Partha Pratim Talukdar

Staff Research Scientist, Google Research, India
Associate Professor, Department of Computational and Data Science (CDS)
Indian Institute of Science, Bangalore 560012, India
+91-80-2293 2789, partha@google.com, ppt@iisc.ac.in, http://www.talukdar.net

RESEARCH INTERESTS	I am interested in Natural Language Processing, Machine Learning, and Knowledge Graphs. My recent research has focused on multilingual learning, learning in low-resource settings, graph-based learning, automatic knowledge harvesting from web-scale datasets, and temporal information processing.

EDUCATION

University of Pennsylvania, Philadelphia, USA   2004 - 2010
PhD, Computer and Information Science
Advisors: Fernando Pereira, Zack Ives, Mark Liberman
Committee: William Cohen, Aravind Joshi, Ben Taskar, Lyle Ungar

University of Pennsylvania, Philadelphia, USA   2004 - 2005
Master of Science and Engineering (MSE), Computer and Information Science

Birla Institute of Technology & Science (BITS), Pilani, India   1999 - 2003
B.E. (Hons.) Computer Science (with Distinction)

EXPERIENCE

Google Research, Bangalore, India   Apr 2020 - Present
Staff Research Scientist. Leading a team on Natural Language Understanding.

Indian Institute of Science (IISc), Bangalore, India
Department of Computational and Data Sciences (CDS)
Department of Computer Science and Automation (CSA) (by courtesy)
    Associate Professor (on leave)   Dec 2019 - Present
    Assistant Professor   Jul 2014 - Dec 2019

KENOME, Bangalore, India   Sep 2017 - Present
Founder and CEO. Kenome is a core AI company helping enterprises make sense of unstructured data using cutting-edge Machine Learning, NLP, and Knowledge Graphs.

Carnegie Mellon University, Pittsburgh, PA   May 2011 – May 2014
Postdoctoral Fellow, Machine Learning Department (Supervisor: Tom Mitchell)
Co-architect of the Never Ending Language Learning (NELL) project.

Microsoft Research Search Labs, Mountain View, CA   Mar 2010 – Mar 2011
Visiting Research Scientist. Worked on machine learning and structured data in Rakesh Agrawal's group.

Google Research, Mountain View, CA   May 2008 – Dec 2008
Research Intern. Worked with Marius Paşca and Fernando Pereira on minimally supervised information extraction at the web-scale.

Google Research, Mountain View, CA   Jun 2006 – Sep 2006
Research Intern. Worked with Thorsten Brants and other members of the Machine

Translation group.

Linguistic Data Consortium (LDC), Philadelphia, PA   Jan 2006 - May 2006
Research Fellow. Worked on the Less Commonly Taught Languages (LCTL) project, developing statistical tools for resource poor languages.

Google Research, Mountain View, CA   Jun 2005 - Sep 2005
Research Intern. Worked with Thorsten Brants on named entity extraction.

Hewlett-Packard (HP) Labs, India   Jul 2003 - Jun 2004
Worked on language processing and learning for building text-to-speech systems.

PUBLICATIONS **h-index: 41, i10-index: 78, Citations: 5460** [Google Scholar: http://goo.gl/8y1na]

Book

Amarnag Subramanya, Partha Pratim Talukdar. *Graph-based Semi-Supervised Learning.* Morgan & Claypool Publishers, 2014.

Thesis

Partha Pratim Talukdar. *Graph-based Weakly-Supervised Methods for Information Extraction & Integration.* PhD Thesis, University of Pennsylvania, Philadelphia, PA, 2010.

Refereed Journal Papers

1. Ashutosh Kumar, Kabir Ahuja, Raghuram Vadapalli, Partha Talukdar. **Syntax-guided Controlled Generation of Paraphrases**. *Transactions of the ACL (TACL) 2020*

2. Tom M. Mitchell, W. Cohen, E. Hruschka, **P. Talukdar**, B. Yang, J. Betteridge, A. Carlson, B. Dalvi, M. Gardner, B. Kisiel, J. Krishnamurthy, N. Lao, K. Mazaitis, T. Mohamed, N. Nakashole, E. Platanios, A. Ritter, M. Samadi, B. Settles, R. Wang, D. Wijaya, A. Gupta, X. Chen, A. Saparov, M. Greaves, J. Welling. **Never-Ending Learning**. *Communications of the ACM*, 61(5), pp. 103-115, May 2018.

3. X. Fu , K. Huang, E.E. Papalexakis, H. Song, **P. Talukdar**, N. D. Sidiropoulos, C. Faloutsos, and T. Mitchell. **Efficient and Distributed Generalized Canonical Correlations Analysis for Big Multiview Data**. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Sep. 2018.

4. Evangelos Papalexakis, Tom Mitchell, Nicholas Sidiropoulos, Christos Faloutsos, Partha Talukdar, Brian Murphy. **Turbo-SMT: Parallel Coupled sparse Matrix-Tensor Factorizations and applications**. *Statistical Analysis and Data Mining: The ASA Data Science Journal, 2016.*

5. Leila Wehbe, Brian Murphy, Partha Talukdar, Alona Fyshe, Aaditya Ramdas, Tom Mitchell. **Simultaneously Uncovering the Patterns of Brain Regions Involved in Different Story Reading Subprocesses**. *PLoS ONE 9(11): e112575. doi:10.1371/journal.pone.0112575*

6. Zhepeng Yan, Nan Zheng, Zachary Ives, Partha Pratim Talukdar, Cong Yu. **Active Learning in Keyword Search-based Data Integration**. *The VLDB Journal Special Issue on Best Papers of VLDB 2013*

7. Zhepeng Yan, Nan Zheng, Zack Ives, Partha Talukdar, Cong Yu. **Actively Soliciting Feedback for Query Answers in Keyword Search-Based Data Integration**. *Proceedings of the VLDB Endowment (PVLDB)*, Volume 6, Number 3, January 2013. [Presented at VLDB 2013. Invited to the special issue of the VLDB Journal with the "Best Papers of VLDB 2013".]

