

<b>EDUCATION</b>	University of Washington - <i>MS Computational Linguistics</i> 06/2018-06/2020 3.87/4 Rensselaer Polytechnic Institute - <i>BS Computer Science</i> 08/2011-05/2015 3.43/4
<b>WORK EXPERIENCE</b>	<div><div><i>Product Manager 2</i>-Microsoft Research &amp; AI, Bing11/2017-Present</div><ul style="list-style-type: none"><li>• Architected relevance metrics stack into a realtime streaming system through new human labeling tasks, novel data sampling methods, and metrics aggregation. Labeling scaled to \$7m/year across 16 markets generating more than 1,000,000 judgments a week. The pipeline went from measuring user experience years old to seconds old. To scale the pipeline created a novel multi class-stratified sampling method using the Horvitz-Thompson estimator, work under review.</li><li>• Built family of MSMARCO Datasets and baselines including QnA, Passage Ranking, and Keyphrase Extraction, which collectively have been used by over 3000 Researchers. Papers Cited over 250 times and competitive leaderboards with over 200 research submissions. Dataset creation using Pandas and baselines models built in Pytorch and Tensorflow.</li><li>• Research and analyzed search relevance and user experience for the executive team across various verticals, languages, and metrics using language embeddings, sampling methodologies, and classification methods. Analysis and visualization done through tSNE, Pandas, Numpy, ANNOY, and Plotly</li></ul></div> <div><div><i>Product Manager 2</i>-Microsoft Cloud + AI, Global Services05/2016-11/2017</div><ul style="list-style-type: none"><li>• Designed, built, and deployed Neural Machine Translation(NMT) into a software localization pipeline all Cloud and Enterprise Microsoft products. Using the NMT models, supplier contracts were renegotiated, saving 20% annualizing to \$3m/year over the next three years.</li><li>• Optimized and scaled the NLP system for customer feedback translation, filtering, and categorization, decreasing team effort by 80% using NLTK, Tensorflow, and sci-kit-learn.</li><li>• Drove end to end localization and internationalization for Azure and Visual Studio products, including automation, bug triage, budgeting, vendor, and release management for \$4m/year worth of translations.</li></ul></div> <div><div><i>Product Manager</i>-Microsoft, Azure RemoteAppSummer 2014 &amp; 08/2015-05/2015</div><ul style="list-style-type: none"><li>• Designed scalable multi-cloud network architecture for cloud-based Remote Desktop Service and drove large scale deployment to alpha customers and Microsoft.</li><li>• Created and scaled a predictive user demand modeling pipeline which, optimized VM usage by 20%.</li></ul></div> <div><div><i>NLP Researcher</i>-Basis Technology01/2014-05/2014</div><ul style="list-style-type: none"><li>• Grew the Rossette Entity Resolution pipeline to decrease document analysis time, increase accuracy, and increase throughput while expanding language support to Spanish, Russian, French, Farsi, and Chinese.</li><li>• Rearchitected the systems machine-learned architecture to state of the art neural methods, which improved multilingual accuracy to English parity.</li><li>• Formalized data creation, data ingestions, retrieval, formatting, and analysis pipeline using Bash, Java, and Python primarily. Optimized models were ANN, KNN, SVMs, LSTMs, and Decision Trees.</li></ul></div> <div><div><i>Software Engineer</i>-Cisco SystemsSummer 2013</div></div>

- Developed JavaScript tooling for collecting user feedback, bug tracking, and HTML5 based screenshotting used with alpha customers during initial product launches.
- Led a team of 12 other interns to develop a novel food ordering system for cafeterias that used containerized Node.js and MongoDB applications allowing efficient scaling.

*CoFounder/Software Engineer- Gapelia* 08/2013-05/2014

- Built an end to end digital multimodal publishing platform which allowed non-technical users to create and publish beautiful digital magazines in a matter of minutes.
- Created REST API in Java, front end in Express, and deployed in AWS with Apache Tomcat, Maven, and MongoDB.
- Won RPI 2014 Business Model Competition 1st place, 2014 Why not change the world grant, Harvard ILab cultural entrepreneurship challenge, and was member of Harvard's Innovation Lab 01/2014-05/2014.

*Software Engineer-GE Capital* Summer 2012

- Developed a reusable web templating system in Bootstrap, Javascript and PHP which allowed GE employees to create websites that matched corporate guidelines around colors, fonts, icons, etc.

*Support Engineer-KIO Networks* Summer 2011

- Debugged and triaged customers networking issues in a secure datacenter.

*Co-founder-Amerikano Experiences* 05/2009-06/2011

- Built up an event hosting business which would host parties for 16-22 year olds in Mexico City.
- Hosted 14+ events, build up a sales network of 50+ students and had \$150,000 revenue.

*Developer Platform Engineer-Microsoft* Summer 2010

- Built proof of concept solutions for customers in the telecom and banking industries using windows products such as Windows Phone 7 and Xbox Kinect

## Research

### Publications

- MS MARCO: A Human Generated Machine Reading COMprehension Dataset- Neurips 2016,2018
- Open Domain Web Keyphrase Extraction Beyond Language Modeling-EMNLP-IJCNLP 2019 Oral Presentation
- Experiments in Inferring Social Networks of Diffusion

### Patents

- Keyphrase Extraction Beyond Language Modeling - U.S. Appln. No. 16/460,853 - Filed July 2nd, 2019

### Datasets

- Optimal Freshness Crawl Under Politeness Constraints

- MSMARCO Question Answering and Natural Language Generation
- MSMARCO Passage Retrieval
- Open Domain Keyphrase Extraction
- Conversational Search with Privacy Minded Intent Representation

## **Academic Activity**

- NIST TREC 2019 Deep Learning-Track Coordinator
- NIST TREC 2020 Deep Learning-Track Coordinator
- ACM SIGIR/SIGKDD Africa Summer School on Machine Learning for Data Mining and Search 2019-Invited Lecturer for Deep Learning in Search
- AACM SIGIR/SIGKDD Africa Summer School on Machine Learning for Data Mining and Search 2020-Invited Lecturer for Deep Learning in Search
- LatinX in AI Research at Neurips-Reviewer