

# Victoria X. Lin

SENIOR RESEARCH SCIENTIST

1 Hacker Way, Menlo Park, CA 94025, USA

✉ victorialin@fb.com | 🏠 victorialin.net | 📷 todpole3 | 🌐 xivictorialin | 🎓 victorialin | 🐦 @VictoriaLinML

## Research Interest

Natural Language Understanding, Few-shot Learning, Self-supervised Learning

## Work Experience

### Facebook AI Research

SENIOR RESEARCH SCIENTIST

- Research focusing on large-scale pre-training, few-shot learning and multilinguality

Menlo Park, CA, USA

Jan. 2021 - present

### Salesforce Research

SENIOR RESEARCH STAFF

- Natural language understanding research focusing on semantic parsing, natural language interface and question answering

Palo Alto, CA, USA

Jun. 2018 - Dec. 2020

RESEARCH SCIENTIST

- Natural language understanding research focusing on knowledge graph reasoning and question answering

Oct. 2017 - Jun. 2018

### Microsoft Research

RESEARCH INTERN

Redmond, WA, USA

Jun. 2015 - Sep. 2015

### Allen Institute for Artificial Intelligence

RESEARCH INTERN

Seattle, WA, USA

Jul. 2014 - Sep. 2014

## Conference Publications

### FeTaQA: Free-form Table Question Answering

Linyong Nan, Chiachun Hsieh, Ziming Mao, [Xi Victoria Lin](#), Neha Verma, Rui Zhang, Wojciech Kryściński, Nick Schoelkopf, Riley Kong, Xiangru Tang, Murori Mutuma, Ben Rosand, Isabel Trindade, Renusree Bandaru, Jacob Cunningham, Caiming Xiong, Dragomir Radev.

TACL 2021

### GraPPa: Grammar-Augmented Pre-Training for Table Semantic Parsing

Tao Yu, Chien-Sheng Wu, [Xi Victoria Lin](#), Bailin Wang, Yi Chern Tan, Xinyi Yang, Dragomir Radev, Richard Socher, Caiming Xiong

ICLR 2021

### Learning to Synthesize Data for Semantic Parsing.

Bailin Wang, Wenpeng Yin, [Xi Victoria Lin](#) and Caiming Xiong.

NAACL 2021 short

### DART: Open-Domain Structured Data Record to Text Generation

Linyong Nan, Dragomir Radev, Rui Zhang, Amrit Rau, Abhinand Sivaprasad, Chiachun Hsieh, Xiangru Tang, Aadit Vyas, Neha Verma, Pranav Krishna, Yangxiaokang Liu, Nadia Irwanto, Jessica Pan, Faiaz Rahman, Ahmad Zaidi, Mutethia Mutuma, Yasin Tarabar, Ankit Gupta, Tao Yu, Yi Chern Tan, [Xi Victoria Lin](#), Caiming Xiong, Richard Socher and Nazneen Fatema Rajani.

NAACL 2021

### Bridging Textual and Tabular Data for Cross-Domain Text-to-SQL Semantic Parsing

[Xi Victoria Lin](#), Richard Socher, Caiming Xiong

EMNLP Findings 2020

### Photon: A Robust Cross-Domain Text-to-SQL System

Jichuan Zeng\*, [Xi Victoria Lin](#)\*, Caiming Xiong, Richard Socher, Michael R. Lyu, Irwin King, Steven C.H. Hoi

ACL Demo 2020

### Double-Hard Debias: Tailoring Word Embeddings for Gender Bias Mitigation

Tianlu Wang, [Xi Victoria Lin](#), Nazeen Fatema Rajani, Bryan McCann, Vicente Ordonez and Caiming Xiong

ACL 2020

### CoSQL: A Conversational Text-to-SQL Challenge Towards Cross-Domain Natural Language Interfaces to Databases

Tao Yu, Rui Zhang, Heyang Er, Suyi Li, Eric Xue, Bo Pang, [Xi Victoria Lin](#), Yi Chern Tan, Tianze Shi, Zihan Li, Youxuan Jiang, Michihiro Yasunaga, Sungrok Shim, Tao Chen, Alexander Fabbri, Zifan Li, Luyao Chen, Yuwen Zhang, Shreya Dixit, Vincent Zhang, Caiming Xiong, Richard Socher, Walter Lasecki and Dragomir Radev

EMNLP 2019

## Editing-based SQL Query Generation for Cross-Domain Context-Dependent Questions

Rui Zhang, Tao Yu, Heyang Er, Sungrok Shim, Eric Xue, [Xi Victoria Lin](#), Tianze Shi, Caiming Xiong, Richard Socher and Dragomir Radev

EMNLP 2019

## SParC: Cross-Domain Semantic Parsing in Context

Tao Yu, Rui Zhang, Michihiro Yasunaga, Yi Chern Tan, [Xi Victoria Lin](#), Suyi Li, Heyang Er, Irene Li, Bo Pang, Tao Chen, Emily Ji, Shreya Dixit, David Proctor, Sungrok Shim, Jonathan Kraft, Vincent Zhang, Caiming Xiong, Richard Socher, Dragomir Radev

ACL 2019

## Multi-Hop Knowledge Graph Reasoning with Reward Shaping

[Xi Victoria Lin](#), Richard Socher and Caiming Xiong

EMNLP 2018

## NL2Bash: A Corpus and Semantic Parser for Natural Language Interface to the Linux Operating System

[Xi Victoria Lin](#), Chenglong Wang, Luke Zettlemoyer and Michael D. Ernst

LREC 2018

## Compositional Learning of Embeddings for Relation Paths in Knowledge Bases and Text

Kristina Toutanova, [Xi Victoria Lin](#), Scott Wen-tau Yih, Hoifung Poon and Chris Quirk

ACL 2016

## Workshop and Other Publications

---

### Testing Cross-Database Semantic Parsers Using Canonical Utterances.

Heather Lent, Semih Yavuz, Tao Yu, Tong Niu, Yingbo Zhou, Dragomir Radev, [Xi Victoria Lin](#).

