

Sunipa Dev

Mobile: +1 (214) 566-4559

Email: sunipadev@gmail.com

Website : sunipa.github.io

EDUCATION AND TRAINING

- **University of California, Los Angeles** Jan 2021 – Jan 2022
Computing Innovation Postdoctoral Fellow; Advisor : Kai-Wei Chang Los Angeles, CA
- **University of Utah** Aug 2016 – Dec 2020
PhD Computer Science; Advisor : Jeff M Phillips Salt Lake City, UT
 - **Coursework:** Data Mining, Machine Learning, Advanced Algorithms, Data Visualization, NLP, Databases
- **Indian Institute of Science Education and Research, Kolkata** Jul 2011 – Jun 2016
Integrated Bachelors and Masters in Mathematics and Statistics Kolkata, India
 - **Coursework:** Differential Geometry, Topology, Statistical Inference, Probability, Stochastic Processes

RESEARCH AND WORK EXPERIENCE

- **Google** Jan 2022 – Present
Research Scientist, Responsible AI and Human Centered Technologies, Google Research San Francisco, CA
 - Assess Large language models for fairness, robustness, and transparency. Build measures to ascertain fair deployment. Lead research on experienced harms based model evaluations. Consult product and associated research teams on Responsible AI.
- **University of California, Los Angeles** Jan 2021 – Jan 2022
Computing Innovation Postdoctoral Researcher; Advisor: Kai-Wei Chang Los Angeles, CA
 - Extend NLP fundamental tasks and applications to perform well beyond just binary gender. Understand and mitigate biases therein. Develop better evaluation paradigms for models towards socially aware NLP. Extend NLP tasks meaningfully to the healthcare domain and ensure fair mental health assessments across different demographic groups.
- **University of Utah** Aug 2016 – Dec 2020
Research Assistant; Advisor: Jeff M Phillips Salt Lake City, UT
 - Understand and enhance interpretability of high-dimensional representations. Analyze, evaluate, measure and mitigate social biases in language representations and their downstream applications in interpretable and cost-effective ways that do not break down embedding structures. Built meaningful evaluation paradigms for the same.
- **Microsoft** June 2020 – August 2020
Data Scientist Intern; Advisor: Yi Zhang Redmond, WA
 - Built adaptive user representations which attend to a user's current needs for providing better Ad engagement.
- **Visa Research** May 2019 – Aug 2019
PhD Research Intern; Advisor: Wei Zhang Palo Alto, CA
 - Built user and merchant representations in interpretable ways so as build better recommender systems to match users to merchants.
- **IISER Kolkata** June 2014 – June 2016
Senior Thesis Researcher; Advisor: Anirban Banerjee Kolkata, India
- **Okinawa Institute of Science and technology** May 2015 – Aug 2015
Research Intern; Advisor: Robert Sinclair Okinawa, Japan
- **Institute of Mathematical Sciences** May 2014 – Aug 2014
Research Assistant; Advisor: Rahul Siddhantan Chennai, India

AWARDS

- DAAD AI-net Postdoc Fellow 2021
- **Computing Innovation Fellowship** Award Recipient 2020 (by CRA, CCC and NSF)
- ACL Outstanding Reviewer 2021
- Grace Hopper Scholar 2019
- INSPIRE Fellowship by DST and Govt. of India for top 1% students in Science. (2011-2016)

TEACHING

- Guest Lecture at course ‘Automatic Speech Recognition’, UCLA; Spring, 2021
- Conducted selected tutorial at **AAAI 2021** and **KDD 2021**: A Visual Tour of Bias Mitigation Techniques in Word Embeddings.
- Teaching Mentee for Data Mining (Spring 2018, 103 students): conducted lectures, tutorials, project evaluations and office hours
- Teaching Mentee for Foundations of Data Analysis (Fall 2017, 43 students): conducted lectures, tutorials, project evaluations and office hours.
- Teaching Assistant for Mathematics II (Spring 2016, ~200 students): conducted tutorials, office hours and graded assignments and exam papers.

SELECTED CONFERENCE PRESENTATIONS AND INVITED TALKS

- **University of Irvine** Feb 2022
Towards Socially-Aware, Inclusive, and Fair Language Representations Irvine, California
- **University of Bocconi** Oct 2021
Towards Interpretable and Fair Language Representations Bocconi, Italy
- **LinkedIn** Aug 2021
Towards Interpretable and Fair Language Representations Bangalore, India
- **EMNLP 2021** Nov 2021
Harms of gender exclusivity and challenges in non-binary representation in language technologies Dominican Republic
- **AAAI 2020** Feb 2020
On Measuring and Mitigating Biased Inferences of Word Embeddings New York City, USA
- **ICDM 2019** Nov 2019
Closed Form Word Embedding Alignment Beijing, China
- **AISTats 2019** April 2019
Attenuating Biases in Word Embeddings Okinawa, Japan