8. Fabian M. Suchanek, James Fan, Raphael Hoffmann, Sebastian Riedel, Partha Talukdar. **Advances in Automated Knowledge Base Construction**. *ACM SIGMOD Records Journal*

9. Todd J. Green, Grigoris Karvounarakis, Nicholas E. Taylor, Val Tannen, Partha Pratim Talukdar, Marie Jacob, Fernando Pereira. **The Orchestra Collaborative Data Sharing System**. *ACM SIGMOD Records Journal*, September 2008.

10. Partha Pratim Talukdar, Marie Jacob, Mohammad Salman Mehmood, Koby Crammer, Zack Ives, Fernando Pereira, Sudipto Guha. **Learning to Create Data-Integrating Queries**. *Proceedings of the VLDB Endowment (PVLDB)*, Volume 1, Number 1, August 2008.

Refereed Conference Papers [acceptance rates, if available, are marked in brackets]

1. Apoorv Saxena, Soumen Chakrabarti and Partha Talukdar. **Question Answering over Temporal Knowledge Graphs**. *ACL 2021*

2. Keshav Kolluru, Martin Rezk, Pat Verga, William W. Cohen, Partha Talukdar. **Multilingual Fact Linking**. *AKBC 2021*

3. Simran Khanuja, Melvin Johnson and Partha Talukdar. **MergeDistill: Merging Language Models using Pre-trained Distillation**. *Findings of ACL 2021*

4. Yash Khemchandani, Sarvesh Mehtani, Vaidehi Patil, Abhijeet Awasthi, Partha Talukdar and Sunita Sarawagi. **Exploiting Language Relatedness for Low Web-Resource Language Model Adaptation: An Indic Languages Study**. *ACL 2021*

5. Sawan Kumar, Partha Talukdar. Reordering Examples Helps during Priming-based Few-shot Learning. *Findings of ACL 2021*

6. Chandrahas, Partha Talukdar. **OKGIT: Open Knowledge Graph Link Prediction with Implicit Types**. *Findings of ACL 2021*

7. Sagar Gubbi Venkatesh, Anirban Biswas, Raviteja Upadrashta, Vikram Srinivasan, Partha Talukdar, Bharadwaj Amrutur. **Spatial Reasoning from Natural Language Instructions for Robot Manipulation**. *ICRA 2021*

8. Naganand Yadati, Tingran Gao, Shahab Asoodeh, Partha Talukdar, Anand Louis. **Graph Neural Networks for Soft Semi-Supervised Learning on Hypergraphs**. *PAKDD 2021*

9. Naganand Yadati, Vikram Nitin, Madhav Nimishakavi, Prateek Yadav, Anand Louis and Partha Talukdar. **NHP: Neural Hypergraph Link Prediction**. *CIKM 2020*.

10. Sawan Kumar, Partha Talukdar. **Natural Language Inference with Faithful Natural Language Explanations**. *ACL 2020*.

11. Apoorv Saxena, Aditay Tripathi and Partha Talukdar. Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings. *ACL 2020*.

12. Zhiqing Sun, Shikhar Vashishth, Soumya Sanyal, Partha Talukdar and Yiming Yang. A Re-evaluation of Knowledge Graph Completion Methods. *ACL 2020.* (Short Paper)

13. Shikhar Vashishth*, Soumya Sanyal*, Vikram Nitin, and Partha Talukdar. **Composition-based Multi-Relational Graph Convolutional Networks**. *International Conference on Learning Representation (ICLR 2020)*, Ethiopia.

14. Ekagra Ranjan, Soumya Sanyal, Partha Talukdar. **ASAP: Adaptive Structure Aware Pooling for Learning Hierarchical Graph Representations**. *AAAI 2020*, USA.

15. Shikhar Vashishth*, Soumya Sanyal*, Vikram Nitin, Nilesh Agrawal, Partha Talukdar. **InteractE: Improving Convolution-based Knowledge Graph Embeddings by Increasing Feature Interactions**. *AAAI 2020*, USA.

16. Vivek Gupta, Ankit Saw, Pegah Nokhiz, Praneeth Netrapalli, Piyush Rai, Partha Talukda. **P-SIF: Document Embeddings using Partition Averaging**. *AAAI 2020*, USA.

17. Naganand Yadati, Madhav Nimishakavi, Prateek Yadav, Vikram Nitin, Anand Louis, Partha Talukdar. **HyperGCN: A New Method of Training Graph Convolutional Networks on Hypergraphs**. *Neural Information Processing Conference (NeurIPS 2019)*, Canada.

18. Sawan Kumar, Sharmistha Jat, Karan Saxena and Partha Talukdar. **Zero-shot Word Sense Disambiguation using Sense Definition Embeddings**. *57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*, Italy. **[Recipient of Outstanding Paper Award]**

19. Sharmistha Jat, Hao Tang, Partha Talukdar and Tom Mitchell. **Relating Simple Sentence Representations in Deep Neural Networks and Brain**. *57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*, Italy.

20. Shikhar Vashishth, Manik Bhandari, Prateek Yadav, Piyush Rai, Chiranjib Bhattacharyya and Partha Talukdar. **Incorporating Syntactic and Semantic Information in Word Embeddings using Graph Convolutional Networks**. *57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*, Italy.

21. Sanket Shah, Anand Mishra, Naganand Yadati and Partha Talukdar. **KVQA: Knowledge-aware Visual Question Answering**. *Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, USA.

22. Sawan Kumar, Varsha Sreenivasan*, Partha Talukdar, Franco Pestilli, Devarajan Sridharan. **ReAl-LiFE: Accelerating the Discovery of Individualized Brain Connectomes on GPUs**. *Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, USA.

23. Swapnil Gupta, Sreyash Kenkre and Partha Talukdar. **CaRe: Open Knowledge Graph Embeddings**. *International Conference on Empirical Methods in NLP (EMNLP 2019)*, Hong Kong.

24. Ashutosh Kumar, Satwik Bhattamishra, Manik Bhandari and Partha Talukdar. **Submodular Optimization-based Diverse Paraphrasing and its Effectiveness in Data Augmentation**. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2019)*, USA.

