

OSHIN AGARWAL

oagarwal3@bloomberg.net | <https://oagarwal.github.io>

EDUCATION

University of Pennsylvania, Philadelphia, PA, USA
Ph.D. Computer and Information Science (Natural Language Processing) *Aug 2018 – Feb 2023*
Advisor: Ani Nenkova
Thesis Title: Practical Named Entity Recognition: The Role of Entity and its Context
M.S.E. Computer and Information Science. GPA: 3.84/4.0 *Aug 2016 – May 2018*
Birla Institute of Technology and Science - Pilani, Goa, India
B.E. (Hons), Computer Science Engineering, GPA: 9.55/10.0 *Aug 2010 – May 2014*

RESEARCH AND INDUSTRY EXPERIENCE (FULL-TIME)

Senior NLP Research Engineer *March 2023 – Present*
Bloomberg LP, New York City, NY, USA
Working on financial information extraction and summarization for news automation
Software Engineer *July 2014 – July 2016*
Paypal, Chennai, India
Worked on the platform for batch processing of credit transactions to facilitate money movement. Developed features for existing vendors, enabled new vendors, and performed internal migration to improve run time

RESEARCH AND INDUSTRY EXPERIENCE (INTERNSHIPS)

Research Scientist Intern *April 2022 – Nov 2022*
Adobe Research, College Park, MD, USA (remote)
Mentor: Ani Nenkova
Worked on incorporating external knowledge for named entity recognition, and the cross-domain evaluation of summarization models
Research Intern/Student Researcher *June 2021 – Nov 2021*
Google Research, Mountain View, CA, USA (remote)
Mentors: Gustavo Hernandez Abrego, Mandy Guo, Jianmo Ni, Yinfei Yang
Incorporating long-form text to leverage multiple retrieved passages in open domain question answering
Research Intern/Student Researcher *May 2020 – Sept 2020*
Google Research, Mountain View, CA, USA (remote)
Mentors: Heming Ge, Rami Al-Rfou, Siamak Shakeri
Knowledge Graph verbalization to integrate knowledge graphs in language model pre-training
Research Intern/Student Researcher *May 2019 – Dec 2019*
Google Research, New York City, NY, USA
Mentor: Dan Bikel
Entity linking via dual and cross-attention encoders
Product Intern *Jan 2014 - June 2014*
Adobe Systems, Noida, India
Data analysis and machine learning to predict subscriber churn for five million Creative Cloud customers
Summer Intern *May 2013 - July 2013*
Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India
Fault detection using chaotic maps on structured adaptive mesh refinement solvers built in C++ when run on multi-core processors
Research Trainee *May 2012 - July 2012*
Bhabha Atomic Research Centre, Mumbai, India
Built real-time video interfaces for stereoscopic video players using GStreamer and QT framework in C++

TEACHING EXPERIENCE

- Teaching Assistant, Senior Design (CIS 400), University of Pennsylvania *Fall 2019*
- Teaching Assistant, Applied Machine Learning (CIS 519), University of Pennsylvania *Fall 2018*
- Teaching Assistant, Software Systems (CIS 505), University of Pennsylvania *Fall 2017 & Spring 2018*

PUBLICATIONS

Journal and Conference Papers

- Named Entity Recognition in a Very Homogeneous Domain
Oshin Agarwal and Ani Nenkova
Findings of EACL 2023, [Link](#)
- Temporal Effects on Pre-trained Models for Language Processing Tasks
Oshin Agarwal and Ani Nenkova
TACL 2022, [Link](#)
- Knowledge Graph Based Synthetic Corpus Generation for Knowledge-Enhanced Language Model Pre-training
Oshin Agarwal, Heming Ge, Siamak Shakeri and Rami Al-Rfou
NAACL 2021, [Link](#)
- The Utility and Interplay of Gazetteers and Entity Segmentation for Named Entity Recognition in English
Oshin Agarwal and Ani Nenkova
Findings of ACL 2021, [Link](#)
- Interpretability Analysis for NER to Understand System Predictions and How They Can Improve
Oshin Agarwal, Yinfei Yang, Byron C. Wallace and Ani Nenkova
CL 2021, [Link](#)
- From Toxicity in Online Comments to Incivility in American News: Proceed with Caution
Anushree Hede, Oshin Agarwal, Linda Lu, Diana C. Mutz and Ani Nenkova
EACL 2021, [Link](#)
- Predicting Annotation Difficulty to Improve Task Routing and Model Performance for Biomedical IE
Yinfei Yang, Oshin Agarwal, Chris Tar, Byron C. Wallace and Ani Nenkova
NAACL 2019, [Link](#)

Workshop Papers

- Towards Robust NER via Temporal Domain Adaptation and Entity Context Understanding
Oshin Agarwal
AAAI Doctoral Consortium 2022, [Link](#)
- Machine Translation Aided Bilingual Data-to-Text Generation and Semantic Parsing
Oshin Agarwal, Mihir Kale, Heming Ge, Siamak Shakeri and Rami Al-Rfou
WebNLG+ 2020, [Link](#)
- Word Embeddings (Also) Encode Human Personality Stereotype
Oshin Agarwal, Funda Durupinar, Norman I. Badler and Ani Nenkova
*SEM 2019, [Link](#)
- Evaluation of Named Entity Coreference
Oshin Agarwal*, Sanjay Subramanian*, Ani Nenkova and Dan Roth
CRAC 2019, [Link](#)
- Browsing Health: Information Extraction to Support New Interfaces for Accessing Medical Evidence
Soham Parikh, Elizabeth Conrad, Oshin Agarwal, Iain Marshall, Byron Wallace and Ani Nenkova
ESSP 2019, [Link](#)

Preprints and Technical Reports

- Entity-Switched Datasets: An Approach to Auditing the In-Domain Robustness of NER Models
Oshin Agarwal, Yinfei Yang, Byron C. Wallace and Ani Nenkova, [Link](#)
- Entity Linking via Dual and Cross-Attention Encoders
Oshin Agarwal and Dan Bikel, [Link](#)

AWARDS

- Haidas & Chryssikou Fellow for excellent progress in the PhD program, University of Pennsylvania *2020*
- Third place poster award at the NDS symposium of New York Academy of Sciences for work on In-domain Robustness of Named Entity Recognition *2019*
- Merit scholarship for being in top 1% of the class, BITS-Pilani Goa Campus *2010 – 2014*

LEADERSHIP AND SERVICE

- Reviewer - ACL (2020), AACL (2020), EMNLP (2019), NAACL (2021), ACL SRW (2020, 2021), EACL SRW (2021)
- Co-chair & Treasurer, Computer and Information Science Doctoral Student Association, University of Pennsylvania *Fall 2021 – Spring 2022*
- Board member, Computer and Information Science Doctoral Student Association, University of Pennsylvania *Fall 2018 – Spring 2022*