

# Suma P. Bhat

Department of Computer Science  
Princeton University  
35 Olden Street,  
Princeton, NJ 08540

Phone: (609) 258-1798  
Email: [sumabhat@princeton.edu](mailto:sumabhat@princeton.edu)

## Education

Ph.D. Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, 2010.

M.A. South and Southeast Asian Studies, University of California, Berkeley, 2000.

M.E. Electrical Engineering, Indian Institute of Science, India, 1996.

B.S. Statistics, Mangalore University, India, 1992.

## Academic Positions

Research Scholar, Department of Computer Science, Princeton University, August 2022-present.

Assistant Professor (tenure-track, on leave), Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, January 2020-present.

Research Assistant Professor, Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, May 2015–December 2019.

Faculty Affiliate, Department of Computer Science, University of Illinois, Urbana-Champaign, November 2020-present.

Faculty Affiliate, Department of Educational Psychology, University of Illinois, Urbana-Champaign, January 2020-present.

Faculty Affiliate, National Center for Supercomputing Applications, University of Illinois, Urbana-Champaign, September 2016-present.

Post-doctoral Fellow, Beckman Institute, University of Illinois, 2010.

## Grants

### *Current Research*

**Lead PI:** Building Idiomaticity into Natural Language Processing, National Science Foundation (EAGER), 2022-2023, \$150,000

**Co-PI:** Development of a Chatbot for Delivering Long-Term Motivational Interviewing for Improving Exercise Adherence in Hemodialysis Patients, The Jump Applied Research for Community Health through Engineering and Simulation program, 2022-2023, \$75,000. Lead PI: Jessie Chin.

**Co-PI:** Automated Assessment of Written Chart Notes: Generating Reliable, Timely, and Useful Feedback, 2021-2023, Edward J. Stemmler, MD Medical Education Research Fund Of the National Board of Medical Examiners (NBME), \$149,497.61. Lead PI: Dr. William Bond, Jump Simulation, Peoria.

### *Completed Research*

**Lead PI:** Using Study Partners to Broaden Participation, Institute for Inclusion, Diversity, Equity, and Access, Grainger College of Engineering (UIUC), 2020-2022, \$13,000.

**Lead PI:** Building a Motivational-Interviewing Conversational Agent (MintBot) for Promoting COVID-19 Vaccination among People with Multiple Sclerosis, The Jump Applied Research for Community Health through Engineering and Simulation program, 2021-2022, \$74,992.

**Co-PI:** COVID-19 Risk Mitigation: Interactive Automatic Counselor, University of Illinois at Urbana-Champaign, 2020-2022, \$5,000. PI: Dan Morrow.

**Lead PI:** Improving Feedback and Efficiency: Automated Grading of Post Simulation Written Chart Notes, The Jump Applied Research for Community Health through Engineering and Simulation program, 2020-2021, \$75,000.

**Co-PI:** Underrepresented Student Learning in Online Introductory STEM College Courses, Institute of Education Sciences, 2018-2021, \$1,399,194. PI: Michelle Perry.

**Co-PI:** Using Conversational Agents to Support Older Adult Learning for Health, Technology Innovation in Educational Research and Design Program, 2020-2021, \$13,602. PI: Dan Morrow.

**Lead PI:** Collaborative Research; BystanderBots: Automated Bystander Intervention for Cyberbullying Mitigation, National Science Foundation (EAGER), 2017-2021, \$200,000.

**Co-PI:** IBM-ILLINOIS Center for Cognitive Computing Systems Research (C3SR), 2016-2021.

**Co-PI:** Supporting Self-regulated Learning in Online Education via Automatically Personalized Interventions, Technology Innovation in Educational Research and Design Program, 2020-2021, \$14,997. PI: Nigel Bosch.

**Lead PI:** A Nuanced Model for Recognizing Levels of Conflict in Decision Making Using Natural Language Processing, Social and Behavioral Sciences Research Initiative (SBSRI) Small Grants Program, 2017-2020, \$12,400.

**Co-PI:** Interactive Technology Support for Patient Medication Self-Management (Round 2 funding), Jump Applied Research for Community Health through Engineering and Simulation (ARCHES), UIC/OSF Hospital, 2018-2019, \$52,224. PI: Daniel Morrow.

**Collaborator:** The Role of Gesture in Mathematics Learning: From Research to Practice. National Science Foundation, Science of Learning Collaborative Network, 2016-2019, \$747,903. PI: Susan Goldin-Meadow.