PUBLICATIONS

1. **S. Dev**, E. Sheng, J. Zhao, A. Amstutz, J. Sun, Y. Hou, M. Sanseverino, J. Kim, A. Nishi, N. Peng, and KW. Chang; “On measures of Biases and Harms in NLP”; AACL 2022, Taiwan
2. **S. Dev**, M. Monajatipoor, A. Ovalle, A. Subramonian, J.M. Phillips and KW. Chang; “Harms of Gender Exclusivity and Challenges in Non-Binary Representation in Language Technologies”; EMNLP 2021
3. **S. Dev**, T. Li, J.M. Phillips and V. Srikumar; “OSCAR: Orthogonal Subspace Correction and Rectification of Biases in Word Embeddings”; EMNLP 2021
4. **S. Dev**, T. Li, J.M. Phillips and V. Srikumar; “On Measuring and Mitigating Biased Inferences of Word Embeddings”; AAAI 2020, New York City, USA; (oral presentation)

5. **S. Dev**, S. Hassan, J.M. Phillips; “Closed Form Word Embedding Alignment”; ICDM 2019, Beijing, China; (oral presentation) **In Special Issue of KAIS for Best Papers from ICDM**
6. **S. Dev**, J.M. Phillips; “Attenuating Bias in Word Vectors”; AISTATS 2019, Okinawa, Japan; (oral presentation)
7. V. Malik, **S. Dev**, A. Nishi, N. Peng, and KW. Chang; “Socially Aware Bias Measurements for Hindi Language Representations”; NAACL 2022, Seattle, USA
8. A. Chowdhery, ..., **S. Dev**, ..., N. Fiedel; “Palm: Scaling language modeling with pathways”;
9. S.Bhatt, **S. Dev**, P. Talukdar, S. Dave, V. Prabhakaran; “Re-contextualizing Fairness in NLP: The Case of India”; AACL 2022, Taiwan
10. D.Wu, W. Ahmad, **S. Dev**, KW. Chang; “Representation Learning for Resource-Constrained Keyphrase Generation”; EMNLP Findings 2022, Abu Dhabi
11. A. Rathore, **S. Dev**, JM. Phillips, V. Srikumar, Y. Zheng, CM. Yeh, J. Wang, W. Zhang, B. Wang; “VERB: Visualizing and Interpreting Bias Mitigation Techniques for Word Representations”; (accepted to appear in TIS)
12. A. Ovalle, **S. Dev**, J. Zhao, M. Sarrafzadeh, KW. Chang; “Auditing Algorithmic Fairness in Machine Learning for Health with Severity-Based LOGAN”; AAAI 2023 Workshop ; Washington D.C., USA
13. **S. Dev**; “The geometry of distributed representations for better alignment, attenuated bias, and improved interpretability”;
14. **S. Dev**; “Brain Data Analysis for Visual Processing and Categorization in the Human Cortex”;

RESEARCH ADVISING

I am extremely fortunate to have had the opportunity to co-advise, and collaborate with following students during my postdoc and as a Research Scientist at Google through internships awarded to the students.

- | | |
|--|--|
| • Akshita Jha
<i>3rd year PhD student</i> | Aug 2022 – Jan 2023
<i>Virginia Tech, Google Student Researcher</i> |
| • Nikil Selvam
<i>Final year undergraduate Student</i> | July 2021 – Jan 2022
<i>UCLA</i> |
| • Masoud Monajatipor
<i>2nd year PhD Student</i> | Jan 2021 – Jan 2022
<i>UCLA</i> |
| • Anaelia Ovalle
<i>2nd year PhD Student</i> | Feb 2021 – Jan 2022
<i>UCLA</i> |
| • Arjun Subramonian
<i>Incoming PhD Student</i> | Feb 2021 – Nov 2022
<i>UCLA</i> |
| • Vijit Malik
<i>4th year Undergraduate Student</i> | April 2021 – Jan 2022
<i>IIT Kanpur, India</i> |

SERVICE

- Affinity Workshop Chair at Neurips 2022
- Chair Widening NLP (2021-2023)
- Workflow Co-chair AAAI 2022
- Action Editor ACL-ARR
- Organizer of workshop at KDD 2021 on Responsible AI
- Reviewer for conferences like AAAI, EACL and ACL + workshops at ACL, ICML, EACL
- Reviewer for IEEE journals

OUTREACH

- Co-founded Women at School of Computing at University of Utah.
- Mentor at LatinX in AI
- Graduate Student Advisory Committee Member (6 students a year) from 2017-2018. Organized events for new and existing students to form sense of community, increase interactions, and aid requirements among other duties. Also contributed to Retention, Promotion Tenure (RPT) procedures of faculty members.