language model pre-training project, which is to inject additional knowledge into pre-trained LMs through continued pre-training to benefit knowledge-intensive tasks. The paper is accepted in ACL 2023.
- Mentor/Manager: Mahdi Namazifar and Dilek Hakkani-Tür

**Research Intern**

Institute of Automation Chinese Academy of Sciences

Beijing, China

*Jul. 2016 - Aug. 2016 & Jul. 2017 - Aug. 2017*

- Developed an Android application to demonstrate the image captioning algorithm for online use.
- Conducted research on text skew correction with Convolutional Neural Networks.

## SELECTED PUBLICATIONS

---

**KILM: Knowledge Injection into Encoder-Decoder Language Models**

ACL 2023

Yan Xu, Mahdi Namazifar, Devamanyu Hazarika, Aishwarya Padmakumar, Yang Liu, Dilek Hakkani-Tür

**Diverse and Faithful Knowledge-Grounded Dialogue Generation via Sequential Posterior Inference**

ICML 2023

Yan Xu<sup>\*</sup>, Deqian Kong<sup>\*</sup>, Dehong Xu, Ziwei Ji, Bo Pang, Pascale Fung, Ying Nian Wu

**Retrieval-Free Knowledge-Grounded Dialogue Response Generation with Adapters**

ACL 2022 DialDoc Workshop

Yan Xu<sup>\*</sup>, Etsuko Ishii<sup>\*</sup>, Zihan Liu, Genta Indra Winata, Dan Su, Andrea Madotto, Pascale Fung

**CAiRE in DialDoc21: Data Augmentation for Information-Seeking Dialogue System**

ACL 2021 DialDoc Workshop

Yan Xu<sup>\*</sup>, Etsuko Ishii<sup>\*</sup>, Genta Indra Winata, Zhaojiang Lin, Andrea Madotto, Zihan Liu, Peng Xu, Pascale Fung

**Multi-hop Question Generation with Graph Convolutional Network**

EMNLP 2020 Findings

Dan Su, Yan Xu, Wenliang Dai, Ziwei Ji, Tiezheng Yu, Pascale Fung

**Learning Knowledge Bases with Parameters for Task-Oriented Dialogue Systems**

EMNLP 2020 Findings

Andrea Madotto, Samuel Cahyawijaya, Genta Indra Winata, Yan Xu, Zihan Liu, Zhaojiang Lin, Pascale Fung

**CAiRE-COVID: A Question Answering and Query-Focused Multi-document Summarization System for Covid-19 Scholarly Information Management**

EMNLP 2020 NLP-COVID Workshop

Dan Su, Yan Xu, Tiezheng Yu, Farhad Bin Siddique, Elham J Barezi, Pascale Fung

**Generalizing Question Answering System with Pre-trained Language Model Fine-tuning**

EMNLP 2019 MRQA

Workshop

Yan Xu<sup>\*</sup>, Dan Su<sup>\*</sup>, Genta Indra Winata, Peng Xu, Hyeondey Kim, Zihan Liu, Pascale Fung

**Incorporating word and subword units in unsupervised machine translation using language model rescoring**

WMT

2019

Yan Xu<sup>\*</sup>, Zihan Liu<sup>\*</sup>, Genta Indra Winata, Pascale Fung

## SELECTED AWARDS

---

**Best Student Paper Award** in DialDoc22 Workshop

ACL 2022

**Third Place** in DialDoc21 Shared Task

ACL 2021

**Winner** of Task 10 in Kaggle CORD-19 Challenge

Kaggle Competition

**Fourth Place** in MRQA-2019 Shared Task

EMNLP 2019

**First Place** in Chatbot Millionaire Challenge

HKSTP 2019

**Top Four** out of 160 teams in SemEval-2019 Task 3

NAACL 2019

National Scholarship (**Top 2%**)

Ministry of Education 2016

National Scholarship (**Top 2%**)

Ministry of Education 2015

### **ACADEMIC SERVICE AND LANGUAGE SKILLS**

---

**Reviewer** ACL 2021, EMNLP 2021, NLPCC 2021, Rep4NLP 2022, ARR, NLPCC 2022, EMNLP 2022

**Languages** Chinese (Mother Tongue) • English (Full-Proficiency) • Cantonese (Basic)