25. Shikhar Vashishth, Prateek Yadav, Manik Bhandari, Partha Talukdar. **Confidence-based Graph Convolutional Networks for Semi-Supervised Learning**.

International Conference on Artificial Intelligence and Statistics (AISTATS 2019), Japan.

26. Prateek Yadav, Madhav Nimishakavi, Naganand Yadati, Shikhar Vashishth, Arun Rajkumar and Partha Talukdar. **Lovasz Convolutional Networks**. *International Conference on Artificial Intelligence and Statistics (AISTATS 2019)*, Japan.

27. Shib Sankar Dasgupta, Swayambhu Nath Ray and Partha Talukdar. **HyTE: Hyperplane-based Temporally aware Knowledge Graph Embedding**. *International Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, Belgium.

28. Swayambhu Nath Ray, Shib Sankar Dasgupta and Partha Talukdar. **AD3: Attentive Deep Document Dater**. *International Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, Belgium.

29. Shikhar Vashishth, Rishabh Joshi, Sai Suman Prayaga, Chiranjib Bhattacharyya and Partha Talukdar. **RESIDE: Improving Distantly-Supervised Neural Relation Extraction using Side Information**. *International Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, Belgium.

30. Madhav Nimishakavi, Bamdev Mishra, Manish Gupta and Partha Talukdar. **Inductive Framework for Multi-Aspect Streaming Tensor Completion with Side Information**. *27th International Conference on Information and Knowledge Management (CIKM 2018)*, Italy.(17%)

31. Shikhar Vashishth, Swayambhu Nath Ray, Shib Sankar Dasgupta and Partha Talukdar. **Dating Documents using Graph Convolution Networks**. *56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)*, Melbourne, Australia.

32. Chandrahas Dewangan, Aditya Sharma and Partha Talukdar. **Towards Understanding Geometry of Knowledge Graph Embeddings**. *56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)*, Melbourne, Australia.

33. Madhav Nimishakavi, Manish Gupta and Partha Talukdar. **Higher-order Relation Schema Induction using Tensor Factorization with Back-off and Aggregation**. *56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)*, Melbourne, Australia

34. Shikhar Vashishth, Prince Jain and Partha Talukdar. **CESI: Canonicalizing Open Knowledge Bases using Embeddings and Side Information**. *The Web Conference 2018 (WWW 2018)*, Lyon, France.

35. Priya Radhakrishnan, Partha Talukdar and Vasudeva Varma. **ELDEN: Improved Entity Linking using Densified Knowledge Graphs**. *NAACL 2018*, New Orleans, USA

36. Prakhar Ojha, Partha Talukdar. KGEval: Accuracy **Estimation of Automatically Constructed Knowledge Graphs**. *International Conference on Empirical Methods in NLP (EMNLP 2017)*, Copenhagen, Denmark.

37. Aditya Sharma, Zarana Parekh, Partha Talukdar. **Speeding up Reinforcement Learning-based Information Extraction Training using Asynchronous Methods**. *International Conference on Empirical Methods in NLP (EMNLP 2017)*, Copenhagen, Denmark [Short Paper].

38. Xiao Fu, Kejun Huang, Otilia Stretcu, Hyun Ah Song, Evangelos Papalexakis, Partha Talukdar, Tom Mitchell, Nicholas Sidiropoulos, Christos Faloutsos, Barnabas Poczos. **BRAINZOOM: High Resolution Reconstruction**

from Multi-modal Brain Signals. *SIAM International Conference on Data Mining (SDM 2017)*, Houston, USA

39. Robert Pienta, Minsuk (Brian) Kahng, Zhiyuan Lin, Jilles Vreeken, Partha Talukdar, James Abello, Ganesh Parameswaran, Duen Horng (Polo) Chau. **Facets: Adaptive Local Exploration of Large Graphs**. *SIAM International Conference on Data Mining (SDM 2017)*, Houston, USA

40. Madhav Nimishakavi, Uday Singh Saini, Partha Talukdar. **Relation Schema Induction using Tensor Factorization with Side Information**. *International Conference on Empirical Methods in NLP (EMNLP 2016)*, Austin, USA

41. Prakhar Ojha, Partha Talukdar. **Quality Estimation of Workers in Collaborative Tasks from Group-Level Feedback**. *AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2016)*, Austin, USA.

42. Danish, Yogesh Dahiya and Partha Talukdar. **Discovering Response-Eliciting Factors in Social Question-Answering: A Reddit Inspired Study**. *10th International AAAI Conference on Web and Social Media (ICWSM-16)*, Cologne, Germany. (17%)

43. Mehdi Samadi, Partha Talukdar, Manuela Veloso, Manuel Blum. **ClaimEval: Integrated and Flexible Framework for Claim Evaluation Using Credibility of Sources**. *30th AAAI Conference Conference on Artificial Intelligence (AAAI-16)*, Phoenix, USA.

44. Manjunath Hegde, Partha Talukdar. **An Entity-centric Approach for Overcoming Knowledge Graph Sparsity**. *International Conference on Empirical Methods in NLP (EMNLP 2015)*, Portugal (Short Paper)

45. Bhushan Kotnis, Pradeep Bansal, Partha Talukdar. **Knowledge Base Inference using Bridging Entities**. *International Conference on Empirical Methods in NLP (EMNLP 2015)*, Portugal (Short Paper)

46. Matt Gardner, Kejun Huang, Evangelos Papalexakis, Xiao Fu, Partha Talukdar, Christos Faloutsos, Nicholas Sidiropoulos, Tom Mitchell. **Translation Invariant Word Embeddings**. *International Conference on Empirical Methods in NLP (EMNLP 2015)*, Portugal (Short Paper)

47. Mehdi Samadi, Partha Talukdar, Manuela Veloso, Tom Mitchell. **AskWorld: Budget-Sensitive Query Evaluation for Knowledge-on-Demand**. *International Joint Conference on Artificial Intelligence (IJCAI 2015)*, Buenos Aires, Argentina.

