

Yuting Yang (杨玉婷)

yangyuting@ict.ac.cn

Ph.D. Candidate:

At Institute of Computing Technology, Chinese Academy of Sciences (ICT, CAS)

Graduation time: July 2023

Advisor: Professor Jintao Li

Visiting Research Scholar:

2021.03~2022.03 NExT++ Research Centre, National University of Singapore (NUS)

Advisor: Professor Tat-Seng Chua

Interests:

Trustworthy NLP (Adversarial Attack and Defense, Neural Network Verification)

Dialogue system

Trustworthy NLP (Main)

- Rethink the robustness of deep neural networks from a quantification perspective and propose the concept of “**weak robustness**”. (*ISSTA*)
- Based on the weak robustness, propose a **model-agnostic method for enhancing robustness**, which fights perturbation with perturbation. (*AAAI*)
- Further apply weak robustness combined with ensemble learning to **resist a variety of text attacks**. (*Under review*)
- Utilize **prompting to explore robustness defects** of deep NLP models. (*FCS Journal*)

Dialogue System

- Design a **dual prompt learning framework for few-shot dialog state tracking** task. (*WWW*)

Text Generation

- Realize a news quality assessment system based on linguistic style.
- Realize guideline-based news headline generation system, which incorporates nondifferentiable professional writing guidelines into automatic generation via reinforcement learning.

Both systems are patented and applied in practical applications of Chinese official news media ([Xinhua Net](#) and [People's Daily](#))

