

Arun Tejasvi Chaganty

Senior Research Scientist

 arunchaganty@gmail.com

 arun.chagantys.org

 github.com/arunchaganty

EXPERIENCE

Google

Senior Research Scientist

Dec 2020–(today)

- Created **dialog inpainting** (co-first author), a technique to generate millions of information-seeking conversations from documents using language models (T5 S-XXL). Implemented the entire bulk inference pipeline (average throughput of ~3k inference calls/s) using Apache Beam. Led human evaluation and safety analysis. Trained masked language models and retrieval models.
- Created the **Conversational Playlist Curation Dataset** (first author; PI), one of the first resources for conversational recommendation with multiple item ratings per-turn. Designed and implemented human-human methodology, including all annotation interfaces.
- Developed **Talk the Walk** (PI), a recipe to generate millions of (music) recommendation-seeking conversations from existing playlists using a combination of random walks and language models. Bootstrapped an end-to-end conversation recommendation system that significantly outperforms baselines in live experiments.
- Defined task and evaluation methodology for **RARR**, a post-hoc attribution and revision method for large language models (PaLM-540B).

Research Intern

Summer 2014

- Explored multi-sentence relation extraction for knowledge bases.

Square

AI Lead

May 2019–Dec 2020

- Led a small team of AI engineers that built **Square Assistant** —a chatbot we launched in October 2019 that helps customers book and reschedule appointments with Square merchants.
- Designed and shipped **conversational rescheduling** feature that increased booking and rescheduling success rates by helping customers find a concrete time for their appointment; the feature understands temporal constraints in user utterances using a model-based semantic parser.
- Developed a **type-safe domain-specific language** to describe asynchrony and interruptions in dialog flows using coroutines. Implemented Java-to-Java compiler. DSL reduced feature code 10–20x and fixed subtle asynchrony bugs.
- Developed most of the AI model deployment, logging and data annotation infrastructure.

EDUCATION

Stanford University

PhD (Computer Science) 2012-2018

Advised by Percy Liang

Indian Institute of Technology,
Madras

MTech. (Computer Science)

BTech. (Computer 2011-2012

Science) 2007-2011

Minor in Physics GPA: 9.24/10

AWARDS

- Stanford Graduate Fellow ('14-'17)
- Robert Padovani Scholar ('09)
- Google Summer of Code ('08)
- Kishore Vaigyanik Protsahan Yojana Scholar ('06-'07)

SKILLS

Natural Language Processing

Conversational AI • Recommendation

Systems • Synthetic Data Generation

• Evaluation • Retrieval •

Crowdsourcing • Semantic Parsing •

Information Extraction

Machine Learning

Deep Learning • Latent Variable

Models • Probabilistic Programming

Programming

Python (PyTorch, Tensorflow) •

Typescript (Angular, React) • SQL •

Bash • Java • C++

Computer Science

Compilers • Operating Systems •

Computer Networks • Cloud

Computing

Eloquent Labs

Head of AI

June 2018–May 2019

- Led a small team of AI engineers that built a conversational AI system for enterprise customer service. Interfaced with clients directly.
- Developed a **human-in-the-loop system to fine-tune question similarity models** for particular clients; led to 2–3x increases in precision and recall for each client.
- Startup acquired by Square in May 2019.

Stanford University

PhD Candidate

Sept 2012–Sept 2018

- Led / part of the Stanford team at TAC-KBP 2013, 2015–17. Our entry was the top-ranked at the TAC-KBP 2015–17 Cold Start tracks.
- Co-author of **CoreNLP Server**, an extremely popular API server for the Stanford CoreNLP package.
- Can we scalably evaluate open-ended language tasks like information extraction or summarization with human feedback? We show fundamental limitations with existing automatic metrics (**ACL 2018**).
- Proposed a human-in-the-loop solution for knowledge-base population evaluation that eliminates *pooling bias* using a novel importance-reweighted estimator that decreases annotation costs by a factor of 4 (**EMNLP 2017**).
- Numeric comparisons, while common in the news, are hard to identify because their definition emerges only in context. We define an explicit representation, called a *textual analogy frames*, for such comparisons and build a semantic parser to identify such frames in text (**EMNLP 2018**).
- People best understand concepts through comparisons: we provide a system to generate compositional comparisons for numerical expressions in text, such as describing Cristiano Ronaldo's signing fee of \$131 million as roughly the amount it would take to pay everyone in Kansas City the median salary for a week (**ACL 2016**).
- Can we efficiently learn latent variable models with guarantees? We show that this is possible for a variety of models satisfying a 'uniformly bottlenecked' assumption including discriminative mixtures of linear experts (**ICML 2013**), high tree-width models, log-linear models and multi-view Markov random fields (**ICML 2014**). In later work, we show guaranteed recovery for any mixture model with polynomial moments is possible via reduction to the generalized moment problem (**NIPS 2015**). All of these methods require tensor factorization, which we show can be more efficiently performed by reduction to simultaneous matrix diagonalization using random productions (**AISTATS 2015**).

Microsoft Research India

Intern

Summer 2009, 2010, 2011

- Used dynamic analysis and concolic execution to efficiently sample from probabilistic programs by avoiding invalid states in both an importance sampling and Metropolis-Hastings setting (**AISTATS 2013**).
- Applied Counter-Example Guided Abstraction Refinement, and generalization (from program analysis) to the Markov Logic Network framework, with significant performance improvements over prior art (**CAV 2013**).