48. Tom Mitchell, William Cohen, Estevam Hruschka, Partha Talukdar, Justin Betteridge, Andrew Carlson, Bhavana Dalvi, Matt Gardner, Bryan Kisiel, Jayant Krishnmurthy, Ni Lao, Kathryn Mazaitis, Tahir Mohammad, Ndapa Nakashole, Emmanouil Antonios Platanios, Alan Ritter, Mehdi Samadi, Burr Settles, Richard Wang, Derry Wijaya, Abhinav Gupta, Xinlei Chen, Abulhair Saparov, Malcolm Greaves and Joel Welling. **Never-Ending Learning**. *Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, Austin, USA. (26.7%)

49. Bhavana Dalvi, Einat Minkov, Partha Talukdar and William Cohen. **Automatic Gloss Finding for a Knowledge Base using Ontological Constraints**. *International Conference on Web Search and Data Mining (WSDM 2015)*, Shanghai, China. (16.4%)

50. Alona Fyshe, Leila Wehbe, Partha P. Talukdar, Brian Murphy and Tom M. Mitchell. **A Compositional and Interpretable Semantic Space**. *Conference of the North American Chapter of the ACL (NAACL 2015)*, Denver, USA. (26%)

51. K. Huang, N. Sidiropoulos, C. Faloutsos, E. Papalexakis, P. Talukdar, T. Mitchell. **Principled Neuro-Functional Connectivity Discovery**. *SIAM International Conference on Data Mining (SDM 2015)*, Vancouver, Canada.

52. Partha Pratim Talukdar, William Cohen. **Scaling Graph-based Semi Supervised Learning to Large Number of Labels Using Count-Min Sketch**. *17th International Conference on Artificial Intelligence and Statistics (AISTATS 2014)*, Reykjavik, Iceland. (35.8%)

53. Matt Gardner, Partha Talukdar, Jayant Krishnamurthy, and Tom Mitchell. **Incorporating Vector Space Similarity in Random Walk Inference over Knowledge Bases**. *International Conference on Empirical Methods in NLP (EMNLP 2014)*, Doha, Qatar. (26%)

54. Alona Fyshe, Partha Talukdar, Brian Murphy, Tom Mitchell. **Interpretable Semantic Vectors from a Joint Model of Brain- and Text-based Meaning**. *52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*, Baltimore, USA. (26.2%)

55. Evangelos Papalexakis, Alona Fyshe, Nicholas Sidiropoulos, Partha Talukdar, Tom Mitchell, Christos Faloutsos. **Good-Enough Brain Model: Challenges, Algorithms and Discoveries in Multi-Subject Experiments**. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2014)*, New York City, USA. (14.6%)

56. Evangelos Papalexakis, Tom Mitchell, Nicholas Sidiropoulos, Christos Faloutsos, Partha Talukdar, Brian Murphy. **Turbo-SMT: Accelerating Coupled Sparse Matrix-Tensor Factorizations by 200x**. *SIAM International Conference on Data Mining (SDM 2014)*, Philadelphia, USA. [Invited to the Statistical Analysis and Data Mining (SAM) Special Issue of "Best of SDM 2014".]

57. Alex Beutel, Abhimanu Kumar, Evangelos Papalexakis, Partha Talukdar, Christos Faloutsos, Eric Xing. **FlexiFaCT: Scalable Flexible Factorization of Coupled Tensors on Hadoop**. *SIAM International Conference on Data Mining (SDM 2014)*, Philadelphia, USA. (15.4%)

58. Matt Gardner, Partha Talukdar, Bryan Kisiel, Tom Mitchell. **Improving Learning and Inference in a Large Knowledge-base using Latent Syntactic Cues**. *International Conference on Empirical Methods in NLP (EMNLP 2013)*, Seattle, USA. [Short Paper]

59. Derry Wijaya, Partha Talukdar, Tom Mitchell. **PIDGIN: Ontology Alignment using Web Text as Interlingua**. *International Conference on Information and Knowledge Management (CIKM 2013)*, San Francisco, USA. (16%)

60. Alona Fyshe, Brian Murphy, Partha Talukdar and Tom Mitchell. **Documents and Dependencies: an Exploration of Vector Space Models for Semantic Composition**. *International Conference on Computational Natural Language Learning (CoNLL 2013)*, Sofia, Bulgaria. (25%)

61. Partha Pratim Talukdar, Derry Wijaya, Tom Mitchell. **Acquiring Temporal Constraints between Relations**. *International Conference on Information and Knowledge Management (CIKM 2012)*, Hawaii, USA. (13.4%)

62. Partha Pratim Talukdar, Derry Wijaya, Tom Mitchell. **Coupled Temporal Scoping of Relational Facts**. *International Conference on Web Search and Data Mining (WSDM 2012)*, Seattle, USA. (20.7%)

63. Brian Murphy, Partha Talukdar, Tom Mitchell. **Learning Effective and Interpretable Semantic Models using Non-Negative Sparse Embedding**. *International Conference on Computational Linguistics (COLING 2012)*, Mumbai, India. (27%)

64. Brian Murphy, Partha Pratim Talukdar, Tom Mitchell. **Selecting Corpus-Semantic Models for Neurolinguistic Decoding**. *Joint Conference on Lexical and Computational Semantics (StarSem) 2012*, Montreal, Canada.

65. Paramveer Dhillon, Partha Pratim Talukdar, Koby Crammer. **Metric Learning for Graph-based Domain Adaptation**. *International Conference on Computational Linguistics (COLING 2012)* [Short Paper], Mumbai, India.

66. Rakesh Agrawal, Ariel Fuxman, Anitha Kannan, John Shafer, Partha Pratim Talukdar. **Associating Structured Records To Text Documents**. *International World Wide Web Conference (WWW 2012) [Poster]*, Lyon, France.