**Co-PI:** Understanding Learning Behavior Patterns in MOOCs to Support Early Interventions, Illinois Learning Sciences Design Initiative, 2016-2017, \$15,000. PI: Michelle Perry.

**Co-PI:** Interactive Technology Support for Patient Medication Self-management. Jump Applied Research for Community Health through Engineering and Simulation (ARCHES), UIC/OSF Hospital, 2016-2017, \$113,958. PI: Daniel Morrow.

## Publications

### *Peer-Reviewed Journal Articles*

1. Z. Zeng, and S. Bhat. (2022). Getting BART to Ride the Idiomatic Train: Learning to Represent Idiomatic Expressions. *Transactions of the Association for Computational Linguistics*, 11.
2. G.M. Henricks, M. Perry, and S. Bhat. (2021). Gender and Gendered Discourse in Two Online Science College Courses. *Computer-Based Learning in Context*,

- 3 (1), 1–16.
3. Z. Zeng, and S. Bhat. (2021). Idiomatic expression identification using semantic compatibility. *Transactions of the Association for Computational Linguistics*, 9, 1546–1562.
  4. S. Yoon and S. Bhat. (2018). A Comparison of Grammatical Proficiency Measures in the Automated Assessment of Spontaneous Speech. *Speech Communication*, Volume 99, May, 221–230.
  5. R. W. Crues, G. Henricks, M. Perry, S. Bhat, C. Anderson, N. Shaik and L. Angrave. (2018). How does Gender, Learning Goals, and Forum Participation Predict Persistence in a Computer Science MOOC? *ACM Transactions on Computing Education*, 18(4).
  6. K. Zechner, S. Yoon, S. Bhat, and C. Leong. (2017). Comparative Evaluation of Automated Scoring of Syntactic Competence of Non-native Speakers. *Computers in Human Behavior*, Volume 76, November, 672–682.
  7. S. Bhat, and S. Yoon. (2015). Automatic Assessment of Syntactic Complexity of Spontaneous Speech Scoring. *Speech Communication* (67), 42–57.
  8. R. Girju, B. Beamer, A. Rozovskaya, A. Fister, and S. Bhat. (2010). A Knowledge-rich Approach to Identifying Semantic Relations Between Nominals. *Information Processing & Management* 46(5), 589–610.

### *Published Conference Papers*

9. P. Hur, H. Lee, S. Bhat, and N. Bosch. (2022). Using Machine Learning Explainability Methods to Personalize Interventions for Students. In *Proceedings of the 15th International Conference on Educational Data Mining (EDM 2022)*. International Educational Data Mining Society.
10. W. Zhu and S. Bhat (2022). Slow Service”  
 →  
 ”GreatFood” : Enhancing Content Preservation in Unsupervised Text Style Transfer. In *15th In*
11. J. Zhou, Z. Zeng, H. Gong, and S. Bhat (2022) Idiomatic Expression Paraphrasing without Strong Supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36, No. 10, pp. 11774–11782.

12. J. Zhou, H. Gong, and S. Bhat, PIE: Parallel Idiomatic Expression Corpus for Idiomatic Sentence Generation and Paraphrasing. In *Proceedings of the 17th Workshop on Multiword Expressions (MWE 2021)*, pp. 33-48.
13. N. Prabhu, M. Perry, R. F. L. Azevedo, L. Angrave, and S. Bhat (2021). Study Partners Matter: Impacts on Inclusion and Outcomes. In *2021 ASEE Virtual Annual Conference Content Access*.
14. W. Zhu, and S Bhat (2021). Euphemistic Phrase Detection by Masked Language Model. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 163-168. 2021.
15. J. Zhou, and S. Bhat (2021). Paraphrase generation: A survey of the state of the art. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 5075-5086. 2021.
16. W. Zhu, H. Gong, R. Bansal, Z. Weinberg, N. Christin, G. Fanti, and S. Bhat (2021). Self-Supervised Euphemism Detection and Identification for Content Moderation. In *2021 IEEE Symposium on Security and Privacy (SP)*, pp. 229-246.
17. W. Zhu and S. Bhat (2021). Generate, Prune, Select: A Pipeline for Counterspeech Generation against Online Hate Speech. In *Findings of ACL 2021*.
18. D. Williams-Dobosz, R. Azevedo, A. Jeng, V. Thakkar, S. Bhat, N. Bosch, and M. Perry (2021). A Social Network Analysis of Online Engagement for College Students Traditionally Underrepresented in STEM. In *Proceedings of the 11th International Learning Analytics and Knowledge (LAK) Conference 2021*.
19. J. Zhou, and S. Bhat (2021). Modeling Consistency of Engagement Patterns in Online Courses. To appear in *Proceedings of the 11th International Learning Analytics and Knowledge (LAK) Conference 2021*.
20. H. Gong, A. Valido, K. Ingram, G. Fanti, S. Bhat, and D. Espelage (2021). Abusive Language Detection in Heterogeneous Contexts: Dataset Collection and the Role of Supervised Attention. In *Proceedings of the AAAI Conference on Artificial Intelligence 35(17)*, 14804–14812.
21. H. Gong, L. Song, and S. Bhat (2020). Rich Syntactic and Semantic Information Helps Unsupervised Text Style Transfer. In *Proceedings of the 13th International Conference on Natural Language Generation*, pp. 113–119.
22. T. Sakakini, J. Y. Lee, A. Duri, R. F.L. Azevedo, V. Sadauskas, K. Gu, S. Bhat, D. Morrow, J. Graumlich, S. Walayat, M. Hasegawa-Johnson, T. Huang, A. Willemsen-

