

AWARDS & RECOGNITION

| | |
|---|-----------|
| Impact Recognition Award - ACM CSCW | Oct 2021 |
| Best Poster Award - UIUC Student Poster Session | Mar 2020 |
| Best student paper award - ASIST SIGMET Workshop | Nov 2018 |
| Graduate Teacher Certificate | May 2018 |
| University of Illinois GIS Day Runner-up (Research Quality) | Nov 2017 |
| Kishore Vaigyanik Protsahan Yojana Scholar | 2007-2012 |
| 3rd rank in Regional Mathematics Olympiad, Uttar Pradesh, India | Dec 2006 |

TEACHING

| | |
|--|--------------------|
| Tutorial presenter, Multiple venues | Sep 2019 - Current |
| <i>Tutorial on hands on advanced machine learning for information extraction from tweets tasks, data, and open source tools. Details at: https://socialmediaie.github.io/tutorials/</i> | |
| Co-instructor - Network Analysis | Spring 2018 |
| Teaching Assistant - Network Analysis | Summer 2017 |
| Teaching Assistant - Foundations of Information Processing | Spring 2017 |
| Co-instructor - Data Mining Applications | Fall 2016 |
| Listed in Teachers Ranked as Excellent! | |

ALL PUBLICATIONS

- [1] R. Eskander et al. “Towards Improved Distantly Supervised Multilingual Named-Entity Recognition for Tweets”. In: Proceedings of the 2nd Multilingual Representation Learning (MRL) Workshop. Association for Computational Linguistics, 2022.
- [2] J. A. Fries et al. “BigBIO: A Framework for Data-Centric Biomedical Natural Language Processing”. In: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks 2 (NeurIPS Datasets and Benchmarks 2022). arXiv, 2022.
- [3] L. Hebert et al. “Robust Candidate Generation for Entity Linking on Short Social Media Texts”. In: Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022). Association for Computational Linguistics, 2022, pp. 83–89.
- [4] J. Li et al. “NTULM: Enriching Social Media Text Representations with Non-Textual Units”. In: Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022). Association for Computational Linguistics, 2022, pp. 69–82.
- [5] S. Mishra et al. “TweetNERD – End to End Entity Linking Benchmark for Tweets”. In: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks 2 (NeurIPS Datasets and Benchmarks 2022). arXiv, 2022.
- [6] V. Kulkarni, S. Mishra and A. Haghighi. “{LMSOC}: An Approach for Socially Sensitive Pretraining”. In: Findings of the Association for Computational Linguistics: EMNLP 2021. Association for Computational Linguistics, 2021, pp. 2967–2975.
- [7] S. Mishra and A. Haghighi. “Improved Multilingual Language Model Pretraining for Social Media Text via Translation Pair Prediction”. In: Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021). Association for Computational Linguistics, 2021, pp. 381–388.
- [8] S. Mishra, S. Prasad and S. Mishra. “Exploring Multi-Task Multi-Lingual Learning of Transformer Models for Hate Speech and Offensive Speech Identification in Social Media”. In: SN Computer Science 2.2, 2021, p. 72.
- [9] K. Yee, U. Tantipongpipat and S. Mishra (equal). “Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency”. In: Proceedings of the ACM on Human-Computer Interaction 5.CSCW2, 2021, pp. 1–24.
- [10] K. Han et al. “WikiCSSH: Extracting Computer Science Subject Headings from Wikipedia”. In: Workshop on Scientific Knowledge Graphs (SKG 2020). 2020.
- [11] S. Mishra. “Information Extraction from Digital Social Trace Data with Applications to Social Media and Scholarly Communication Data”. In: ACM SIGIR Forum 54.1, 2020.
- [12] S. Mishra. “Non-neural Structured Prediction for Event Detection from News in Indian Languages”. In: Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation. CEUR Workshop Proceedings, CEUR-WS.org, 2020.
- [13] S. Mishra and D. Collier. “A Framework for Generating Annotated Social Media Corpora with Demographics, Stance, Civility, and Topicality”. In: SSRN Electronic Journal, 2020.
- [14] S. Mishra and S. Mishra. “Scubed at 3C task A - A simple baseline for citation context purpose classification”. In: Proceedings of the 8th International Workshop on Mining Scientific Publications. Association for Computational Linguistics, 2020, pp. 59–64.