67. Anitha Kannan, Partha Pratim Talukdar, Nikhil Rasiwasia, Qifa Ke. **Improving Product Classification Using Images**. *International Conference on Data Mining (ICDM 2011)*, Vancouver, Canada. (12.5%)

68. Anton Bakalov, Ariel Fuxman, Partha Pratim Talukdar, Soumen Chakrabarti. **SCAD: Collective Discovery of Attribute Values**. 20th *International World Wide Web Conference (WWW 2011)*, Hyderabad, India. (12.5%)

69. Partha Pratim Talukdar, Fernando Pereira. **Experiments in Graph-based Semi-Supervised Learning Methods for Class-Instance Acquisition**. *The 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, Uppsala, Sweden. (25%)

70. Paramveer Dhillon, Partha Pratim Talukdar, Koby Crammer. **Learning Better Data Representation using Inference-Driven Metric Learning**. *The 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)* [Short Paper], Uppsala, Sweden. (22%)

71. Partha Pratim Talukdar, Zack Ives, Fernando Pereira. **Automatically Incorporating New Sources in Keyword Search-Based Data Integration**. *ACM SIGMOD 2010*, Indianapolis, USA. (20.8%)

72. Partha Pratim Talukdar, Koby Crammer. **New Regularized Algorithms for Transductive Learning**. *European Conference on Machine Learning (ECML-PKDD) 2009*, Bled, Slovenia. (24.9%)

73. Z. Ives, C. Knoblock, S. Minton, M. Jacob, P. Talukdar, R. Tuchinda, J. L. Ambite, M. Muslea, C. Gazen. **Interactive Data Integration through Smart Copy and Paste**. *Conference on Innovative Database Research (CIDR) 2009*, Asilomar, USA. (27%)

74. Ted Sandler, John Blitzer, Partha Pratim Talukdar, Lyle H. Ungar. **Regularized Learning with Networks of Features**. *Advances in Neural information Processing Systems (NIPS) 2009*, Vancouver, Canada. (23.8%)

75. Partha Pratim Talukdar, Joseph Reisinger, Marius Pasca, Deepak Ravichandran, Rahul Bhagat, Fernando Pereira. **Weakly Supervised Acquisition of Labeled Class Instances using Graph Random Walks**. *Empirical Methods in Natural Language Processing (EMNLP) 2008*, Honolulu, USA. (30%)

76. Partha Pratim Talukdar, Marie Jacob, Mohammad Salman Mehmood, Koby Crammer, Zack Ives, Fernando Pereira, Sudipto Guha. **Learning to Create Data-Integrating Queries**. 34th *International Conference on Very Large Databases (VLDB 2008)*, Auckland, New Zealand. (16.5%)

77. Koby Crammer, Partha Pratim Talukdar, Fernando Pereira. **A Rate-Distortion One-Class Model and its Applications to Clustering**. *International Conference on Machine Learning (ICML) 2008*, Helsinki, Finland. (26.5%)

78. Koby Crammer, Mark Dredze, Kuzman Ganchev, Partha Pratim Talukdar and Steve Caroll. **Automatic Code Assignment to Medical Text**. *BioNLP 2007*, Prague. (29%)

79. Partha Pratim Talukdar, Thorsten Brants, Mark Liberman and Fernando Pereira. **A Context Pattern Induction Method for Named Entity Extraction**. *Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, New York City, 2006. (35%)

80. Deepa S.R., A.G. Ramakrishnan, Kalika Bali, Partha Pratim Talukdar, **Automatic Generation of Compound Word Lexicon for Hindi Speech Synthesis**, *LREC 2004*, Portugal.

81. N. Sridhar Krishna, Partha Pratim Talukdar, Kalika Bali, A.G. Ramakrishnan, **Duration Modeling for Hindi Text-to-Speech Synthesis**, *8th International Conference on Spoken Language Laguage Processing (ICSLP 2004)*, Jeju Island, Korea.

82. Satinder Singh, Partha Talukdar, Sridhar Krishna, Sandeep Manocha, Kalika Bali, Sitaram R.N.V., **Optimal Creation of Speech Databases for Indian Language Speech Technology**, *O-COCOSDA 2004*, New Delhi, India.

83. Sriram S., Partha Pratim Talukdar, Sameer Badaskar, Kalika Bali, A.G. Ramakrishnan, **Phonetic Distance Based Cross-lingual Search**, *International Conference on Natural Language Processing* 2004, Hyderabad India.

84. K. Panchapagesan, Partha Pratim Talukdar, N. Sridhar Krishna, Kalika Bali, A.G. Ramakrishnan, **Hindi Text Normalization**, *Fifth International Conference on Knowledge Based Computer Systems (KBCS 2004)*, Hyderabad India.

Preprints

- Simran Khanuja, Diksha Bansal, Sarvesh Mehtani, Savya Khosla, Atreyee Dey, Balaji Gopalan, Dilip Kumar Margam, Pooja Aggarwal, Rajiv Teja Nagipogu, Shachi Dave, Shruti Gupta, Subhash Chandra Bose Gali, Vish Subramanian, Partha Talukdar. **MuRIL: Multilingual Representations for Indian Languages**. *Preprint: arXiv:2103.10730* [Model]

- Sharmistha Jat, Erika J C Laing, Partha Talukdar, Tom Mitchell. **Spatio-temporal Characteristics of Noun and Verb Processing during Sentence Comprehension in the Brain**. *doi.org/10.1101/2020.06.22.163808*

Refereed Workshop & Symposium Papers

1. Srinivas Ravishankar, Chandrahas Dewangan and Partha Talukdar. **Revisiting Simple Neural Networks for Learning Representations of Knowledge Graphs**. *6th Workshop on Automated Knowledge Base Construction (AKBC) 2017.*

2. Sharmistha Jat, Siddhesh Khandelwal and Partha Talukdar. **Improving Distantly Supervised Relation Extraction using Word and Entity Based Attention**. *6th Workshop on Automated Knowledge Base Construction (AKBC) 2017*

3. Matt Gardner, Partha Talukdar and Tom Mitchell. **Combining Vector Space Embeddings with Symbolic Logical Inference over Open-Domain Text**. *AAAI Spring Symposium on Knowledge Representation and Reasoning*, Stanford, USA

4. Leila Wehbe, Partha Talukdar, Brian Murphy, Alona Fyshe, Gustavo Sudre, and Tom Mitchell. **Tracking Story Reading in the Brain**. *NIPS 2012 Workshop on Machine Learning and Interpretation in NeuroImaging (MLINI 2012)*, Lake Tahoe, USA.

