

# Sunny Rai

299, Towne Hall, 3330 Walnut St., Philadelphia, PA 19104

✉ [sunnyrai@seas.upenn.com](mailto:sunnyrai@seas.upenn.com)

## Education

---

|   |               |
|---|---------------|
| <b>Ph.D. in Computer Engineering</b><br><i>Faculty of Technology, University of Delhi, India</i>                                    | <b>Feb'21</b> |
| <b>Master of Technology in Information Security (with Honors)</b><br><i>Guru Gobind Singh Indraprastha University, Delhi, India</i> | <b>Jul'14</b> |
| <b>Bachelor of Technology in Information Technology</b><br><i>Guru Gobind Singh Indraprastha University, Delhi, India</i>           | <b>Jul'12</b> |

## Research Experience

---

|   |                          |
|---|--------------------------|
| <b>The World Bank, USA.</b><br><i>Consultant - Gender Norms in Entertainment Media</i>  | <b>July'24 - present</b> |
| <b>University of Pennsylvania, USA.</b><br><i>Postdoctoral Researcher, Department of Computer and Information Science</i><br>o Advisors: Dr. Sharath Chandra Guntuku and Dr. Lyle Ungar | <b>Oct'22 - present</b>  |
| <b>École Centrale School of Engineering, Mahindra University, India.</b><br><i>Assistant Professor, Department of Computer Science and Engineering</i>                                  | <b>Dec'20 - Sept'22</b>  |
| <b>University of Delhi, India.</b><br><i>Graduate Teaching Assistant</i>  | <b>Dec'14- Jul'18</b>    |

## Publications

---

Peer-Reviewed Conference/Workshop.....

- Y'24 Havaladar, S., Giorgi, S., **Rai, S.**, Talhelm, T., Guntuku, S.C., & Ungar, L. (2024). Building Knowledge-Guided Lexica to Model Cultural Variation. NAACL'24.
- Y'23 Cho, Y.M., **Rai, S.**, Ungar, L., Sedoc J., & Guntuku, S.C. (2023). An Integrative Survey on Mental Health Conversational Agents to Bridge Computer Science and Medical Perspectives. EMNLP'23, Singapore.
- Y'23 Havaladar, S., **Rai, S.**, Singhal, B., Liu, L., Guntuku, S. C., & Ungar, L. (2023, July). Multilingual Language Models are not Multicultural: A Case Study in Emotion. WASSA, ACL'23. **Best Paper Award**
- Y'23 Joseph, R., Liu, T., Ng, A. B., See, S., & **Rai, S.** (2023, July). NewsMet: A 'Do It All'dataset of contemporary Metaphors in News headlines. In Findings of the Association for Computational Linguistics: ACL 2023 (pp. 10090-10104).
- Y'23 Aggarwal, A., **Rai, S.**, Giorgi, S., Havaladar, S., Sherman, G., Mittal, J., & Guntuku, S. C. (2023, April). A Cross-Modal Study of Pain Across Communities in the United States. In Companion

Proceedings of the ACM Web Conference 2023 (pp. 1050-1058).

- Y'23 Bogireddy, N. R., Suresh, S., & **Rai, S.** (2023, April). I'm out of breath from laughing! I think? A dataset of COVID-19 Humor and its toxic variants. In Companion Proceedings of the ACM Web Conference 2023 (pp. 1004-1013).
- Y'22 **Rai, S.**, Joseph, R., Thakur, P. S., & Khaliq, M. A. (2022). Identifying Human Needs through Social Media: A study on Indian cities during COVID-19. In Proceedings of the Tenth International Workshop on Natural Language Processing for Social Media (pp. 65-74).
- Y'21 Khaliq, M., Joseph, R., & **Rai, S.** (2021, December). #covid is war and #vaccine is weapon? COVID-19 metaphors in India. In Proceedings of the 18th International Conference on Natural Language Processing (ICON) (pp. 431-438).
- Y'17 **Rai, S.**, & Chakraverty, S. (2017). Metaphor detection using fuzzy rough sets. In Rough Sets: International Joint Conference of Rough Sets, IJCRS 2017, Olsztyn, Poland, July 3–7, 2017, Proceedings, Part I (pp. 271-279). Springer International Publishing.
- Y'16 **Rai, S.**, Chakraverty, S., & Tayal, D. K. (2016, June). Supervised metaphor detection using conditional random fields. In Proceedings of the Fourth Workshop on Metaphor in NLP (pp. 18-27).