- [15] S. Mishra and S. Mishra. “Scubed at 3C task B - A simple baseline for citation context influence classification”. In: Proceedings of the 8th International Workshop on Mining Scientific Publications. Association for Computational Linguistics, 2020, pp. 65–70.
- [16] S. Mishra, S. Prasad and S. Mishra. “Multilingual Joint Fine-tuning of Transformer models for identifying Trolling, Aggression and Cyberbullying at TRAC 2020”. In: Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying. European Language Resources Association (ELRA), 2020, pp. 120–125.
- [17] N. N. Parulian et al. “Effectiveness of the Execution and Prevention of Metric-Based Adversarial Attacks on Social Network Data †”. In: Information 11.6, 2020, p. 306.
- [18] M. V. Avram et al. “Adversarial perturbations to manipulate the perception of power and influence in networks”. In: 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2019, pp. 986–993.
- [19] D. Collier et al. “Who is Most Likely to Oppose Federal Tuition-Free College Policies? Investigating Variable Interactions of Sentiments to America’s College Promise”. In: SSRN Electronic Journal, 2019.
- [20] D. A. Collier et al. “Americans ‘support’ the idea of tuition-free college: an exploration of sentiment and political identity signals otherwise”. In: Journal of Further and Higher Education 43.3, 2019, pp. 347–362.
- [21] S. Mishra. “Multi-dataset-multi-task Neural Sequence Tagging for Information Extraction from Tweets”. In: Proceedings of the 30th ACM Conference on Hypertext and Social Media - HT ’19. ACM Press, 2019, pp. 283–284.
- [22] S. Mishra and J. Diesner. “Capturing Signals of Enthusiasm and Support Towards Social Issues from Twitter”. In: Proceedings of the 5th International Workshop on Social Media World Sensors - SIdEWays’19. ACM Press, 2019, pp. 19–24.
- [23] S. Mishra and S. Mishra. “3Idiots at HASOC 2019: Fine-tuning Transformer Neural Networks for Hate Speech Identification in Indo-European Languages”. In: Proceedings of the 11th annual meeting of the Forum for Information Retrieval Evaluation. 2019, pp. 208–213.
- [24] S. Mishra and J. Diesner. “Detecting the Correlation between Sentiment and User-level as well as Text-Level Meta-data from Benchmark Corpora”. In: Proceedings of the 29th on Hypertext and Social Media - HT ’18. ACM Press, 2018, pp. 2–10.
- [25] S. Mishra et al. “Expertise as an aspect of author contributions”. In: Metrics 2018: Workshop on Informetric and Scientometric Research (SIG/MET). 2018.
- [26] S. Mishra et al. “Self-citation is the hallmark of productive authors, of any gender”. In: PLoS ONE 13.9, 2018, e0195773.
- [27] A. Addawood et al. “Developing an Information Source Lexicon”. In: Prioritising Online Content workshop co-located at NIPS. 2017.
- [28] S. Mishra. “SCTG: Social Communications Temporal Graph – A novel approach to visualize temporal communication graphs from social data”. In: UIUC Data Science Day. 2017.
- [29] S. Mishra and J. Diesner. “Semi-supervised Named Entity Recognition in noisy-text”. In: Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT). The COLING 2016 Organizing Committee, 2016, pp. 203–212.
- [30] S. Mishra and V. I. Torvik. “Quantifying Conceptual Novelty in the Biomedical Literature.” In: D-Lib magazine : the magazine of the Digital Library Forum 22.9-10, 2016.
- [31] S. Mishra et al. “Sentiment Analysis with Incremental Human-in-the-Loop Learning and Lexical Resource Customization”. In: Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT ’15. ACM Press, 2015, pp. 323–325.
- [32] S. Mishra et al. “Enthusiasm and support: alternative sentiment classification for social movements on social media”. In: Proceedings of the 2014 ACM conference on Web science - WebSci ’14. ACM Press, 2014, pp. 261–262.