5. Partha Pratim Talukdar, William Cohen. **Crowdsourced Comprehension: Predicting Prerequisite Structure in Wikipedia**. *HLT-NAACL 2012 Workshop on Innovative Use of NLP for Building Educational Applications (BEA7)*, Montreal, Canada.

6. Mark Dredze, Partha Pratim Talukdar, Koby Crammer. **Sequence Learning from Data with Multiple Labels**. *European Conference on Machine Learning (ECML-PKDD) 2009* workshop on *Learning from Multi-Label Data*, Bled, Slovenia.

7. Partha Pratim Talukdar, John Blitzer, Ted Sandler, Mark Dredze, Koby Crammer, Fernando Pereira. **DRASO: Declaratively Regularized Alternating Structural Optimization**. ICML 2008 *Workshop on Prior Knowledge for Text and Language Processing*, Helsinki, Finland.

8. Mark Dredze, John Blitzer, Partha Pratim Talukdar, Kuzman Ganchev, Joao Graca, and Fernando Pereira. **Frustratingly Hard Domain Adaptation for Dependency Parsing**. *CoNLL Shared Task Session of EMNLP-CoNLL 2007*, Prague.

9. Kedar Bellare, Partha Pratim Talukdar, Giridhar Kumaran, Fernando Pereira, Mark Liberman, Andrew McCallum and Mark Dredze. **Lightly-Supervised Attribute Extraction**. NIPS 2007 *Workshop on Machine Learning for Web Search*, Whistler, Canada.

10. Kalika Bali, A.G.Ramakrishnan, Partha Pratim Talukdar, N. Sridhar Krishna, **Towards the Development of a Hindi Speech Synthesis System**, *5th ISCA Speech Synthesis Workshop*, 2004, Carnegie Mellon University, USA.

Technical Reports

1. Paramveer Dhillon, Partha Pratim Talukdar, Koby Crammer. **Inference-Driven Metric Learning (IDML) for Graph Construction**. *University of Pennsylvania CIS Technical Report: MS-CIS-10-18*.

2. Partha Pratim Talukdar. **Topics in Graph Construction for Semi-Supervised Learning**. *University of Pennsylvania CIS Technical Report MS-CIS-09-13*.

| **HONORS AND AWARDS** | • [Outstanding Paper Award at ACL 2019, Italy](#) (One of five such awardees out of 1737 submissions) |

• IBM Faculty Award, 2016

• MSE Visiting Fellowship, University of Melbourne, Australia, 2016

• Google Focused Research Award, 2014, 2015, 2016 (By Invitation Only)

• Accenture Open Innovation Research Award, 2015

• Spotlight Talk Award, Machine Learning Symposium (2014), New York Academy of Sciences

• Graduate Fellowship, CIS Department, University of Pennsylvania

• CIDR 2009 Graduate Student Scholarship

• Best Poster Award, North East DB-IR Day (Fall 2008)

• Young Engineering Fellowship (YEFP) awarded by Indian Institute of Science (IISc) to a total of 25 students in India for the 2002-2003 year.

• Govt. of India (North Eastern Council) Scholarship for academic excellence

PATENT	Anitha Kannan, Partha Pratim Talukdar, Nikhil Rasiwasia, Qifa Ke, Rakesh Agrawal. **Accurate Text Classification Through Selective Use of Image Data**. United States Patent Application 20120314941.

SOFTWARE	**Junto: The Label Propagation Toolkit** (): This toolkit consists of standalone as well as Hadoop-based implementations of various graph-based semi-supervised learning (SSL) algorithms. [Downloaded 680+ times as of February 2014]

STUDENT ADVISING & MENTORING

• **PhD**

1. Madhav N.	Aug 2015 - Oct 2019
2. Shikhar Vashishth	Aug 2017 - Dec 2019
3. Chandrahas D.	Aug 2015 - Aug 2021
4. Sharmistha	Aug 2015 - Sep 2021
5. Ashutosh Kumar	Aug 2016 - Present
6. Y. Naganand	Aug 2016 - Mar 2021
7. Sawan Kumar	Aug 2017 - Present
8. Apoorv Umang	Aug 2018 - Present

• **Masters**

1. Aditya Sharma (MTech (Research))	2017 - Present
2. Chittersu Raghu	2019 - 2020
3. Aniruddha Bala	2019 - 2020
4. Swapnil Gupta	2018 - 2019
5. Vipul Rathore	2018 - 2019
6. Aniket Pramanick	2018 - 2019
7. Ishaan Sood	2018 - 2019
8. Rahul Chittimalla	2018 - 2019
9. Swayambhu Nath Ray	2017 - 2018
10. Shib Shankar Dasgupta	2017 - 2018
11. Prince Jain	2016 - 2017
12. Manjunath Hegde	2015 - 2016
13. Arabinda Moni	2015 - 2016
14. P. Ojha (MSc (Engg.))	2014 - 2017
15. Ranjitha Kulkarni (VLIS, Carnegie Mellon)	Fall 2012
16. Lingjuan Peng (VLIS, Carnegie Mellon)	Fall 2012
17. Shuai Wang (VLIS, Carnegie Mellon)	Fall 2012
18. Guo Chen (VLIS Masters, Carnegie Mellon)	Spring 2012
19. Yilun Cui (VLIS Masters, Carnegie Mellon)	Spring 2012
20. Liuyi Sun (VLIS Masters, Carnegie Mellon)	Spring 2012

• **Undergraduates and Interns**

1. Prateek Yadav (IISc Bangalore)	2017 - 2019

2. Arash Enayati, (CMU Qatar)	Summer 2012
3. Malcolm Greaves (Carnegie Mellon)	Fall 2011
4. Kunal Lala (University of Pennsylvania)	Summer 2007
5. Deepa S.R. (BITS, Pilani)	Spring 2004