#### Journal.....

- Y'24 **Rai, S.**, Stade, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C. (2024). Key language markers of depression on social media depend on race. PNAS Brief Reports.
- Y'24 **Rai, S.**, Kornides, M., Morgan, J., Kumar, A., Cappella, J., and Guntuku, S.C. (2024). Detecting and Monitoring Concerns Against HPV Vaccination on Social Media Using Large Language Models. Scientific Reports.
- Y'20 **Rai, S.**, & Chakraverty, S. (2020). A survey on computational metaphor processing. ACM Computing Surveys (CSUR), 53(2), 1-37.
- Y'19 **Rai, S.**, Chakraverty, S., Tayal, D. K., Sharma, D., & Garg, A. (2019). Understanding metaphors using emotions. New Generation Computing, 37, 5-27.
- Y'18 **Rai, S.**, Chakraverty, S., Tayal, D. K., & Kukreti, Y. (2018). A study on impact of context on metaphor detection. The Computer Journal, 61(11), 1667-1682.
- Y'15 Jain, A., Tayal, D. K., & **Rai, S.** (2015). Shrinking digital gap through automatic generation of WordNet for Indian languages. AI & SOCIETY, 30, 215-222.

#### Preprints.....

- Y'24 **Rai, S.\***, Shelat, K.\*, Jain, D. R., Cho, Y. M., Redkar, M., Sawant, S., & Guntuku, S. C. (2024). Studying Differential Mental Health Expressions in India. arXiv preprint arXiv:2402.11477.
- Y'24 **Rai, S.**, Zaveri, K. J., Havaladar, S., Nema, S., Ungar, L., & Guntuku, S. C. (2024). Social Norms in Cinema: A Cross-Cultural Analysis of Shame, Pride and Prejudice. arXiv preprint arXiv:2402.11333.
- Y'23 Baheti, B., **Rai, S.**, Innani, S., Mehdiratta, G., Nasrallah, M. P., & Bakas, S. (2023). Detecting Histologic Glioblastoma Regions of Prognostic Relevance. arXiv preprint arXiv:2302.00669.

\*co-first authors.

#### Non Archival.....

- Y'24 **Rai, S.**, Zaveri, K. J., Havaladar, S., Nema, S., Ungar, L., & Guntuku, S. C. (2024). A Cross-Cultural Analysis of Social Norms in Bollywood and Hollywood Movies. Oral Talk at IC2S2'24, Philadelphia.

- Y'24 **Rai, S.\***, Shelat, K.\*, Jain, D. R., Cho, Y. M., Redkar, M., Sawant, S., & Guntuku, S. C. (2024). Studying Differential Mental Health Expressions in India. Poster at IC2S2'24, Philadelphia.
- Y'24
- Y'24 Jamalova, R., **Rai, S.\***, & Guntuku, S. C. (2024). Characterizing the Geographical Landscape of COVID-19 Misinformation from X. Poster at IC2S2'24, Philadelphia
- Y'24 Baheti, B., **Rai, S.**, Innani, S., Mehdiratta, G., Guntuku, S. C., Nasrallah, M. P., & Bakas, S. (2024). "Detecting Histologic & Clinical Glioblastoma Patterns of Prognostic Relevance", Oral presentation at the 20th edition of the European Congress on Digital Pathology.
- Y'23 Baheti, B., **Rai, S.**, Innani, S., Mehdiratta, G., Guntuku, S. C., Nasrallah, M. P., & Bakas, S. (2023). EPCO-15. DETECTING HISTOLOGIC & CLINICAL GLIOBLASTOMA PATTERNS OF PROGNOSTIC RELEVANCE. *Neuro-Oncology*, 25(Supplement\_5), v126-v126.
- Y'23 **Rai, S.**, Francisco, A., Giorgi, S., Curtis, B., Ungar, L., and Guntuku, S.C. (2023). Studying Racial Heterogeneity in Language Markers of Depression. Early Career Plenary Talk at IC2S2'23, Denmark. Acceptance Rate: 1.74%
- Y'23 Aggarwal, A., **Rai, S.**, Giorgi, S., Havaldar, S., Sherman, G., Mittal, J., and Guntuku, S.C. (2023). A Cross-Modal Study of Pain Across Communities in the United States. [Workshop on Trustworthy AI](#), Jan 5-6, 2023 - Microsoft Research, Bengaluru, India. POSTER.
- Y'22 Joseph, R., **Rai, S.**, Thakur, P.S., & Khaliq, M.A. (2022). Pandemic, People and Pattern of needs : Emerging trends in Indian society during COVID-19. In *8<sup>th</sup> IC2S2 2022*.
- Y'19 Gollapudi, U. M , Arulkumaran, R., Perla, M., Jonnalagadda, S.H., Das, A., and **Rai, S.** "*Understanding User Vulnerability Towards Radicalization On Twitter*". IC2S2 2019.