Eval4NLP 2021

### NeurIPS 2020 NLC2CMD Competition: Translating Natural Language to Bash Commands.

Mayank Agarwal, Tathagata Chakraborti, Quchen Fu, David Gros, [Xi Victoria Lin](#), Jaron Maene, Kartik Talamadupula, Zhongwei Teng, Jules White.

NeurIPS 2020  
Competition Track

### ColloQL: Robust Text-to-SQL Over Search Queries

Karthik Radhakrishnan, Arvind Srikantan, [Xi Victoria Lin](#)

Intex-Sempar 2020

### Program Synthesis from Natural Language Using Recurrent Neural Networks

[Xi Victoria Lin](#), Chenglong Wang, Deric Pang, Kevin Vu, Luke Zettlemoyer, Michael D. Ernst

UWCSE-TR 2017

### Multi-label Learning with Posterior Regularization

[Xi Victoria Lin](#), Sameer Singh, Luheng He, Ben Taskar, and Luke Zettlemoyer

NeurIPS-MLNLP 2014

### Fine-grained Named Entity Classification in Machine Reading

[Xi Victoria Lin](#)

Oxford Thesis 2011

## Patents

---

### Multi-hop knowledge graph reasoning with reward shaping

[Xi Victoria Lin](#), Richard Socher, Caiming Xiong

US Patent App. 16/051,309

## Talks

---

### The Age of Data Conversation: Talk to Your Structured Data (invited talk)

Knowledge ConneXions 2020

### Creating The Future Of AI: How Salesforce Research Advances AI For CRM (co-speaker)

Dreamforce 2019

## Technical Writing

---

### Talk to Your Data: One Model, Any Relational Database.

Einstein.ai 2020

## Software

---

### Photon v1.1: <https://naturalsql.com/>

Salesforce, 2020

Photon is a deep learning based cross-domain natural language interface to databases that focuses on factual look-up questions. It allows end users to query a number of relational DBs in natural language, including DBs it has never been trained on.

Tellina is an end-user scripting assistant that can be queried via natural language. It translates a natural language sentence typed by the user into a piece of short, executable script.

## Education

---

### University of Washington

PH.D. GRADUATED AS M.SC. IN COMPUTER SCIENCE

Seattle, WA, USA

Aug. 2017

### University of Pennsylvania

PH.D. GRADUATED AS M.SC. IN COMPUTER SCIENCE

Philadelphia, PA, USA

May. 2013

### University of Oxford

M.Sc. IN COMPUTER SCIENCE

Oxford, UK

Sep. 2011

### The Hong Kong Polytechnic University

B.ENG. IN ELECTRONIC AND INFORMATION ENGINEERING

Kowloon, HK

Aug. 2010

### Xi'an Jiaotong University

SPECIAL CLASS FOR THE GIFTED YOUNG

Xi'an, Shaanxi, CN

Jan. 2008

## Honors & Awards

---

|      |   |                        |
|------|---|------------------------|
| 2014 | <b>Travel Award</b> , NIPS Women in Machine Learning Workshop                                       | Montreal, Québec, CA   |
| 2011 | <b>Doctoral Fellowship</b> , University of Pennsylvania   | Philadelphia, PA, USA  |
| 2010 | <b>Best Academic Performance Award</b> , Department of Electronic and Information Engineering, HKPU | Kowloon, HK            |
| 2010 | <b>Scholarship</b> , Hong Kong SAR Government   | Kowloon, HK            |
| 2009 | <b>Scholarship</b> , Hong Kong Polytechnic University Post-Entry                                    | Kowloon, HK            |
| 2009 | <b>Scholarship</b> , Hong Kong & Kowloon Electrical Appliances Merchants Association                | Kowloon, HK            |
| 2009 | <b>Student Travel Award</b> , Apple Worldwide Developers Conference                                 | San Francisco, CA, USA |

## Service

---

### CO-CHAIR

System Demonstration Track

NAACL 2021

### ORGANIZING COMMITTEE

1st Workshop on Interactive and Executable Semantic Parsing (Intex-Sempar)

EMNLP 2020

Competition for Automatic Translation of English to Bash (NLC2CMD)

NeurIPS 2020

### PROGRAM COMMITTEE

|      |                           |
|------|---------------------------|
| 2020 | ACL, EMNLP, AACL, ACL-NLI |
| 2019 | ICML, ACL, NAACL          |
| 2018 | ACL, EMNLP, COLING, CoNLL |
| 2017 | ACL, EMNLP                |
| 2016 | EMNLP                     |
| 2015 | EMNLP                     |

## Teaching

---

### CIS 520: Machine Learning

TEACHING ASSISTANT

University of Pennsylvania

Sep. 2012 - Dec. 2012

- Making exam problems; answering Piazza questions; holding office hours; grading

## Programming Skills

---

|                      |  |
|----------------------|--|
| <b>Languages</b>     | Python, Java, C++, Matlab, R, Lua, JavaScript, HTML, CSS |
| <b>Deep Learning</b> | Pytorch, Tensorflow                                      |
| <b>Web</b>           | Django   |