

SHAYNE R. LONGPRE

(650) 690 5632

shayne.longpre@cs.stanford.edu

shaynelongpre.com

CAREER INTERESTS

Applied research in Machine Learning, Deep Learning, Natural Language Understanding, and Reinforcement Learning. Primary applications: neural reading comprehension, systematic generalization, multilinguality, and model interpretability. Also, exploring economic and competitive implications of advancements in AI.

EDUCATION

MASSACHUSETTS INSTITUTE of TECHNOLOGY | *Cambridge, MA*
Doctor of Philosophy candidate | MIT Media Lab – Social Machines

Sep. 2021 – Present

STANFORD UNIVERSITY | *Stanford, CA*
Master of Science | Computer Science (Artificial Intelligence)
Bachelor of Arts | Economics + Minor in History

Sep. 2012 – Jan. 2018

EXPERIENCE

APPLE INC. | *Seattle, WA*
Senior Applied Machine Learning Scientist | Web Ranking and Answers

Feb. 2018 – July 2021

- Applied research into passage retrieval and reranking, conversational and multilingual systems, textual semantic similarity, low-resource NLP, model compression and interpretability.
- Lead a weekly reading group covering the latest NLP research.

STANFORD NLP GROUP | *Stanford, CA*

Jun. 2017 – Dec. 2017

Research Assistant (RA) mentored by Danqi Chen and Chris Manning (<https://nlp.stanford.edu/>)

- Researched question answering models with reinforcement learning for paragraph selection (TriviaQA).
- Contributed a QA code repository to the Stanford NLP repo, parts subsequently [used](#) in [CoQA](#).

SALESFORCE RESEARCH (formerly **METAMIND**) | *Palo Alto, CA*

Jun. 2016 – Sept. 2016

Deep Learning Research Intern advised by Caiming Xiong and Richard Socher.

- Research into variants of deep sequence models for classification and language modeling tasks.

MICROSOFT | *Redmond, WA*

Jun. 2015 – Sept. 2015

Software Engineer Intern | OneDrive Mac Client Team

- Designed and implemented Finder Integration for the OneDrive Mac client (shipped Dec 2015).

TEACHING and SERVICE

Program Committee Member / Reviewer

- Conferences: **ICLR 2021**, **Apple ML Summit** (2019, 2020)
- Workshops: [MRQA](#) at EMNLP 2021, Insights from [Negative Results in NLP](#) at EMNLP 2021

Reviewer & Moderator | 1st Apple NLU Workshop

2020

Computer Vision Co-lead and Mentor for [AI4ALL](#), *Inclusion + Diversity Program*

2017

- Taught course on Computer Vision fundamentals to young women aspiring to work in AI.

Teaching Assistant (TA) for [CS231N](#): Convolutional Neural Networks for Visual Recognition | Stanford CS

2017

Teaching Assistant (TA) for [CS224N](#): Natural Language Processing with Deep Learning | Stanford CS

2017

- Weekly office hours, exam development, project advisement, and [mini-lecture](#) to 700+ students

RESEARCH

AI/ML Preprints

Pivot through English: Reliably Answering Multilingual Questions without Document Retrieval 2020
Ivan Montero, [Shayne Longpre](#), Ni Lao, Drew Frank, Chris DuBois. [[preprint](#)]

AI/ML Select Publications

MKQA: A Linguistically Diverse Benchmark for Multilingual Open Domain Question Answering 2021
[Shayne Longpre](#), Yi Lu, Jo Daiber
To appear, **TACL 2021**. [[arXiv](#)] [[26 language evaluation set](#)]

Evaluating Entity Disambiguation and the Role of Popularity in Retrieval-Based NLP 2021
Anthony Chen, Pallavi Gudipati, [Shayne Longpre](#), Xiao Ling, Sameer Singh
Presented at **ACL 2021**. [[ACLWeb](#)] [Dataset coming soon]

Open Domain Question Answering Goes Conversational via Question Rewriting | 2021
Raviteja Anantha*, Svitlana Vakulenko*, Zhucheng Tu, [Shayne Longpre](#), Stephen Pulman, Srinivas Chappidi.
Presented at **NAACL 2021**. [[ACLWeb](#)] [[Dataset](#)]

On the Transferability of Minimal Prediction Preserving Inputs in Question Answering 2021
[Shayne Longpre](#)*, Yi Lu*, Chris DuBois.
Presented at **NAACL 2021**. [[ACLWeb](#)]

Question Rewriting for Conversational Question Answering 2021
Svitlana Vakulenko, [Shayne Longpre](#), Zhucheng Tu, Raviteja Anantha.
Presented at **WSDM 2021**. [[arXiv](#)]

A Comparison of Question Rewriting Methods for Conversational Passage Retrieval 2021
Svitlana Vakulenko, Nikos Voskarides, Zhucheng Tu, [Shayne Longpre](#).
Presented at **ECIR 2021**. [[arXiv](#)]

How Effective is Task-Agnostic Data Augmentation for Pretrained Transformers? 2020
[Shayne Longpre](#)*, Yu Wang*, Chris DuBois. [Discussed in [Sebastian Ruder's NLP Newsletter](#)]
Presented at **Workshop on Insights from Negative Results** and **EMNLP (Findings) 2020**. [[ACLWeb](#)]

A Wrong Answer or a Wrong Question? 2020
Svitlana Vakulenko, [Shayne Longpre](#), Zhucheng Tu, Raviteja Anantha.
Presented at **SCAI Workshop @ EMNLP 2020**. Received **Best Paper Award**. [[ACLWeb](#)]

An Exploration of Augmentation Techniques for Domain-Agnostic Question Answering 2019
[Shayne Longpre](#)*, Yi Lu*, Zhucheng Tu*, Chris DuBois.
2nd place / 14 in shared task competition - **MRQA Workshop @ EMNLP 2019**. [[ACLWeb](#)] [[leaderboard](#)]

Economics

Invigorating Competition in Social Networking: An Interoperability Remedy 2021
Cristian Santesteban, [Shayne Longpre](#).
Published in **Competition Policy International (CPI) Antitrust Chronicle**, June 2021 Issue. [[SSRN](#)] [[CPI](#)]

How Big Data Confers Market Power to Big Tech: Leveraging the Perspective of Data Science 2020
Cristian Santesteban, [Shayne Longpre](#)
Published in **The Antitrust Bulletin** (Vol 65, Issue 3), Sept 2020. [[SSRN](#)] [[Antitrust Bulletin](#)]