## Invited Talks/Panels

---

1. "**Key Language Markers of Depression on Social Media Depend on Race**", Society for Personality and Social Psychology (SPSP) Annual Convention 2025 (pre-conference), Denver.
2. "**Shame, Social Norms and Cultural Alignment of LLMs**", The 1st Workshop on Reclaiming the Narrative: Digital Recovery, AI, and Mitigating Harm in Social Media, ICWSM, Buffalo, NY, June 3, 2024
3. Panelist for Penn Data Scientist Meetup, Penn Data-Driven Discovery Initiative. May 29, 2024
4. Invited Research Talk on "**Using LLMs for analyzing human behaviors: A case study of social norms** ", Center for Social Norms and Behavioral Dynamics, University of Pennsylvania. Apr. 19, 2024.
5. **Cross-Cultural Study of Social Norms in Hollywood and Bollywood**, Computational Linguistics Lunch'24 (CLunch) at the University of Pennsylvania.
6. **Multimodal explainable AI for prognostic stratification of glioblastoma patients**, Penn ASSET/IBI Symposium on Trustworthy AI for Health Care. Oct 6, 2023
7. **Language Models/ NLP Chatbots**, IBI-GCB Retreat at the University of Pennsylvania. Sept, 2023.
8. **Racial Heterogeneity in Language Markers of Depression** , Computational Linguistics Lunch'23 (CLunch) at the University of Pennsylvania.
9. **Multilingual Zero-Shot Learning**, ATAL-FDP, a Government of India Initiative. Nov'21

## Awards & Scholarships

---

1. 1<sup>st</sup> prize at Rachleff Scholars Program at the University of Pennsylvania for “Characterizing the Geographical Landscape of COVID-19 Misinformation from Twitter”. Role: co-advisor
2. Best Paper Award at WASSA, ACL’2023.
3. Abstract selected for Early Career Research Plenary Talk at IC2S2, 2023, Denmark. (16 Plenaries were selected out of 918 submissions)
4. 1<sup>st</sup> at Yale Undergraduate Research Symposium for Project “Mental Health Expressions in India”. Role: Co-advisor
5. Awarded Grace Hopper Celebration’22 Faculty Scholarship
6. UGC-NET (2018)- Ranked in Top 6% in National level Exam for Assistant Professorship
7. Travel Grant (USD 1500) to attend NAACL 2016, Netaji Subhas Institute of Technology, University of Delhi.
8. Awarded Graduate Scholarship (2012, 2014) by Government of India
9. Director’s Award for Academic Excellence, 2011

## Volunteering/Service

---

### Chair.....

1. Session chair for “Computational Social Science and Cultural Analytics” track, NAACL 2024.
2. Publication chair, Session chair for “Computational Sequence Modeling”, ICETCI’22

### Reviewer.....

1. Conferences: COLM’24, ACL’[23,24], ICWSM’[23-24], EACL’23, AACL’[21-22], EMNLP’[21-24]
2. Workshops: NAACL-SRW’22, ACL-SRW’24
3. Non-archival venues: IC2S2’24
4. Journals
  - (a) NLP: ACM Transactions on Asian and Low Resource Language Information Processing (TALLIP), Natural Language Engineering, IEEE Transactions on Affective Computing
  - (b) Health Sciences: Frontiers in Psychology, Journal of Language and Social Psychology, Psychiatry Research, ACM Transactions on Computing for Healthcare, Communications Psychology, Nature, Scientific Reports, Nature, Human Behavior and Emerging Technologies, Hindawi,
  - (c) Others: IEEE Access, Neurocomputing, Computer Applications in Engineering Education (Wiley), Neurocomputing, Expert Systems, Computer Application in Engineering Education (Wiley)

### Organizer.....

1. Co-Organizer, Summer Institute for Computational Social Science (SICSS), University of Pennsylvania, 2023

## Press/Blog Mentions

---

- o PNAS Paper entitled “Key markers of depression on social media depend on race” covered by [45 News outlets](#) including Thomson-Reuters, Science News, etc.
- o My Research on cross-cultural social norms was covered and endorsed in a [blog](#) in “The Great

Gender Divergence”.

## **Professional Affiliations**

---

1. Associate Fellow, Leonard Davis Institute of Health Economics.
2. Member, Association of Computing Machinery
3. Member, Association of Computational Linguistics.
4. Member, The Institution of Engineers (India) ([IEI](#))