

WANYU DU

Email: wd5jq@virginia.edu; Homepage: [wyu-du.github.io](https://github.com/wyu-du)

RESEARCH INTERESTS

My research interest is natural language generation. In particular, I am interested in few-shot natural language generation, open-domain dialogue generation and evaluation, deep latent-variable models for stylistic text generation.

EDUCATION

University of Virginia (UVa), US <i>PhD student in Computer Science, Advisor: Yangfeng Ji</i>	August 2020 - Present
University of Virginia (UVa), US <i>Master of Science in Computer Science</i>	August 2018 - May 2020
Dongbei University of Finance and Economics (DUFE), China <i>Bachelor of Management in E-Commerce (Top 10%)</i>	August 2014 - July 2018

PUBLICATIONS

- **Wanyu Du** and Yangfeng Ji. SideControl: Controlled Open-domain Dialogue Generation via Additive Side Networks, *Findings of the Association for Computational Linguistics: EMNLP 2021*, 2021.
- Stephanie Schoch, **Wanyu Du**, Yangfeng Ji. Contextualizing Variation in Text Style Transfer Datasets, *Proceedings of the 14th International Conference on Natural Language Generation (INLG 2021)*, 2021.
- Sebastian Gehrmann et al., The GEM Benchmark: Natural Language Generation, its Evaluation and Metrics, *arXiv*, 2021.
- **Wanyu Du**, Jianqiao Zhao, Liwei Wang and Yangfeng Ji. Diverse Text Generation via Variational Encoder-Decoder Models with Gaussian Process Priors, *ARR under review*, 2021.
- **Wanyu Du** and Yangfeng Ji. An Empirical Comparison on Imitation Learning and Reinforcement Learning for Paraphrase Generation, *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)*, 2019.
- Xinyu Wang, Tsan-Ming Choi, Xiaohang Yue, Mengji Zhang, and Wanyu Du. An Effective Optimization Algorithm for Application Mapping in Network-on-Chip Designs, *IEEE Transactions on Industrial Electronics*, 2019.

RESEARCH EXPERIENCE

Iterative Text Editing <i>NLP Research Intern</i>	June 2021 - August 2021 <i>Grammarly Engineering</i>
---	---

- Proposed a novel approach to model human iterative text editing behaviors.
- Constructed new data resources for iterative text editing tasks.
- Analyzed human iterative text editing patterns and designed new iterative text editing systems.

Controlled Open-domain Dialogue Generation <i>Research Assistant</i>	January 2021 - May 2020 <i>UVa ILP Lab</i>
--	---

- Proposed a novel approach to control the generation of Transformer-based pretrained language models under the low-data setting.

- Designed novel control attributes losses to incorporate useful control signals into large pretrained language models.
- Conducted empirical experiments to show that our approach has better controllability, higher generation quality and better sample-efficiency than existing gradient-based and weighted-decoding methods. [EMNLP 2021 (Findings)]

Communicative Intent Flow for Multi-turn Dialogue Evaluation

June 2020 - Dec 2020

NLP Research Intern

Tencent AI Lab

- Proposed a novel evaluation dimension, communicative intent flow, to evaluation the coherence between interactive conversations in open-domain dialogues.
- Designed a communicative intent annotation schema and conducted crowd-sourcing human annotations on two benchmark open-domain dialogue datasets.
- Conducted empirical experiments to analyze the correlation between communicative intent flow and the overall quality of multi-turn conversations, and showed that the communicative intent flow is an essential evaluative dimension to multi-turn conversations.

Gaussian Process Priors for Diverse Text Generation

June 2019 - May 2020

Research Assistant

UVa ILP Lab

- Proposed a novel approach to enrich contextual representation learning of encoder-decoder models by introducing a stochastic function to map encoder hidden states into random context variables.
- Applied Gaussian process to explicitly model the dependency among random context variables.
- Conducted empirical experiments on downstream text generation tasks (e.g. paraphrase generation, text style transfer and dialog generation), and demonstrated that the proposed method outperforms baseline models in both quality and diversity.

Efficient Learning Algorithms for Paraphrase Generation

January 2019 - May 2019

Research Assistant

UVa ILP Lab

- Developed a unified framework for different learning algorithms (e.g. REINFORCE, DAgger, etc.) in a sequence-to-sequence model.
- Proposed variant learning algorithms based on reinforcement learning, and conducted empirical comparison on two benchmark datasets to see which algorithms would best alleviate exposure bias.
- Figured out the most effective learning algorithm for paraphrase generation, which outperformed the state-of-art model with 12.21 BLEU score. [EMNLP 2019]

HONORS

- **UVa Academic Excellence Fellowship**, 2018.
- **DUFE Outstanding Graduates**, 2018.
- **DUFE First Place Scholarship**, 2014-2018.
- **National Data Driven Innovation Research Competition: National 2nd Prize**, 2018.
- **Mitsubishi UFJ Trust Scholarship**, 2017.
- **Citi Financial Innovation Application Competition: National 3rd Prize**, 2016.

PROFESSIONAL ACTIVITIES

- Organizing team member, The GEM Workshop at ACL 2021, 2021.
- Reviewer, EMNLP 2021, ACL 2021.

- Talk on "Efficient Learning Algorithms for Paraphrase Generation", UVa Artificial Intelligence and Machine Learning Seminar, 2019.