- Dunlap, and D. Halpin (2020). Context-Aware Automatic Text Simplification of Health Materials in Low-Resource Domains. In *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis*, pp. 115–126.
23. H. Gong, S. Bhat, and P. Viswanath (2020). Enriching Word Embeddings with Temporal and Spatial Information. In *Proceedings of the 24th Conference on Computational Natural Language Learning*, pp. 1-11.
24. W. Zhu, and S. Bhat (2020), GRUEN for Evaluating Linguistic Quality of Generated Text. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 94–108.
25. W. Zhu, H. Gong, J. Shen, C. Zhang, J. Shang, S. Bhat, and J. Han (2020). FUSE: Multi-Faceted Set Expansion by Coherent Clustering of Skip-grams. In *Proceedings of ECML-PKDD 2020*.
26. D. Morrow, R. F. L. Azevedo, L. Sari, K. Gu, T. Sakakini, M. Hasegawa-Johnson, S. Bhat, J. Graumlich, T. Huang, A. Hariharan, Y. Shao, and E. Cox. (2020). Closing the Loop in Computer Agent/Patient Communication. In *Proceedings of the 2020 Human Factors and Ergonomics Society Annual Meeting, Chicago, IL*.
27. F. Brahman, N. Varghese, S. Bhat and S. Chaturvedi. (2020). Effective Forum Curation via Multi-task Learning. In *Proceedings of the Educational Data Mining Conference 2020*. (Short paper, 22% acceptance rate)
28. V. Jay, G. Henricks, C. Anderson, L. Angrave, N. Bosch, N. Shaik, D. Williams-Dobosz, S. Bhat, and M. Perry. (2020). Online Discussion Forum Help-Seeking Behaviors of Students Underrepresented in STEM. In *Proceedings of the 2020 International Conference of the Learning Sciences (ISLS)*.
29. G. Henricks, S. Bhat, and M. Perry, Gender and Gendered Discourse in Online STEM College Courses.(2020). In *Proceedings of the 2020 International Conference of the Learning Sciences*.
30. H. Gong, K. Gupta, A. Jain, and S. Bhat. (2020). Illinimet: Illinois system for metaphor detection with contextual and linguistic information. In *Proceedings of the Second Workshop on Figurative Language Processing*.
31. O. Anjum, H. Gong, S. Bhat, J. Xiong, and W. Hwu. (2019). PaRE: A Paper-Reviewer Matching Approach Using a Common Topic Space. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. (Long paper, 24% acceptance rate)

