

Xinyi (Cindy) Wang

Education

- 08/2017–06/2022 **Carnegie Mellon University, Pittsburgh, PA.**
Ph.D., Language Technologies
Advisor: Graham Neubig
GPA: 4.00/4.33
- 08/2013–05/2017 **University of Notre Dame, Notre Dame, IN.**
summa cum laude, B.S. in Computer Science, B.A. in Economics
GPA: 3.99/4.00

Experience

- September 2022–Now **Google DeepMind, Research Scientist**, New York, NY.
Factuality of LLMs in Google's Generative Search Experience
Gemini training and evaluation
- Summer 2020 **Google Translate, Research Intern**, Remote from Pittsburgh.
Hosts: Ankur Bapna, Melvin Johnson, Orhan Firat
- Summer 2019 **Facebook AI Research, Research Intern**, New York, NY.
Hosts: Yacine Jernite, Jason Weston, Michael Auli

Honors and Awards

- 2020 **Apple Scholars in AI/ML Fellowship.**
- 2019 **Best Demo Paper Runner-up, NAACL.**
- 2017 **Outstanding CS Graduate, University of Notre Dame.**
- 2017 **Theodore S. Weber Award for the economics graduate, University of Notre Dame.**
- 2016–2017 **Tau Beta Pi Engineering Scholarship, University of Notre Dame.**
- Fall 2015 **Grace Hopper Scholarship, Travel Award to Grace Hopper Conference.**
- 2013–2016 **Dean's List, All Semesters, University of Notre Dame.**
- 2013 **Greater China Scholarship, University of Notre Dame.**

Publication

- 2023 **Xinyi Wang**, John Wieting, Jonathan H Clark. FIAT: Fusing learning paradigms with Instruction-Accelerated Tuning. In Arxiv
- 2023 Jonas Pfeiffer, Francesco Piccinno, Massimo Nicosia, **Xinyi Wang**, Machel Reid, Sebastian Ruder. mMT5: Modular Multilingual Pre-Training Solves Source Language Hallucinations. In EMNLP
- 2023 Benjamin Muller, John Wieting, Jonathan H Clark, Tom Kwiatkowski, Sebastian Ruder, Livio Baldini Soares, Roei Aharoni, Jonathan Herzig, **Xinyi Wang**. Evaluating and Modeling Attribution for Cross-Lingual Question Answering. In EMNLP

- 2023 Sebastian Ruder, Jonathan H Clark, Alexander Gutkin, Mihir Kale, Min Ma, Massimo Nicosia, Shruti Rijhwani, Parker Riley, Jean-Michel A Sarr, **Xinyi Wang**, John Wieting, Nitish Gupta, Anna Katanova, Christo Kirov, Dana L Dickinson, Brian Roark, Bidisha Samanta, Connie Tao, David I Adelani, Vera Axelrod, Isaac Caswell, Colin Cherry, Dan Garrette, Reeve Ingle, Melvin Johnson, Dmitry Panteleev, Partha Talukdar. XTREME-UP: A User-Centric Scarce-Data Benchmark for Under-Represented Languages. In EMNLP
- 2022 **Xinyi Wang**, Sebastian Ruder, Graham Neubig. Expanding Pretrained Models to Thousands More Languages via Lexicon-based Adaptation . In ACL
- 2022 Brian Yan, Patrick Fernandes, Siddharth Dalmia, Jiatong Shi, Yifan Peng, Dan Berrebbi, **Xinyi Wang**, Graham Neubig, Shinji Watanabe. CMUs IWSLT 2022 Dialect Speech Translation System. 2022. In IWSLT.
- 2021 **Xinyi Wang**, Yulia Tsvetkov, Sebastian Ruder, Graham Neubig. Efficient Test Time Adapter Ensembling for Low-resource Language Varieties. In EMNLP Findings
- 2021 **Xinyi Wang**, Sebastian Ruder, Graham Neubig. Multi-view Subword Regularization. In NAACL
- 2021 Hieu Pham*, **Xinyi Wang***, Yiming Yang, Graham Neubig. Meta Back Translation. In ICLR
- 2020 **Xinyi Wang**, Yulia Tsvetkov, Graham Neubig. Balancing Training for Multilingual Neural Machine Translation. In ACL
- 2020 **Xinyi Wang***, Hieu Pham*, Paul Mitchel, Antonis Anastasopoulos, Jaime Carbonell, Graham Neubig. Optimizing Data Usage via Differentiable Rewards. In ICML
- 2020 Junxian He*, **Xinyi Wang***, Graham Neubig, Taylor Berg-Kirkpatrick. A Probabilistic Formulation of Unsupervised Text Style Transfer. In ICLR
- 2020 Luyu Gao, **Xinyi Wang**, Graham Neubig. Improving Target-side Lexical Transfer in Multilingual Neural Machine Translation. 2020. In Findings of EMNLP.
- 2019 **Xinyi Wang**, Graham Neubig. Target Conditioned Sampling: Optimizing Data Selection for Multilingual Neural Machine Translation. In ACL
- 2019 **Xinyi Wang**, Hieu Pham, Philip Arthur, Graham Neubig. Multilingual Neural Machine Translation with Soft Decoupled Encoding. In ICLR
- 2019 Zi-Yi Dou, **Xinyi Wang**, Junjie Hu, Graham Neubig. Domain Differential Adaptation for Neural Machine Translation. 2019. In WNGT.
- 2019 Graham Neubig, Zi-Yi Dou, Junjie Hu, Paul Michel, Danish Pruthi, **Xinyi Wang**, John Wieting. compare-mt: A Tool for Holistic Comparison of Language Generation Systems. 2019. In NAACL. **Best Demo Runner-up**
- 2018 **Xinyi Wang**, Hieu Pham, Pengcheng Yin, Graham Neubig. A Tree-based Decoder for Neural Machine Translation. 2018. In EMNLP.
- 2018 **Xinyi Wang***, Hieu Pham*, Zihang Dai, Graham Neubig. SwitchOut: an Efficient Data Augmentation Algorithm for Neural Machine Translation. 2018. In EMNLP.
- 2018 **Xinyi Wang**, Salvador Aguinaga, Tim Weninger, David Chiang. Growing Better Graphs With Latent-Variable Probabilistic Graph Grammars. 2018. In KDD Workshop on Mining and Learning with Graphs.
- 2018 Graham Neubig, Matthias Sperber, **Xinyi Wang**, Matthieu Felix, Austin Matthews, Sarguna Padmanabhan, Ye Qi, Devendra Singh Sachan, Philip Arthur, Pierre Godard, John Hewitt, Rachid Riad, Liming Wang. XNMT: The eXtensible Neural Machine Translation Toolkit. 2018. In AMTA.

Professional Services

Program Committee **WIT 2022, EMNLP, NAACL, MTSummit, AMTA**

Reviewer **ACL, ACL Rolling review, NeurIPS (outstanding reviewer 2020), ICML, ICLR, CSCW**

Journal Reviewer **IEEE Transactions on Audio, Speech and Language Processings**

Teaching

Fall 2020	Multilingual Natural Language Processing	<i>Carnegie Mellon University, TA</i>
Fall 2019	Machine Translation and Seq-to-seq Models	<i>Carnegie Mellon University, TA</i>
Fall 2016	Design and Analysis of Algorithms	<i>University of Notre Dame, TA</i>
Spring 2016	Theory of Computing	<i>University of Notre Dame, TA</i>