TUTORIALS

- **Intl. Conf. on Data Sciences (CODS-COMAD 2020)**
  Co-Instructor
- **Intl. Conf. on Empirical Methods in NLP (EMNLP 2019)**
  Co-Instructor (421 registrants)
- Invited Tutorial, **Intl. Conf. on Machine Learning (ICML 2019)**
  Co-Instructor, along with Prof. Tom Mitchell (CMU) ($\sim$1000+ attendees)
- **Intl. Conf. on Knowledge Discovery and Data Mining (KDD 2018)**
  Instructor (700+ registrants)
- **Conference on Information and Knowledge Management (CIKM 2017)**
  Co-instructor (200 attendees)
- **European Summer School in Logic, Language and Info. (ESSLLI 2014)**
  Co-Instructor
- **Intl. Conf. on Acoustics, Speech, Signal Processing (ICASSP 2013)**
  Co-Instructor (Participants: 63)
- **Lisbon Machine Learning School (2012)**
  Instructor (Participants: 110)
- **Intl. Conf. of Association of Computational Linguists (ACL 2012)**
  Co-Instructor (Participants: 70)

TEACHING

- E1 246: Natural Language Understanding, IISc Bangalore — Spring 2019
- DS 222: Machine Learning with Large Datasets, IISc Bangalore — Fall 2018
- E1 246: Natural Language Understanding, IISc Bangalore — Spring 2018
- DS 222: Machine Learning with Large Datasets, IISc Bangalore — Fall 2017
- SE 256: Scalable Systems for Data Science, IISc Bangalore — Spring 2016
- E1 246: Natural Language Understanding, IISc Bangalore — Fall 2015
- SE 294: Data Analysis and Visualization, IISc Bangalore — Spring 2015
- SE 305: Topics in Web-scale Knowledge Harvesting, IISc Bangalore — Fall 2014
- 10-605: Machine Learning with Large Datasets, Carnegie Mellon — Spring 2012
  Guest Lecturer (Instructor: Prof. William Cohen)
- Certificate in College and University Teaching — 2010
  Center for Teaching and Learning (CTL), University of Pennsylvania
- CSE 112: Networked Life, University of Pennsylvania — Spring 2006
  Teaching Assistant for Prof. Michael Kearns
- CSE 140: Intro. to Cognitive Science, University of Pennsylvania — Fall 2005
  Teaching Assistant for Prof. Lyle Ungar and Prof. Virginia Richards
- BITS C481: Computer Networks, BITS Pilani, India — Fall 2002

GRANTS

- Principal Investigator, Shell sponsored project on "Machine Learning Approaches for Material Discovery and Development". INR 53,20,000 (2019-2019)
- Co-Principal Investigator, Siemens sponsored project on "Knowledge-aided Zero-shot Part Detection". INR 31,96,000 (2019-2019)
- Co-Principal Investigator, IMPRINT-II (SERB) sponsored project on "Conversation Agents for Urban Navigation". INR 68,20,000 (2019-2021)
- Principal Investigator, Siemens sponsored project on "Knowledge Graph Induction from Technical Literature". INR 55,15,440 (2018-2019)
- Co-Principal Investigator, Intel sponsored project on "Network Representation Learning". INR 54,00,000 (2018-2020).
- Principal Investigator, Indian-centric Never Ending Learning (INELL), Defense Research & Development Organization (DRDO). INR 107,00,000 (2017-2019)
- Principal Investigator, Google Focused Research Award on "Knowledgable Neural Conversation Model". USD 50,000 (2016-2017)
- Principal Investigator, Google Focused Research Award on "Towards A Unified and Interpretable Embedding of Knowledge and Text". USD 50,000 (2015-2016)
- Principal Investigator, Robert Bosch India sponsored project on "Predicate Discovery from Instructional Documents". INR 17,00,000 (2015-16)
- Principal Investigator, Microsoft Research India Virtual Centre 2016. USD 10,000 (2015-16)
- Principal Investigator, IBM Research Open Science Collaboration Project. USD 4,000 (2015-16)
- Co-Investigator of a Centre of Excellence in DNS Security funded by the Department of Electronics and Information Technology, Government of India. INR 437,00,000 (2015-2018)
- Principal Investigator of University Research Grant from Accenture on "Task-directed Knowledge Graph Expansion for Improved Cognitive Computing". USD 25,000 (2015-2016)
- Principal Investigator of Google Focused Research Award on "Goal-directed Ontology Extension for Natural Language Understanding". USD 50,000 (2014-2015)
- Contributed to preparation of the National Science Foundation (NSF) Big Data grant "Big Tensor Mining: Theory, Scalable Algorithms and Applications". PIs: Christos Faloutsos (CMU), Tom Mitchell (CMU). USD 894,892 (2012-2016)

INVITED TALKS

2021
Keynote Talk, IJCAI-21 Workshop on Applied Semantics Extraction and Analytics (ASEA)

2020
Carnegie Mellon University, IIT Patna, IIT Madras, Google Zurich, Ashoka University

2019
Goldman Sachs New York City, American Express, Yodlee, NXP Semiconductors, IIT Alumni Association Bangalore, UiPath, Goldman Sachs Bangalore, Ericsson, Mindtree, Webscience Conference at IIT Madras, Intuit, Dubai Police Symposium on AI and Policing, NVidia, JNU Delhi, Rakuten, ITechLaw Conference, Infosys.

2018

LG, CAIR DRDO, Cargill, ALEF Education, Samsung Research, Shell, Microsoft, Flipkart, Anthill Inside Conference, IIIT Hyderabad, Shell, Microsoft, GlobalEdge, JP Morgan Chase, Airbus, IIT Guwahati, ISI Kolkata, Pi Ventures, Google Bangalore, LG, Siemens.

2017

Google Research NYC, University of Melbourne, SAP, TedX NIT Warangal, Capital One, Adobe Research, IIT Bombay Alumni Association, National Law School, Khosla Labs, JP Morgan Chase, Accenture AI Labs, Siemens, Goldman Sachs, Sony.