32. T. Sakakini, H. Gong, J. Y. Lee, R. Schloss, J. Xiong, and S. Bhat. (2019). Equipping Educational Applications with Domain Knowledge. In *Proceedings of the Workshop on Innovative Use of NLP for Building Educational Applications*. (Long paper)
33. R. F. L. Azevedo, D. Morrow, K. Gu, T. Huang, M. Hasegawa-Johnson, P. Soni, S. Tang, T. Sakakini, S. Bhat, A. Willemsen-Dunlap, and J. Graumlich. (2019). The Influence of Computer Agent Characteristics on User Preferences in Health Contexts. In *Proceedings of the 2019 Human Factors and Ergonomics Society Health Care Symposium*.
34. H. Gong, S. Bhat, L. Wu, J. Xiong and W. Hwu. (2019). Reinforcement Learning Based Text Style Transfer without Parallel Training Corpus, In *Proceedings of the 17th (2019) Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. (Long paper, 26% acceptance rate)
35. H. Gong, Y. Li, S. Bhat, and P. Viswanath. (2019). Context-Sensitive Malicious Spelling Error Correction. In *Proceedings of the Web Conference (WWW), 2019*. (Short paper, 18% acceptance rate)
36. Z. Zeng, S. Chaturvedi, S. Bhat, and D. Roth. (2019). DiAd: Domain Adaptation for Learning at Scale, In *Proceedings of the 9th International Learning Analytics and Knowledge (LAK) Conference 2019*. (Long paper, 32% acceptance rate)
37. H. Gong, J. Mu, S. Bhat, and P. Viswanath. (2018). Preposition Sense Disambiguation and Representation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. (Long paper, 26% acceptance rate)
38. R. F. L. Azevedo, T. Sakakini, J. Y. Lee, V. Sadauskas, K. Gu, Y. Zhang, D. Morrow, S. Bhat, M. Hasegawa-Johnson, T. S. Huang, A. Willemsen-Dunlap, D. J. Halpin, and J. Graumlich. (2018). Using Conversational Agents to Explain Medication Instructions to Older Adults. *Proceedings of the AMIA Annual Symposium*. San Francisco, CA: American Medical Informatics Association.
39. H. Gong, T. Sakakini, S. Bhat, and J. Xiong. (2018). Document Similarity for Texts of Varying Lengths via Hidden Topics. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. (Long paper, 25% acceptance rate)
40. R. W. Crues, N. Bosch, M. Perry, L. Angrave, N. Shaik, and S. Bhat. (2018). Refocusing the Lens on Engagement in MOOCs. In R. Luckin, K. R. Koedinger, & S. Klemmer (Eds.), *Proceedings of the 5th (2018) ACM Conference on Learning@Scale*. New York, NY: ACM. (Long paper, 22% acceptance rate)

41. R. W. Crues, N. Bosch, C. J. Anderson, M. Perry, S. Bhat, and N. Shaik. (2018). Who They Are and What They Want: Understanding the Reasons for MOOC Enrollment. In K. E. Boyer & M. V. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining (EDM 2018)*. International Educational Data Mining Society. (Long paper, 16% acceptance rate)
42. H. Gong, S. Bhat, and P. Viswanath. (2018) Embedding Syntax and Semantics of Prepositions via Tensor Decomposition. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. (Long paper, 32% acceptance rate)
43. N. Bosch, R. W. Crues, G. Henricks, M. Perry, L. Angrave, N. Shaik, S. Bhat, and C. J. Anderson. (2018). Modeling Key Differences in Underrepresented Students' Interactions with an Online STEM Course. In *Proceedings of the TechMindSociety'18 conference*. New York, NY: ACM. (27% acceptance rate)
44. R. F. L., Azevedo, K. Gu, Y. Zhang, V. Sadauskas, T. Sakakini, D. Morrow, M. Hasegawa-Johnson, T. S. Huang, S. Bhat, A. Willemsen-Dunlap, D. J., Halpin, J. Graumlich, and W. Schuh. (2017). Using Computer Agents to Explain Clinical Test Results. *AMIA Annual Symposium Proceedings*. Washington, DC: American Medical Informatics Association.
45. T. Sakakini, R. F. L. Azevedo, V. Sadauskas, K. Gu, Y. Zhang, S. Bhat, D. Morrow, M. Hasegawa-Johnson, T. S. Huang, A. Willemsen-Dunlap, D. J. Halpin, and J. Graumlich. (2017). Dr. Babel Fish: A Machine Translator to Simplify Providers' Language. In *AMIA Annual Symposium Proceedings*. Washington, DC: American Medical Informatics Association.
46. Z. Zeng, S. Chaturvedi and S. Bhat. (2017). Learner Affect Through the Looking Glass: Characterization and Detection of Confusion in Online Courses, *Proceedings of the Educational Data Mining Conference 2017*. (Short paper, 42% acceptance rate)
47. J. Mu, S. Bhat, and P. Viswanath. (2017). Representing Sentences as Low-Rank Subspaces. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 629-634. (Short paper, 19% acceptance rate)
48. T. Sakakini, S. Bhat, and P. Viswanath. (2017). MORSE: Semantic-ally Drive-n MORpheme SEgment-er, *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 552-561. (Long paper, 25% acceptance rate)
49. J. Mu, S. Bhat, and P. Viswanath. (2017). Geometry of Polysemy. In *Proceedings of*