Goal-directed Knowledge Graphs

IIT Kanpur	Jul 2016
Distinguished Lecture Series, IBM Research India	Jun 2016
Allen Institute for Artificial Intelligence (AI2), USA	Jun 2016

Graph-based Learning for Web-scale Knowledge Harvesting   Mar 2015

Indian Institute of Technology (IIT), Kharagpur

From Big Text to Big Knowledge

Intel India Academic Forum	Oct 2015
Indian Statistical Institute (ISI), India	Aug 2015
Ericsson Research, India	Feb 2015
Alumni Research Talk, BITS Pilani, India	Jan 2015
CSE Department, Indian Institute of Technology (IIT) Madras	Dec 2014
Google, India	Nov 2014

Knowledge Representation and Processing in Machines and Humans

Center for Neurosciences (CNS) Seminar, Indian Institute of Science	Dec 2014

Large-scale Learning over Graphs

TextGraphs 2014, Doha, Qatar	Oct 2014

Web-Scale Knowledge Harvesting

IBM Research, India	Aug 2014

Overcoming Knowledge Base Sparsity using Web Text as Interlingua

Google Research, Mountain View, USA	Apr 2014
SIGMOD Workshop on Automatic Creation and Curation of KBs (WACCK)	June 2014

Never-Ending Language Learning

Mid-Atlantic Student Colloquium on Speech, Language and Learning	Oct 2013
NIPS 2013 Workshop on Knowledge Extraction from Text	Dec 2013

Temporal Scoping of Relational Facts

NLP Lab, Simon Fraser University	May 2013

Tutorial on Graph-based Semi-Supervised Learning

Database Seminar, Carnegie Mellon University	Nov 2012

Coupled Temporal Scoping of Relational Facts

Google Pittsburgh	Feb 2012

Collective Discovery of Attribute Values

Information Systems Seminar, Tepper School of Business (CMU)	Oct 2011
Indian Institute of Technology (IIT), Hyderabad	Mar 2011

Graph-based Methods for Information Extraction & Integration

Machine Learning Lunch, Carnegie Mellon University	Oct 2011
Harvard University	Dec 2010
Powerset (Microsoft)	May 2010
Microsoft Research Search Labs	Apr 2010

Learning to Create Data-Integrating Queries

UC Berkeley Machine Learning Tea	Sep 2009

PROFESSIONAL SERVICES

Program Committee (PC) Chair
- International Conference on Data Sciences (CoDS)   2017

Action Editor/Area Chair/Senior Program Committee
- Action Editor, Transaction of the Association for Computational Linguistics (TACL)
  June 2020 - Present
- Automated Knowledge Base Construction (AKBC)   2021
- International Joint Conference on Artificial Intelligence (IJCAI)   2016
- Association for Computational Linguistics (ACL)   2016
- Empirical Methods in Natural Language Processing (EMNLP-CoNLL)   2012

Guest Editorial Board
- Information Processing and Management (Special Issue: Time and Information Retrieval)   2015

Program Committee Member/Conference Reviewing
- American Association of Artificial Intelligence (AAAI)   2010-11, 2013-18
- Association for Computational Linguistics (ACL)   2011-17, 2019
- Artificial Intelligence and Statistics (AISTATS)   2011-17
- Automated Knowledge Base Construction (AKBC)   2013-17, 2019
- International Conference on Natural Language Learning (CoNLL)   2013, 2016-18
- International Conference on Data Sciences (CODS)   2015-17
- Conf. of the European Association for Computational Linguistics (EACL)   2012, 2014, 2017
- Empirical Methods in Natural Language Processing (EMNLP) 2008, 2009, 2012-19
- North American Association for Comp. Linguists (NAACL)   2009-10, 2012-13, 2015-18
- International Conference on Computational Linguistics (COLING)   2016
- International Conference on Natural Language Processing (ICON)   2003, 2015-18
- International Conference on Machine Learning (ICML)   2013-18
- International Conference on Learning Representations (ICLR)   2018
- International Joint Conference on Artificial Intelligence (IJCAI)   2009, 2015, 2016
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2013-2018
- Neural Information Processing Systems (NeurIPS)   2008, 2009, 2011-18
- World Wide Web Conference (WWW)   2013-18
- Web Search and Data Mining (WSDM)   2015-18

Journal Reviewing
- ACM Computing Surveys
- ACM Transactions on the Web

- Artificial Intelligence
- IEEE Transactions on Knowledge on Data Engineering (TKDE)
- IEEE Transactions on Neural Networks and Learning Systems
- IEEE Letters of the Computer Society
- Information Processing and Management
- Journal of Artificial Intelligence Research (JAIR)
- Journal of Machine Learning Research (JMLR)
- Journal of Natural Language Learning (JNLE)
- Journal of Web Semantics
- Machine Learning
- Pattern Recognition
- PLoS ONE
- Transactions of the ACL (TACL)

Meeting Organization
- Workshop on Complex Graphs, AAAI 2021, Co-organizer
- Conference on Automated Knowledge Base Construction (AKBC 2019), Workshop Co-chair
- 2nd Conference on Data Sciences (CoDS 2015), Student Travel Awards Coordinator
- NeurIPS Workshop on Automatic Knowledge Base Construction (AKBC 2014), Co-organizer
- CIKM Workshop on Automatic Knowledge Base Construction (AKBC 2013), Co-organizer
- NAACL-HLT Workshop on Automatic Knowledge Base Construction, Co-organizer 2012
- North East Student Colloquium on Artificial Intelligence (NSECAI), Co-organizer 2007
- Computational Linguistic Lunch, University of Pennsylvania, Co-organizer 2005 – 2006

Other
- Grand Award Judge, Intel International Science and Engineering Fair (ISEF) 2012
- Graduate Student Engineering Group Representative, Univ. of Pennsylvania 2007 – 2008

MISCELLA-NEOUS	Volunteer, Western Pennsylvania Humane Society, Pittsburgh, PA	2011 - 2014
REFERENCES	Available upon request.	

Last updated on September 14, 2021