*the International Conference on Learning Representation, ICLR.* (Long paper, 39% acceptance rate)

50. H. Gong, S. Bhat, and P. Viswanath. (2017). Geometry of Compositionality. In *Proceedings of Thirty-First AAAI Conference on Artificial Intelligence (AAAI)*. (Long paper, 25% acceptance rate)
51. S. Bhat, S. Yoon, and D. Napolitano. (2015). Automatic Detection of Grammatical Structures from Non-Native Speech. *Proceedings of the Workshop on Speech and Language Technology in Education, Interspeech*.
52. S. Bhat, P. Chinprutthiwong and M. Perry. (2015). Seeing the Instructor in Two Video Styles: Preferences and Patterns. *Proceedings of the 8th International Conference on Educational Data Mining (EDM)*, 305–312. (Long paper, 16% acceptance rate)
53. B. Amnueypornsakul and S. Bhat. (2014). Machine-guided Solution to Mathematical Word Problems. *Proceedings of the Pacific Asia Conference on Language, Information and Computing (PACLIC)*.
54. B. Amnueypornsakul, S. Bhat, P. Chinprutthiwong. (2014). Predicting Attrition Along the Way: The UIUC Model. *Proceedings of the Workshop on Analysis of Large Scale Social Interaction in MOOCs (EMNLP)*.
55. S. Bhat, H. Xue and S. Yoon. (2014). Shallow Analysis Based Assessment of Syntactic Complexity for Automated Speech Scoring. *Proceedings of the Annual meeting of the Association of Computational Linguistics (ACL)*. (Long paper, 26% acceptance rate)
56. S. Bhat, Statistical Stemming for Kannada. (2013). *Proceedings of the 4th Workshop on South and Southeast Asian Natural Language Processing, International Joint Conference on Natural Language Processing*.
57. S. Bhat, G. Herman. (2013). Student Perceptions of Differences in Visual Communication Mode for an Online Course in Engineering. *Frontiers in Education*.
58. S. Bhat. (2012). Morpheme Segmentation for Kannada Standing on the Shoulder of Giants. *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing, COLING*.
59. S. Yoon, S. Bhat. (2012). Assessment of ESL Learners' Syntactic Competence Based on Similarity Measures. *Proceedings of the Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*.

- 60. S. Yoon, S. Bhat and K. Zechner. (2012). Vocabulary Profile as a Measure of Vocabulary Sophistication. *Proceedings of 7th Workshop on Innovative Use of NLP for Building Educational Applications*.
- 61. S. Bhat, M. Hasegawa-Johnson and R. Sproat. (2010). Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech. *INTERSPEECH Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology*.
- 62. S. Bhat, R. Sproat, M. Hasegawa-Johnson and F. Davidson. (2010). Automatic Fluency Assessment Using Thin-Slices of Spontaneous Speech. *Language Technology Research Colloquium*.
- 63. S. Bhat, R. Sproat. (2009). Knowing the Unseen: Estimating Vocabulary Size over Unseen Samples. *Proceedings of the 47th Annual Meeting of the Association of Computational Linguistics (ACL)*, page 109–117.
- 64. S. Bhat, K. Church. (2008). Variable Selection for Ad Prediction. *Proceedings of the 2nd International Workshop on Data Mining and Audience intelligence For Advertising*.
- 65. B. Beamer, S. Bhat, B. Chee, A. Fister, A. Rozovskaya and R. Girju. (2007). UIUC: A Knowledge-rich Approach to Identifying Semantic Relations between Nominals. *Proceedings of the Semantic Evaluation Workshop (SemEval 2007)*.
- 66. S. Bhat and R. Sproat. (2007). Consistent Estimation of the Number of Unseen Elements. *Proceedings of the Fourth Midwest Computational Linguistics Colloquium*.

### *Conference Presentations*

- 67. S. Bhat, W. F. Bond, V. Thakkar, R. Ebert-Allen, P. Bhagchandani, T. Sakakini (2022). Automated Assessment of Written Patient Notes for Reliable Medical Learner Feedback. International Meeting for Simulation in Healthcare. Los Angeles, CA.
- 68. W. F. Bond, S. Bhat, V. Thakkar, T. Sakakini, P. Bhagchandani, R. Ebert-Allen (2021). Improving Feedback and Efficiency: Automated Grading of Post Simulation Written Chart Notes. Oral Presentation. University of Illinois Urbana-Champaign, Grainger College of Engineering, 8th Health Care Engineering Systems Symposium.